

J.-P. BENZÉCRI

F. BENZÉCRI

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

Les cahiers de l'analyse des données, tome 18, n° 2 (1993),
p. 143-176

http://www.numdam.org/item?id=CAD_1993__18_2_143_0

© Les cahiers de l'analyse des données, Dunod, 1993, 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/>

TYPLOGIE DE TEXTES GRECS D'APRÈS LES OCCURRENCES DE FORMES DE MOTS

[TEXTES GRECS 2]

J.-P. & F. BENZÉCRI

Le présent travail fait suite à [TEXTES GRECS], publié en 1991 dans CAD, Vol.XVI, n°1. Depuis cette date, d'une part, le corpus a été quelque peu étendu; et, d'autre part, compte tenu des travaux effectués sur d'autres langues, notamment en français, pour la recherche du thème des articles parus dans CAD (cf. [CAD XII-XVII (1...3)], Vol. XVIII, n°1, 1993), on a traité les dénombrements de mots pleins avec ceux de mots outil.

Nous décrivons d'abord l'état du corpus (§1) puis le choix des lexiques de formes dénombrées et l'écrêtage des tableaux (§2). Les résultats d'analyses et de classifications occupent les §§3 à 6. Ces §§ sont de longueur très inégale, parce que, entre de multiples études, on a choisi de n'en présenter en détail qu'une seule, celle fondée sur le dénombrement simultané de mots outils et de mots pleins (§5).

1 Le corpus des textes

Notre corpus est présentement distribué en 5 parties; avec, pour chacune de celles-ci une lettre symbolique qui sert d'initiale aux sigles de tous les chapitres ou fragments qu'elles comprennent.

1.1 Les Évangiles et les Actes : &

Sont compris dans cette partie, avec les quatre Évangiles, les Actes des Apôtres; soit:

St Matthieu, {&M01... &M28}; St Marc, {&m01... &m16}; St Luc, {&L01... &L24}; St Jean, {&J01... &J21}; Actes, {&a01... &a28}.

Au total, ≈ 63000 occurrences; en 117 chapitres, dont aucun n'a été découpé en fragments.

1.2 Les Épîtres et l'Apocalypse : π

Nous énumérons les textes, avec leurs sigles (en désignant éventuellement une Épître par le nom de ses destinataires ou de son auteur):

Romains, { $\pi Rm1... \pi R16$ }; 1-ère et 2-ème aux Corinthiens, { $\pi Ca1... \pi CaG$; $\pi Cb1... \pi CbD$ }; Galates, { $\pi GL1... \pi GL6$ }; Philippiens, { $\pi f\pi1... \pi f\pi4$ }; Colossiens, { $\pi CL1... \pi CL4$ }; 1-ère et 2-ème aux Thessaloniciens, { $\pi \emptyset a1... \pi \emptyset a5$; $\pi \emptyset b1... \pi \emptyset b3$ }; 1-ère et 2-ème à Timothée, { $\pi Ta1... \pi Ta6$; $\pi Tb1... \pi Tb4$ }; Tite, { $\pi Tt1... \pi Tt3$ }; Philémon, { πPhm }; Hébreux { $\pi Hb1... \pi HbD$ }; St Jacques, { $\pi Jk1... \pi Jk5$ }; 1-ère et 2-ème de St Pierre, { $\pi Pa1... \pi Pa5$; $\pi Pb1... \pi Pb3$ }; 1-2-3 de St Jean, { $\pi Ia1... \pi Ia5$; πIHb ; πIHc }; St Jude, { $\pi Io\Delta$ }; Apocalypse, { $\pi A\Omega 1... \pi A\Omega 9$, $\pi A\Omega a... \pi A\Omega o$ }.

Au total, \approx 55000 occurrences, en 143 chapitres, dont aucun n'a été découpé en fragments.

Ensemble, { $\&$, π } renferment tous les chapitres du Nouveau Testament.

1.3 Livres prophétiques de l'Ancien Testament : ¶

Ont été saisis trois livres:

Isaïe, { $\¶ I\Omega 1... \¶ I66$ }; Malachie, { $\¶ Ma1... \¶ Ma4$ }; Daniel, { $\¶ \partial 01... \pi \partial 14$ }.

Comme l'original hébraïque n'est pas connu pour la totalité du texte grec du livre de Daniel, on a considéré, outre la division en 14 chapitres, une division en 20 fragments, dont certains ne s'identifient pas à des chapitres, notamment: cantique des hébreux dans la fournaise, { $\¶ \partial 3x \¶ \partial 3y \¶ \partial 3z$ }; histoire de Susanne, { $\¶ \partial \Sigma x \¶ \partial \Sigma y$ }; Bel et le dragon, { $\¶ \partial \beta x \¶ \partial \beta y$ }.

Au total, \approx 41000 occurrences, en 84 chapitres; le livre de Daniel étant divisé soit en 14 cap., soit en 20 fragments.

1.4 Littérature d'époque hellénistique : §

Ces textes, plus ou moins illustres, ont déjà été présentés par nous dans [TEXTES GRECS]; nous les énumérons ici avec les sigles adoptés désormais:

Antiquités Judaïques de Flavius JOSÈPHE, Livre X: histoire d'une période où est rapporté le ministère du prophète Isaïe: 4 chapitres, { $\§ F1... \§ F4$ }; ou 7 fragments, découpés comme indiqué dans [TEXTES GRECS], p.66.

Antiquités Judaïques, Livre XVIII: débute par le recensement effectué par Quirinius en Judée, recensement pendant lequel le Christ naît à Béthléhem (Luc II); fait allusion à la Passion du Christ sous Ponce Pilate; se termine par la mort de Tibère: 8 cap., { $\§ f1... \§ f8$ }; ou 18 fragments, découpés comme indiqué dans [TEXTES GRECS], p.67.

Le Pasteur, d'HERMAS: écrit allégorique, d'inspiration chrétienne, daté du début du deuxième siècle; 25 cap., { $\§ Hr1... \§ Hr9$, $\§ HrA... \§ HrP$ }.

Les Éthiopiennes (ou Histoire des aventures de Théagène et Chariclée), d'HÉLIODORE: les 33 cap. du Livre I, { $\§ \acute{E}T1... \§ \acute{E}T9$, $\§ \acute{E}ta... \§ \acute{E}tx$ }.

Au total, les textes d'époque hellénistique que nous avons saisis comptent ≈28000 occurrences, en 70 chapitres.

1.5 Textes philosophiques d'ARISTOTE et de PLATON : @

ARISTOTE: Livre I de l'Éthique à Nicomaque, {@N01... @N13}; Livre I de la Politique, {@P01... @P13}; Livre I du traité "de l'Âme", 5 cap. {@A1... @A5}, ou 14 fragments; Traité "de l'Interprétation", (de ce bref traité, préliminaire à la logique, le titre doit s'entendre comme: 'de la valeur des jugements en tant qu'interprète de ce que l'âme connaît'), 14 cap., {@<A... @<N}, ou 11 fragments, {@<a... @<k}; Livre A (i.e. I) de la Métaphysique, 10 cap., {@MA1... @MA9, @MAX}, ou 14 fragments.

PLATON: le Phédon, découpé par nous en 40 cap., {@f01... @f40}. Dans ce dialogue Phédon, témoin de la dernière journée vécue par Socrate, rapporte à ses amis les ultimes enseignements du Maître; lequel, condamné à boire la ciguë, mourut à la tombée de la nuit.

Au total, pour ARISTOTE, ≈ 40000 occurrences, en 80 chapitres; pour PLATON, ≈ 22000 occurrences, en 40 chapitres.

Et, pour l'ensemble du corpus, ≈ 250000 occurrences, en 509 chapitres.

2 Choix du lexique et dénombrement des formes

En bref, on poursuit l'étude du corpus des textes grecs, en recourant à des méthodes appliquées pour l'analyse du contenu des articles de *CAD* (cf. [*CAD* XIII-XVII (2)]

2.1 Rappel: le lexique d4 de 96 formes de mots outil

Le lexique d4 a déjà servi au §4 de [TEXTES GRECS]; nous le croisons ici avec un ensemble de 538 fragments. Plus précisément, sont mis en supplément les trois textes {&M01 &L03 ¶∂3z}, i.e.: le chapitre 1 de l'Évangile selon St Matthieu, et le 3 de St Luc, dont chacun renferme une généalogie du Christ; et la fin de la prière des hébreux dans la fournaise, en Daniel; ces trois textes comportant la répétition de formules. De plus, comme au §4 de [TEXTES GRECS], sont en supplément les 5 formes {γὰρ εἶναι ἐκεῖνος μὲν τε}; et au dénombrement des formes on a adjoint celui des lettres et des mots {πlet πmot} (leurs nombres étant respectivement divisés par 100 et 10).

2.2 Lexique ∂v de 120 formes de mots outil

À la différence de ce qu'on a fait généralement jusqu'ici (notamment pour le lexique d4), on a conservé dans ∂v les formes, pronominales ou verbales, des premières et deuxième personnes du singulier et du pluriel. Voici pourquoi.

Un auteur recourt à la première et à la deuxième personne, non tant par propension stylistique ou pour se conformer aux exigences du sujet traité, que

parce qu'il a choisi de montrer, avec les paroles, ceux qui les disent ou les écoutent.

La première personne est celle qui parle; la deuxième personne est le destinataire. Dans le cas le plus simple, la première personne est l'auteur; et la deuxième personne est le lecteur - ou auditeur - de son ouvrage. Mais parfois, l'auteur met en scène, dans son texte, une ou plusieurs personnes auxquelles il attribue le rôle d'émettre ou de recevoir des paroles: depuis les répliques théâtrales proprement dites, jusqu'aux citations introduites dans un récit, voire un exposé doctrinal. D'ailleurs, à ce jeu, la morphologie des personnes 1 et 2 n'est pas indispensable: "le Seigneur appelle ses serviteurs" équivaut à "Je vous appelle".

Afin de saisir avec exactitude cette mise en scène, parfois complexe, il conviendrait d'analyser le contenu du texte, d'en dénombrer, dans chaque chapitre, les acteurs, comme on le fait pour une pièce de théâtre. À défaut, on attend du dénombrement des outils personnels qu'il saisisse quelque chose de la vivacité du jeu des paroles.

Dans le cours d'une œuvre, ce jeu peut grandement varier: les chapitres ou fragments n'en seront-ils pas rangés dans plusieurs classes distinctes? c'est ce qu'on doit observer.

2.3 L'ensemble des formes de mots pleins

Nous avons dénombré dans le corpus ≈ 18000 occurrences de $\kappa\alpha\iota$ = et; ≈ 4900 de la particule $\delta\epsilon$; ≈ 4700 de la préposition $\epsilon\nu$ = dans;... Au rang 45, ≈ 800 , vient la forme la plus fréquente qui ne soit pas un mot outil: $\theta\epsilon\omicron\nu$, cas génitif du nom 'Dieu'; puis, ≈ 700 , $\epsilon\iota\pi\epsilon\nu$, 'il dit' (passé); et, ≈ 500 , $\theta\epsilon\omicron\sigma$, nominatif du nom 'Dieu' et $\lambda\epsilon\upsilon\epsilon\iota$, 'il dit' (présent)... Ces résultats s'accordent avec la composition du corpus.

La distinction entre mot outil et mot plein n'est certes pas strictement définie: ainsi, la locution 'sur la face de' peut jouer le rôle de la préposition 'sur'; mais elle renferme un mot plein; et évoque la langue des prophètes: on a conservé $\pi\rho\acute{o}\sigma\omega\pi\omicron\nu$ dans le lexique des formes de plein. (De façon précise, la même locution $\epsilon\pi\grave{\iota}$ $\pi\rho\acute{o}\sigma\omega\pi\omicron\nu$ $\tau\eta\varsigma$ $\gamma\eta\varsigma$, 'sur la face de la terre', veut dire 'partout' dans Isaïe 23:17; tandis que dans Isaïe 49:23 il s'agit d'une prostration).

Sans nous arrêter à de telles nuances, nous avons distingué un ensemble $\Delta\Delta$ de 882 formes de mots pleins dont la fréquence, dans le corpus est ≥ 20 . De ces formes, les 535 qui ont une fréquence ≥ 25 constituent le lexique $\Delta\mu$.

Bien que ne songeant pas à analyser un tableau de contingence croisant l'ensemble I de nos 500 chapitres avec 500, voire 800 mots, nous avons construit les tableaux de contingence $I \times \Delta\Delta$ et $I \times \Delta\mu$ afin de choisir des

sous-lexiques de moindre effectif. Dans la suite, nous rendrons compte seulement des expériences faites à partir du lexique $\Delta\mu$.

2.4 Choix d'un lexique $\Delta\mu Da$ de 171 formes caractéristiques

L'article [CAD XII-XVII (2)] (in CAD Vol.XVIII, n°1, 1993) évoque l'usage du programme 'ranger'.

Le programme 'ranger', destiné à choisir un sous-lexique, commence par ce dialogue:

sont pris en compte dans chaque chapitre (colonne)
les mots dont le rang en fréquence est \leq à un seuil;
le seuil rs choisi (de 10 à 15) est:

seront pris en compte dans le lexique
les mots qui sortent dans un nombre de chapitres $>$ à un seuil;
le seuil qs choisi (de 4 à 8) est:

Ainsi, en deux questions, on choisit, d'une part, un seuil 'rs' pour délimiter les quelques mots pleins qui, pour chaque chapitre, sont considérés comme en caractérisant le contenu; et, d'autre part, un seuil 'qs' pour retenir, parmi les mots trouvés comme caractéristiques de plusieurs chapitres, ceux pouvant servir à la différenciation sémantique au sein du corpus.

On donne au sous-lexique obtenu un sigle se terminant par deux lettres qui indiquent les seuils utilisés: la 1-ère, de A à E, pour qs (de 4 à 8); la seconde, de a à f pour rs (de 10 à 15). Ainsi, pour le lexique $\Delta\mu Da$, qs=7 rs=10: on retient les mots de $\Delta\mu$ qui, dans plus de 7 chapitres, se rangent parmi les 10 plus fréquents. Le sous-lexique est d'autant plus restreint que rs est plus bas et qs plus élevé.

Le critère de choix ne tient pas compte de la répartition des chapitres en œuvres: ainsi, on peut conserver une forme parce qu'elle est caractéristique de plusieurs chapitres, bien que ceux-ci appartiennent tous à la même œuvre. Par exemple, (cf. §5.0) la forme $\alpha\pi\omicron\upsilon\alpha\sigma\iota\sigma$, négation, se trouve exclusivement dans le traité de l'Interprétation dont elle caractérise plusieurs chapitres. L'écarter compromettrait la reconnaissance du thème de la logique; cependant, il faudrait étendre notre corpus pour que ce thème fût présent dans plusieurs traités (cf. §7.2).

2.5 Écrêtage du tableau de correspondance entre chapitres et mots pleins

La forme $\alpha\upsilon\alpha\theta\omicron\nu$ (de l'adjectif bon) est, dans l'ordre alphabétique, la première du lexique $\Delta\mu Da$. Cette forme a 32 occurrences dans le dernier chapitre, @<N, du traité de l'Interprétation, où ARISTOTE étudie toutes les sortes de négations et de contraires sur l'exemple de {bon, non bon, mauvais}. On trouve encore, respectivement 6 et 4 occurrences dans des chapitres 6 et 7.

@N06 @N07, du Livre I de l'Éthique à Nicomaque (ce qui sied à un traité de morale). Dans les autres chapitres du corpus, la fréquence est, au plus de 3.

Dans une classification fondée sur un tableau non écrêté, @<N est séparé du reste du *de Interpretatione*, et s'agrège, à un niveau élevé, à une branche qui contient les chapitres de l'Éthique à Nicomaque.

On peut, de bien des manières, écrêter un tableau. Dans [CAD XII-XVII (2)], §4.4, on utilise une unique formule semi-linéaire: les $k(i,j)$ de 1 à 10 sont conservés; au-delà, ce qui dépasse de 10 est divisé par 5. Nous disposons maintenant d'un programme offrant d'écrêter valeurs brutes ou composantes des profils des lignes ou des colonnes, en choisissant un seuil... Nous dirons seulement comment on a procédé ici. Pour chaque mot (ou forme), considéré séparément, l'écrêtage porte sur les nombres d'occurrences qui sont le plus forts, le 2-ème et le 3-ème; en tenant compte du 4-ème. Soit $\{a_1, a_2, a_3, a_4\}$ ces nombres: a_1 est réduit à a_3 ; et plus généralement, une transformation linéaire réduit à $[a_4, a_3]$ l'intervalle $[a_4, a_1]$. Ainsi pour $\alpha\chi\alpha\theta\omicron\nu$, 32 devient 4; et 6 et 4 sont réduits entre 4 et 3: c'est-à-dire, pratiquement, à 3.

Réduction du lexique et écrêtage des nombres d'occurrences ne sont appliqués que pour les mots pleins. Quand, dans une analyse, figurent simultanément outils et mots pleins, on construit séparément deux tableaux de contingence; est écrêté celui afférent aux mots pleins; puis on le juxtapose à l'autre.

§3 Analyses fondées sur un lexique de mots outil

Quant à la clarté de l'interprétation, les résultats des analyses du §3 nous paraissent inférieurs à ceux obtenus au §5, avec un lexique comprenant des mots outil et des mots pleins. Il suffit donc d'exposer au §3 ce qui peut aider à découvrir expérimentalement les principes de la taxinomie statistique des textes. C'est pourquoi nous ne publions ici, avec les tableaux de valeurs propres, que des classifications commentées des chapitres ou fragments.

§3.1 Le lexique d4, 96 formes, sans aucune des premières et deuxièmes personnes

538 fragments × (96 formes de d4 + (π itr π mot))
 en supplément : (ψ ap ϵ ivna ϵ l ϵ xelvo σ μ ev τ e) ; (ϵ m01 ϵ L03 η 3z) ;
 trace : 7.471 e-1
 rang : 1 2 3 4 5 6 / 8 9 10
 lambda : 782 452 363 311 217 201 191 176 153 143 e-4
 taux : 1047 605 486 416 290 269 255 236 205 191 e-4
 cumul : 1047 1652 2138 2553 2844 3113 3368 3604 3809 4000 e-4

De toutes les analyses considérées dans le présent article, celle du §3.1 offre la trace la plus faible. En effet, l'hétérogénéité des thèmes et des genres se manifeste, dans la distribution des mots pleins, ou même dans celle des formes outil des premières et deuxièmes personnes, par un contraste beaucoup plus

c Partition en 24 classes : Sigles des fragments de la classe c (\$3.1)														
1030	εM02	εa20	εa18	εa13	εM28	εm16	εa11	εfv5	εa12	εa08	εa21	εa19	εfv1	εf4a
	εfv2	εET5	εHrA	εf7c	εa15	εa04	εa17	εI49	εa16	εa05	εa14	εI20	εa02	εa01
	εL24	εa28	εa23	εa25	εHb2	εHb5	εa24	Actes des Apôtres						
1024	εPa1	εTb2	εHb1	εR16	εθa5	εHbD	εTb4	εa26	εa27	εHb7	εf37	εHbA	εHb6	εHb9
	εTa6	εHbB	εHbC	εI37	εI36	εI52	εI21	εHrN	εHrO	εHrL	εETk	εHrK	εAe	
1041	εP12	εET1	εETs	εETr	εETt	εETc	εETa	εETg	εETf	εET8	εETb	εF4b	εFW3	εf7b
	εf6b	εF1a	εFW2	εF1b	εf35	εF4c	εET2	εETn	εETi	εET3	εf3b	εETm	εF4a	εET7
	εETv	εETu	εET9	εET1	εET4	εETo	εETj	εETe	εETx	εETd	εETh	εf8c	εf8e	εf8f
	εf8a	εf8b	εf7a	εf3a	εf4b	εf6a	εf8d	Éthiopiennes, flavius						
1043	εPb2	εIoΔ	εm04	εANa	εAN6	εHrI	εHrM	εHr5	εHb4	εJk5	εHrP	εM03	εJk1	εI53
	εL02	εL04	εL01	εL11	εM23	εL06	εM13	εM16	εM10	εL12	εJ02	εM22	εL07	εL10
	εL13	εa07	εL17	εM12	εCa8	εHb3	εM18	εM05	εM06	εHr4	εHr9	εL21	εM24	εm13
	εM11	εI55	εI32											
1039	εL14	εL15	εa06	εθβy	εθβx	εL20	εm12	εa03	εa22	εL19	εa10	εL09	εL16	εL18
	εL08	εM27	εm15	εa09	εL23	εm01	εm05	εm03	εm09	εm06	εm08	εM20	εM15	εm07
	εm11	εM21	εM04	εM14	εM08	εM10	εM19	εM25	εL05	εM09	εm02	εM17	εL22	εM26
	εm14	synoptiques: (εM εm εL) St Matthieu, St Marc, St Luc												
1000	εAN2	εAN3	εANm	εANh	εANg	εANd	εAN4	εAN8	εAN5	εANf	εAN7			
1028	εMa3	εθ2w	εθ1	εθ2a	εθ6	εI31	εI09	εI25	εθ9	εθ10	εMa2	εI39	εMa1	εθ12
	εθ8	εθ7	εANo	εAN1	εθ5	εθ4	εI12	εANk	εθ34	εθ11				
886	εHrF	εHrC	εHrD	εHrE										
1023	εθΣx	εθΣy	εANi	εANn	εANc	εI15	εI34	εANj						
986	εI14	εI10	εI19	εI29	εI07	εI08	εI11	εI04	εI22	εI28	εI30	εI17	εI35	εI27
	εI24	εI23	Isaie (début)											
1017	εHrG	εHr8	εHrH	εHr7	εHr6	εHb8	εI65	εHrJ	εI02	εHr1	εHr3	εI66	εI61	εANb
	εAN9	le Pasteur d'HERMAS												
1033	εI26	εI47	εI33	εθ3x	εI38	εI06	εI62	εI46	εI48	εI16	εI54	εI51	εI40	εI58
	εI60	εI18	εI01	εI45	εI44	εI43	εθ3y	εI60	εI03	εI63	εMa4	εI13	εI05	εI41
	εM07	εI56	εI57	εI42	εI59	Isaie								
83	εJ17	prière sacerdotale												St Jean
1042	εIHb	εJ15	εJ16	εJ10	εJ14	εJ05	εJ08	εIa5	εIa4	εIa3	εIa1	εIa2		
1021	εJ06	εJ07	εJ01	εJ09	εJ21	εJ20	εJ11	εJ04	εJ13	εJ12	εJ03	εJ18	εJ19	
1037	εEp6	εCL4	εPhm	εTb3	εfπ4	εTa4	εTt3	εR12	εCb3	εCb7	εθb2	εθa3	εfπ1	εPa5
	εTb1	εPb1	εθa1	εCb1	εIHc	εCb8	εCa5	εCaA	εRm3	εθa4	εRm1	εR15	εEp4	εCb9
	εCL2	εRm2	εCL3	εCa1	εCb4	εCa2	εθb1	εEp2	εEp3	εCL1	εEp1			
1047	εθa2	εCa6	εRm8	εRm7	εR13	εR14	εCaC	ε<j	εCaF	εJk2	εJk3	εR11	εGL1	εI50
	εJk4	εGL5	εCaB	εRm6	εGL4	εGL3	εRm9	εRm5	εR10	εfπ3	εRm4	εCbA	εCaD	εCaG
	εCaE	εCa7	εCa9	εCa3	εCbB	εPa2	εPb3	εEp5	εTa2	εCa4	εCb6	εTa5	εGL6	εTa3
	εCbC	εTt2	εTt1	εGL2	εPa4	εfπ2	εθb3	εTa1	εPa3	εCb5	εCbD	εCb2		

une subdivision 743 qui coïncide, à peu près, avec les chapitres 12 et 13, dont la suite confirmera l'originalité. Il est, en revanche, difficile de cerner, même approximativement, chacune des autres œuvres d'ARISTOTE, même en fragmentant les classes { 1036 1038... } qui les renferment.

Le reste du corpus constitue la classe 1068, partagée en 1067 et 1066.

De 1066, chacun des deux descendants a un contenu homogène: 1056 = { 1037 1047 } = Épîtres; 1058 = { 83 1042 1021 } = St Jean. Plus précisément,

on a, dans 1058, l'Évangile selon St Jean et la 1-ère Épître de cet Apôtre; mais non l'Apocalypse: structure déjà vue dans [TEXTES GRECS], confirmée dans la suite, et que nous considérerons au §5.3 dans l'usage même des mots.

La classe 1067 se partage en 1065 et 1063.

Sans reprendre les détails, portés sur l'arbre et le tableau, nous dirons que 1063, {100 1028 886 1023 986 1017 1033}, comprend, avec des livres prophétiques de l'Ancien Testament (dont ne manquent que 11 chapitres), l'Apocalypse (20/22); et la moitié du Pasteur d'HERMAS, texte qui prend parfois le ton de la révélation ou de la prophétie.

Enfin 1065 comprend, d'une part, dans 1059 {1043 1039}, la majeure partie des Évangiles synoptiques; et d'autre part, dans 1057 {1030 1024 1041}, les Actes des Apôtres, auxquels fait pendant la prose hellénistique des Éthiopiennes et des Antiquités Judaïques.

§3.2 Le lexique ∂v : 120 formes de mots outil, dont certaines des premières et deuxièmes personnes

509 chapitres X 120 mots de ∂v ; $\epsilon\iota\nu\alpha\iota$ en suppl.

trace :	1.229 e+0									
rang :	1	2	3	4	5	6	7	8	9	10
lambda :	1172	771	587	570	406	308	287	280	243	231 e-4
taux :	953	627	477	463	330	251	233	228	197	188 e-4
cumul :	953	1580	2058	2521	2851	3102	3335	3563	761	3949 e-4

Ainsi qu'on l'a annoncé au §3.1, la trace augmente de moitié avec l'introduction des outils des premières et deuxièmes personnes.

Quant à la classification des textes, la structure globale est semblable; mais la composition des classes se laisse parfois mieux résumer.

On retrouve ARISTOTE et PLATON associés dans 1010; où, comme précédemment, se signalent le *de Interpretatione* et le Phédon; de ce dialogue manquent 4 chapitres, dont {@f35 36 37}, qui rejoignent dans 992 la littérature hellénistique: on reverra, dans la suite, {@f34-37} séparés du reste du Phédon, de par leur contenu.

Comme au §3.1, les Épîtres, 1009, et St Jean, 1004, sont l'aîné et le benjamin d'un même nœud, 1014.

La branche 1012 {994 989 938 997} a été étiquetée: prophéties. Plus précisément, 994 se partage en deux classes inégales: dont l'une contient deux chapitres de Daniel; et l'autre 19 cap. de l'Apocalypse ($\pi\text{A}\Omega\dots$); et 989 contient 13 chapitres des Évangiles synoptiques {&M &m &L}.

La structure de la branche 1013 est nouvelle: d'une part, dans 992, la littérature hellénistique (i.e. Éthiopiennes et FLAVIUS); d'autre part, 1007, où une moitié du Pasteur (1005 = {796 991}) fait pendant à une grande classe 1003;

c	Partition en 20 classes : Sigles des chapitres de la classe c : \$3.2																																																																																																
992	€M01	\$ÉTB	\$ÉTC	\$ÉTA	\$ÉTH	\$ÉTO	\$ÉTE	\$ÉTF	\$ÉTP	\$ÉTG	\$ÉT7	\$ÉTX	\$F3	\$F4	\$F2	\$F1	@P12	\$ÉT6	\$ÉT1	\$ÉTK	\$ÉTS	\$ÉTn	\$ÉTr	\$ÉTT	\$ÉTu	\$ÉTi	\$ÉT1	\$ÉTj	\$ÉT2	\$ÉT3	@f35	@f37	@f36	@f16	\$f3	\$f4	\$f6	\$f7	\$f8	\$ÉT4	\$ÉTV	\$ÉTM	\$f1	\$f5	\$f2	Ethiopiennes et flavus																																																			
796	\$HrC	\$HrF	\$HrD	\$HrE	Hermas																																																																																												
991	€J17	\$HrB	\$HrG	\$Hr4	\$Hr5	\$Hr9	\$ÉTd	πTb4	¶I50	\$HrI	\$HrP																																																																																						
996	¶D03	€a25	€a24	€a23	€a22	€a26	¶I36	πHb1	¶I37	\$Hr2	\$Hr1	πIHb	πTa1	πTb1	πIHc	\$Hr3	πPhm	€a27	€a20	€a08	€a12	€a21	€a19	€a15	€a28	€m16	€M28	€a11	€a18	€a13	€a17	€a04	€a16	€a05	€a02	€a01	€a14	€L24	¶D13	€L02	€L04	€L01	€m08	€m06	€m03	€m05	€m01	€M14	€M08	€M09	€M04	¶D14	€L15	€m02	€L05	€a09	€a10	€L20	€m12	€L14	€L08	€L18	€L16	€L19	€m09	€L09	\$HrL	\$HrO	\$HrA	€M15	€m07	€a03	€m11	€M21	€a07	€M03	πJK1	€M22	€L07	€M02	€J02	€M16	€M12	€m04	€M13	€L17	€L13	€M19	€M20	€m10	€L23	€M27	€m15	€M17	€L22	€m14	€M26
994	(¶D05	¶D02	¶I39	¶D12	¶D06	¶I06	¶D08	¶D11	¶D04	¶D07)	(Daniel)+(AΩ)										€L03	πAΩ2	πAΩ3	πAΩ1	πAΩo	\$HrM	πAΩa	πAΩ6	€a06	¶D01	πAΩ8	πAΩg	πAΩh	πAΩ4	πAΩ5	πAΩf	πAΩc	πAΩk	πAΩm	πAΩd	\$Hr6	πAΩb	πHb8	πPb2	πAΩ7	πAΩ9																																																			
989	πfπ4	¶Ma1	¶Ma2	¶I55	¶I01	€M06	€M18	€M05	€L12	€M11	€L10	€M10	\$HrN	\$HrK	€L11	€M07	€L06	€L21	€m13	€M24																																																																													
938	πAΩn	πAΩi	¶I15	¶I34	πAΩj																																																																																												
997	¶I13	¶I35	¶I24	¶I23	¶I08	¶I28	¶I30	¶I20	¶I31	¶I25	¶I02	¶I27	¶I19	πTb3	¶I10	¶I18	¶I04	¶I11	¶I17	¶I12	¶I22	¶I07	€M25	¶I40	¶I42	¶I56	¶I44	¶I61	¶I21	¶I66	¶I65	¶I54	¶I32	¶I53	¶I59	¶I03	¶I09	¶I14	¶I05	¶I29	Isaïe																																																								
999	¶I16	¶I52	¶I26	¶I62	¶I58	¶I53	¶I63	¶I64	¶D09	(fin) d'Isaïe										¶I41	¶I46	¶I45	¶D10	¶I49	¶I38	¶I51	¶I48	¶I57	¶I43	\$Hr7	¶I47	¶I54	¶I60																																																																
966	€J01	€J21	€J09	€J20	€J11	€J04	€J18	€J19	€J03	€J12	€J06	€J07	St Jean																																																																																				
975	€J13	€J05	€J10	€J08	€J16	€J15	€J14	discours en St Jean																																																																																									
890	πIa1	πIa2	πIa4	πIa5	πIa3	1-ère épître de St Jean																																																																																											
987	πTt3	πGL2	πHb5	πHb2	πRm2	πCa2	πTb2	πHbA	πHb4	πTa6	πHbB	πHbC	πHb7	πHb9	πHb6	πCa5	πCa1	πCb5	πCb4	πTt2	πfπ3	πRm4	πCb3	πRm5	πRm8	€<B	πCa6	πRm3	πCaA	πCaD	πCaC	πRm6	πR12	πR14	πR13	πRm7	πRm9	πGL3	πCa8	πHb3	πR10	πR11	πJk3	πJk2	πCa3	πCaF	πGL4	πGL6	πCaB	πGL1	πGL5	\$ÉT5	πTa5	€<F	πTt1	πTa3	πCa7	πCa9	πCaE																																						
998	πEp2	πEp3	πEp4	πEp1	πCL1	πJk5	\$HrJ	\$HrH	πCL3	€M23	πfπ2	πTa2	πTa4	πEp6	πPa3	πJk4	πEp5	πPa4	πPb3	πPa5	πPa2	πCbA	πPa1	πCL2	πR16	πCb9	πPa4	πRm1	πR15	πfπ1	πCL4	πCaG	πCbD	πCb2	πCbC	πCbB	πCa4	πCb6	πIoA	¶I33	πHbD	πPb3	πPa5	πPb2	πPa2	πPb1	πCb8	πCb7	πPb1	πCb1	πPa1	πPa3																																													
993	€<N	€<D	€<C	€<G	€<J	€<I	€<L	€<M	de Interpretatione																																																																																								
971	\$ÉTW	\$ÉTq	@N09	@MA3	@P07	@P04	@P02	@P06	@P10	@P08	@P09	@P05	@MA6	@A2	@MA4	@MA5	@N01	@N04	@N11	@N10	@MA1	@N12	@P13	@N05	@N08	@N03	@N13	@MA2	@P11	@P03	@MAX	@MA8	Aristote...																																																																
995	@P01	€<K	@f14	@MA7	@f05	€<A	@f12	@A1	@MA9	@N06	@N02	@N07	@A3	@A4	@A5	@f32	€<E	€<H	@f11	@f31	@f29	@f24	Aristote et Platon																																																																										
961	@f09	@f10	@f06	@f08	@f15	@f17	@f07	@f04	@f03	@f34	@f28	@f33	@f25	@f18	@f22	@f20	@f13	@f19	@f26	@f27	@f23	@f38	@f21	@f30	@f39	@f40	\$ÉT9	\$ÉT8	\$Hr8	@f02	@f01	Phédon (28/40)																																																																	

laquelle associe, en bref, les Actes (23 chapitres sur 28), 996, aux synoptiques, 981 (45 chapitres; + 3 ch. dans 996).

```

992 _Éthiopiennes, flavius_ :13_ :15_ :16_ //_.
796 _Herm_ :05_ :07_ |
991 _Hermas..._ | |
996 _Actes_ :03_ | |
981 _Synopt_ | |
994 _Daniel Apoc_ :11_ :12_ | prophéties.. |
989 _&..._ :06_ | |
938 _ | |
997 _Isaie_ | |
999 _Isaie_ | |
966 _ :04_ St Jean_ :14_ |
975 _ | &v et émi |
890 _ :01_ |
987 _ :09_ Épitres_ |
998 _ |
993 _de Interpr_ :10_ Philosophie_ //_.
971 _ |
995 _ :02_ |
961 _Phédon_ |
    
```

§3.2 : CAH d'après 93 facteurs : arbre de la partition en 19 classes
 NB sur les noeuds de l'arbre les numéros {1000... 1016} sont écrits { :00... :16}

§4 Analyse fondée sur un lexique de formes de mots pleins

509 chapitres X 171 mots pleins de ΔμDa ; tableau écrété; sans suppl ; §4
 trace : 8.909e+0

rang	1	2	3	4	5	6	7	8	9	10
lambda	5014	3077	2734	2573	2356	1914	1697	1624	1456	1409 e-4
taux	563	345	307	289	264	215	190	182	163	158 e-4
cumul	563	908	1215	1504	1768	1983	2174	2356	2519	2677 e-4

Avec un lexique composé exclusivement de formes de mots pleins, à l'exclusion des outils, la trace approche de 9; alors qu'aux §§3 et 5, elle est, respectivement, inférieure à 1,5 ou 2,5.

```

992 _Actes_ :01_ :05:09:10:11_ :13_ :14_
979 _I $f_ | | | |
999 _Éthiopienn..._ | | | |
994 _Apocalypse..._ | | | |
977 _Isaie_ | | | |
803 _HERMAS_ :06_ | | | |
947 _HERMAS..._ | | | |
913 _Daniel_ | | | |
967 _ :08_ Épitres_ | | | |
997 _ | | | |
893 _St Jean_ :03_ Évangiles_ | | | |
987 _id_ :00_ | | | |
973 _&MmL_ | | | | synoptiques
998 _de Interpretatione_ :16_ Philosophie_
982 _Métaphysique_ :12_ :15_
959 _Éthique_ :02_ :07_
981 _Politique_ | |
985 _de l'Âme_ | |
996 _Phédon_ | |
    
```

§4 CAH d'après 93 facteurs
 partition en 19 classes

Quant à l'interprétation de la CAH, on reconnaît d'abord la classe 1016 de la philosophie: n'y manquent que 9 chapitres éparpillés ailleurs (dans 1014) et le

c Partition en 19 classes : Sigles des chapitres de la classe numéro c	
992	&M01 \$ÉTO &M28 &a12 &M02 &L01 &L02 &a08 &a19 ¶I63 ¶I11 ¶ð33 \$ÉTm &a28 &a09 &m16 &L24 &a05 &a04 &a23 &a17 &a13 πHbc πAΩb &a11 &a10 &a02 &a16 &a01 πAΩ4 &a03 πHbA \$HrL \$HrK πHb3 &M03 &L03 &m01 &a21 &a20 &a27 &m04 &M14 &m06
979	πAΩc ¶I38 SF2 πCb6 ¶I04 πHb1 &a07 ¶I07 ¶I08 ¶I37 πAΩm πPb3 ¶I23 ¶I20 ¶ð12 ¶I32 &M07 ¶I64 ¶ð09 ¶ð08 ¶ð04 ¶ð03 SF1 SF4 SF3 SF3 SF6 SF7
999	\$ÉT6 SF5 πCa7 πCaB \$Hr2 \$Hr1 \$HrF \$HrA \$Hr9 \$Hr5 \$HrJ \$HrP \$Hr8 πIa1 ¶I53 πHb9 πRm6 \$ÉTn \$ÉT7 \$ÉTa πCb7 \$ÉTv \$ÉTu \$ÉT3 ¶I28 \$ÉT1 πCbB \$ÉTh \$ÉT2 \$ÉTc @<C SF8 SF4 \$ÉTf \$ÉTr \$ÉTp \$ÉTd @f22 πIHc πR13 πCaD SF1 \$ÉTg \$ÉTe @<B \$ÉTx \$ÉTB @P01 \$ÉTw \$ÉTj @f01 @f12 @MA9 SF2 @<D @<E
994	@f36 @f37 ¶I34 @f34 @f35 ¶I24 \$ÉT5 ¶I46 ¶I26 πAΩ8 πAΩ9 πAΩf πAΩa πAΩh πAΩn πAΩj πAΩg πAΩl πAΩ7 πAΩk πAΩ5 πAΩ6
977	¶Ma2 πHb8 ¶Ma1 ¶Ma4 ¶I47 ¶Ma3 ¶I15 ¶I55 ¶I48 ¶I43 ¶I44 ¶I41 ¶I45 ¶I14 ¶I17 ¶I06 ¶I02 ¶ð10 ¶I18 ¶I13 ¶I29 ¶I65 ¶I54 ¶I56 ¶I59 ¶I21 ¶I31 ¶I66 ¶I49 πIoΔ πPb2 ¶I03 ¶I60 πJk5 ¶I05 ¶I50 ¶I30 ¶I58 ¶I40 ¶I42 ¶I51 ¶I35 ¶I01 ¶I52 ¶I25 ¶I10 ¶I33 ¶I62 ¶I09 ¶I27 ¶I12 ¶I22 ¶I19
803	\$HrE \$HrG \$HrC \$Hr4 \$HrB
947	¶ð07 πAΩd πAΩi \$HrM \$HrO \$HrN
913	¶ð01 ¶ð11 ¶ð02 ¶ð05 ¶ð06 ¶I36 ¶I39 ¶ð14
967	πHb7 &a24 πJk4 πCa9 πRm2 πRm7 πJk2 πRm3 πGL3 πGL2
997	πIHb πIa2 \$Hr7 πCb2 πIa5 πIa3 πCa8 πIa4 πR10 πRm4 πPa3 \$HrH \$Hr6 ¶I16 ¶I57 ¶I61 πRm9 πCaC πGL4 πRm8 πGL6 πGL5 πfπ3 πPa4 πEp2 πCL2 \$ÉTi πHb5 πHb2 πCa2 πCb3 πR15 &L04 &M04 &a06 πðb2 πJk1 &a18 &a14 &a15 πJk3 &m03 πCb8 πða4 πðb3 πCbD πða5 πCaG πAΩo πCaE πR11 πPa2 πCa4 πða2 πCaA πCa1 πTa5 &M23 πTb4 πGL1 πHbD πCaF πTa2 \$HrD πHb4 πTt2 πTt3 πCa3 πCb9 πR12 πEp3 πEp4 πða1 πðb1 πCb1 πRm1 πPal πHb6 \$Hr3 πTb2 πTb3 πTt1 πHbB πða3 πTa4 πTa6 \$ÉT4 πPa5 πEp1 πCL1 πCb4 πTb1 πRm5 πPb1 πTa1 πfπ1 πfπ2 πR14 πPhm πfπ4 πR16 πCa6 πCb5 \$HrI πCbA πCL4 πEp6 πCL3 πEp5
893	πCa5 &J17 \$ÉT9 &M06 &J05 &J08 &J10 &J15 &J14 &J16 discours en St Jean
987	&J18 &J19 &a25 &a26 &a22 &m15 &M27 &L23 πCbC πAΩ2 πAΩ3 &J04 &J02 &J11 &J20 &J13 &J21 &J09 &J07 &J12 &J01 &J06 &J03 St Jean...
973	&L20 &M22 &m12 &L06 &L16 &L14 &M15 &m07 &L08 &L05 &L18 &L07 &m02 &m05 &M09 &M10 &M05 &m10 &M19 &m09 &L12 &M18 &L21 &m13 &M24 &M13 &M16 &m08 &L11 &M12 &M21 &m11 &M26 &m14 &M11 &L15 &L17 &L10 &L13 &M25 &M20 &M08 &L19 &L22 &L09 &M17 Évangiles synoptiques
998	@f08 @<K @<N @<A @<J @<H @<G @<F @<L @<M @<I de Interpretatione
982	@f28 \$ÉTk @MA7 @f27 @N04 @MA2 @MA1 @MA4 @MA3 @MAX @MA8 @MA5 Métaphysique
959	@N11 @N09 @MA6 @N01 @N07 @N03 @N08 @N12 @N06 @N10 @N05 @N02 Éthique à N.
981	@P06 @P07 @P11 @P12 @P13 πTa3 @P08 \$ÉTs @P10 @P02 @P04 @P03 @P09 Polit.
985	@P05 @A1 @N13 @A2 @A5 @A3 @A4 de Anima
996	\$ÉT1 \$ÉTt @f21 @f25 @f19 @f23 @f24 @f15 @f20 @f16 @f17 @f06 @f07 @f05 @f09 @f14 \$ÉTq @f39 @f29 @f18 @f33 @f02 @f38 @f03 \$ÉT8 @f04 @f13 @f32 @f30 @f31 @f26 @f40 @f11 @f10 @f : 30/40

bloc {@f34-37},(du Phédon,) qui dans 994 va avec 13 chapitres de l'Apocalypse (AΩ).

Outre le *de Interpretatione*, ≈998, chacune des œuvres d'ARISTOTE (plus exactement les Livres séparés que nous en avons saisis), s'identifie approximativement à une subdivision de 1012: Métaphysique ≈ 982; Éthique à Nicomaque ≈ 959; Politique ≈ 981; de l'Âme ≈ 985.

Se détachent ensuite les classes 1003, Évangiles; et 1008, {967 997}, Épîtres.

Dans 1003, {893 987 973}, St Jean et les trois autres Évangiles (dits synoptiques, du fait du parallélisme qui les unit), sont très proches, à la différence de ce qui est le cas quand on considère l'usage des mots outil: en effet, en dehors des chapitres {&J14-16, &J17}, discours après la Cène et Prière sacerdotale (893), l'Évangile selon St Jean se distingue des synoptiques par la forme stylistique plus que par le contenu (cf. infra, §5.3).

Reste la branche 1011. Les subdivisions en sont notées: Daniel ≈ 913; HERMAS ≈ 1006; Isaïe ≈ 977; Apocalypse ≈ 994; Éthiopiennes ≈ 999; Actes ≈ 992. Mais à ces dénominations, il y a de nombreuses exceptions: en particulier, les chapitres des Antiquités Judaïques, de Flavius Josèphe, sont partagés entre 999 (Éthiopiennes) et une classe 979 où sont des chapitres de Daniel, d'Isaïe, de l'Apocalypse... Et les récits de la Nativité {&M01 &M02 &L01 &L02} et de la Résurrection {&M28 &m16 &L24}, issus des synoptiques, sont dans 992, avec (19/28) des chapitres des Actes.

§5 Analyses prenant en compte simultanément mots pleins et mots vides

On se bornera (cf. §2) à considérer un tableau de base croisant l'ensemble I des 509 chapitres du corpus avec 291 formes: ∂v , 120 formes de mots outil, et $\Delta\mu Da$, 171 de mots pleins; tableau construit en juxtaposant le tableau de contingence $I \times \partial v$, pris tel quel, avec le tableau $I \times \Delta\mu Da$ écrêté (cf. §§2.4-2.5).

Nous ne présenterons en détail qu'une analyse (§5.1) et les classifications afférentes (§§5.2, 5.3). Mais nous citerons des essais qui ont orienté notre choix et dont certaines conclusions sont à retenir (§5.0).

§5.0 Analyse préliminaire

Dans l'analyse considérée ici, sont en principal les 271 formes et la totalité des 509 chapitres.

509 chapitres \times 291 mots de ($\Delta\mu Da$, ∂v) ; écrêtage pour les mots pleins
 trace : 2.236e+0
 rang : 1 2 3 4 5 6... 10... 15... 25... 50.. 93...
 lambda: 1694 901 819 690 533 426... 327... 234... 175... 112.. 70 e-4
 taux : 758 403 366 308 239 190... 146... 105... 78... 50.. 31 e-4
 cumul : 758 1160 1527 1835 2074 2264... 2913... 3498... 4383... 5902.. 7579 e-4

La trace et les premières valeurs propres sont à un niveau intermédiaire entre ceux des §§3 et 4.

Prenons l'exemple de l'axe 1: dans toutes les analyses, son interprétation est la même, il oppose les textes philosophiques du corpus (les plus éloignés de l'origine) aux textes des Évangiles et des prophéties de l'Ancien Testament; les Épîtres et la littérature hellénistique étant au voisinage de l'origine. Une forme déterminée, qu'il s'agisse d'un mot outil ou d'un mot plein a, sur I, le même profil dans toutes les analyses où elle figure; et la distance du chi² change peu, les modifications du lexique respectant à peu près les poids relatifs des chapitres.

On a donc un nuage J des points figuratifs des formes (que celles-ci soient en principal ou en supplément) qui est quasi invariable, avec notamment, le même axe principal d'allongement et mêmes coordonnées sur celui-ci (ce que l'on peut vérifier en comparant les valeurs de $F1(j)$ issues des diverses analyses). Les variations de la trace et des valeurs propres résultent donc essentiellement des variations de masse des j.

En bref, la masse, m_j , des mots outil retenus est 7 fois celle, m_p , des mots pleins (après écrêtage): l'inertie des deux nuages pris ensemble est donc (7/8) de celle du nuage des mots vides (considérés seuls) + (1/8) de celle des mots pleins. Voici ce que l'on trouve en appliquant cette formule pour calculer l'inertie de l'analyse du §5.0 en fonction des résultats des §§3.2 et 4:

$$\text{trace} : 2236 =? (7/8) \cdot 1229 + (1/8) \cdot 8909 = 2189 ;$$

$$v.pr1 : 169 =? (7/8) \cdot 117 + (1/8) \cdot 501 = 155 ;$$

L'ordre de grandeur est bien compris; et le faible déficit s'explique en partie parce que $\epsilon\upsilon\alpha\upsilon$ est en supplément au §3.2 et non au §5.0.

À l'axe 1, l'infinitif $\epsilon\upsilon\alpha\upsilon$ du verbe être, apporte une contribution relative $CTR1 = 111 (1/9)$. Cette forme, se rencontrant principalement, dans le corpus, en dehors des textes bibliques, peut masquer des gradations continues que nous désirons observer: au §5.1, comme aux §§3.2 et 3.1, elle sera mise en supplément.

Sur le demi axe ($F5 > 0$), se détachent $\{ @<L, @<M \}$, chapitres XII et XIII du traité "de l'Interprétation" avec une contribution totale $CTR5 = 229$. S'y associent, entre autres, les formes $\{ \alpha\lambda\eta\theta\epsilon\sigma, \alpha\pi\omicron\upsilon\alpha\sigma\iota\sigma \}$: vrai, négation. Ce dernier mot est caractéristique du traité: le mettre en supplément nuirait à la reconnaissance du thème général de la logique... Dans un corpus dix fois plus étendu, on pourrait avoir plusieurs traités de logique: et prendre, sans hésiter, les mots trouvés comme caractéristiques de chapitres de plusieurs de ceux-ci. On a conservé $\alpha\pi\omicron\upsilon\alpha\sigma\iota\sigma$ en principal dans l'analyse du §5.1; mais sont en supplément les chapitres $\{ @<L, @<M \}$ (qui forment, à la CAH, une classe isolée).

Sur le demi axe ($F5 > 0$) $\epsilon\varphi\eta$, 'dit-il', donne $CTR5 = 135$; cette forme sera en supplément, ainsi que la particule $\tau\epsilon$ ($F5 < 0$; $CTR5 = 54$).

La question se pose du nombre des facteurs à utiliser pour la CAH. Prendre 93 facteurs (nombre maximum prévu dans le format actuel de nos fichiers) équivaut à accepter, à peu près, telle quelle, la distance distributionnelle; se limiter à quelques facteurs, le nombre en étant choisi d'après la décroissance des valeurs propres, réalise un filtrage qui pourrait être utile. On comparera ici les classifications des chapitres réalisées dans l'espace engendré par 12 ou par 93 axes.

Pour l'interprétation, 93 semble préférable: ainsi, notamment, les 5 chapitres $\{\pi Ia1 \dots \pi Ia5\}$ de la première de St Jean sont agrégés avec les chapitres de l'Évangile selon St Jean (le second excepté); ce qu'on a déjà souvent trouvé et que la classification avec 12 facteurs ne donne pas. Mais il est peu sûr de poursuivre des essais, sans critère numérique, jusqu'à ce que sortent de bons résultats.

c	1017	1016	1015	1014	1013	1012	1011	1010	1009	1008	1007	1006	1005	1004
car	509	401	295	211	108	94	141	84	14	102	50	106	51	70
T	1992	947	743	585	542	505	357	300	287	286	252	185	128	128

ci-dessus : avec 12 axes utilisés pour la CAH : brusque dénivellation quand le nombre des classes correspond à un simplexe dans l'espace retenu, à 12 dimensions;

ci-dessous: avec 93 axes utilisés pour la CAH : décroissance régulière de T (après le nœud 1 = axe1: COD1 = 931 CTD1 = 751)

c	1017	1016	1015	1014	1013	1012	1011	1010	1009	1008	1007	1006	1005	1004
car	509	369	240	140	47	93	118	122	14	41	81	129	93	55
T	806	357	326	264	239	208	207	168	145	119	116	116	116	103

Considérons, dans les deux CAH, la suite des taux d'inertie, T, afférents aux nœuds les plus hauts. Avec 12 facteurs, on a une chute rapide entre les nœuds 11 et 13, comme pour un nuage assimilé, schématiquement, à un simplexe à 13 sommets; (ce qui, dans l'espace de dimension 12, est l'analogue de ce qu'est un tétraèdre pour l'espace de dimension 3). Or de ce schéma, rien ne subsiste pour la CAH avec 93 facteurs, où les niveaux des nœuds décroissent régulièrement (après la dénivellation du premier au second). Ceci suggère que le schéma à 13 pointes est un artefact créé par la projection sur l'espace des axes 1 à 12.

5.1 Analyse principale avec mots outil et mots pleins

5.1.0 Décroissance des valeurs propres; choix des graphiques

On voit sur le tableau que la 1-ère valeur propre est nettement détachée des deux suivantes; elles-mêmes séparées de la 4-ème, comme celle-ci l'est de la 5-ème; la décroissance se faisant ensuite par degrés. Nous publions les plans (1,2) et (1,3), avec, pour chacun, deux images: d'une part, un nuage des

509 chapitres x 291 mots de {ΑμDa, δv} ; écrêtage pour les mots pleins ;
 en supplément : {Εψη, ΕΙvαι, τε} ; {<L, <M} CTR5 ;
 trace : 2.174e+0
 rang : 1 2 3 4 5 6.. 10.. 15. 25... 50... 93...
 lambda: 1479 892 814 666 446 381. 291. 229. 171.. 111 70 e-4
 taux : 681 410 374 307 205 175 134.. 105. 79.. 51... 32 e-4
 cumul : 681 1091 1465 1772 1977 2152.. 2780... 3357... 4246... 5800.. 7529 e-4

chapitres, où chacun est désigné par l'initiale de son sigle, ce qui montre, approximativement la distribution des genres de textes; d'autre part, les nuages des centres des classes des partitions retenues pour I et J, les nuages des chapitres n'étant alors figurés que par des points; ce qui suffit à donner l'échelle relative des coordonnées des individus et des centres. En effet, avec un grand nombre de chapitres et de mots, qui tous méritent de retenir l'attention, l'interprétation de l'analyse factorielle gagne à anticiper sur les résultats des classifications. On reconnaît les classes de mots à ce que leurs sigles sont formés de la lettre j, suivie d'un numéro de l'ordre de 500; tandis que pour les chapitres, la lettre est i, et le numéro voisin de 1000 (et toujours >700).

5.1.1 Le premier axe

Sur le demi-axe (F1<0) se détache l'ensemble des textes d'Aristote, ainsi que la quasi-totalité du Phédon de Platon: sont seuls assez proches de l'origine les chapitres {@f34... 37}=753, où Platon décrit, comme par une géologie fabuleuse, les lieux que hantent les âmes après la mort.

À ce niveau de l'axe, se place également la littérature Hellénistique: 936, 904, 970, les Antiquités Judaïques de Flavius, le roman des Éthiopiennes, le Pasteur d'HERMAS; avec quelques chapitres des Épîtres: 944, 980.

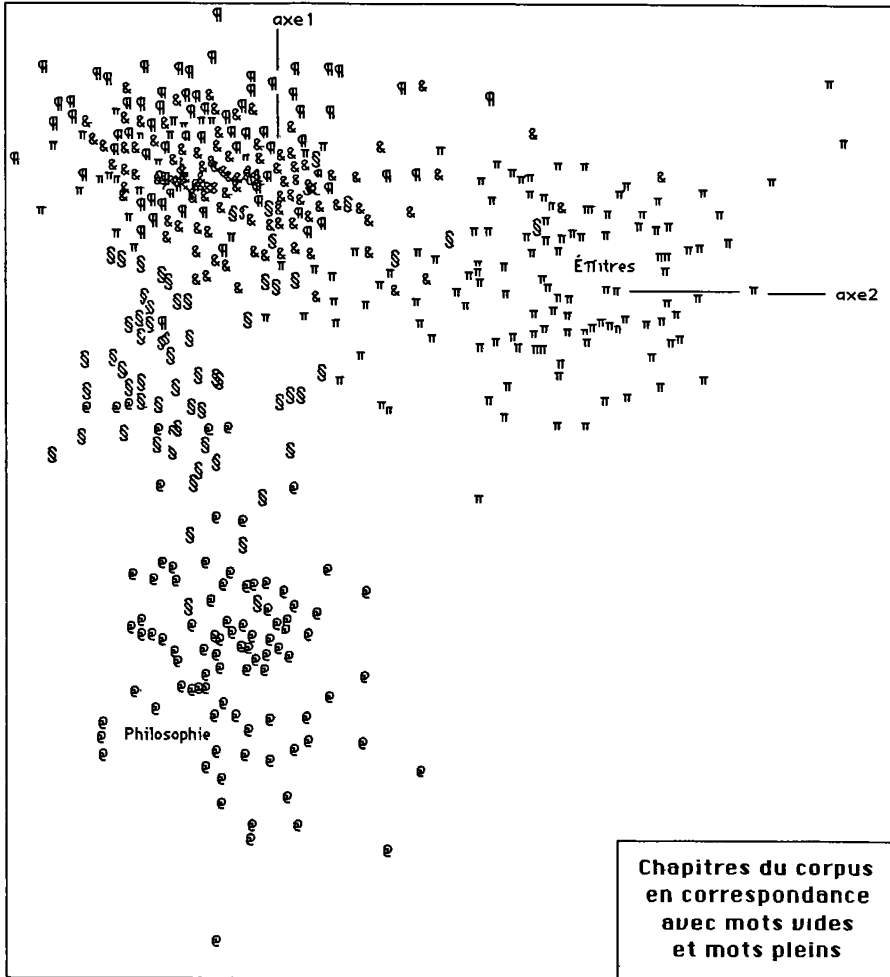
À cette exception près, les textes bibliques, de l'Ancien et du Nouveau Testament, sont du côté (F1>0).

Quant aux classes de formes, sur lesquelles nous reviendrons à propos de l'étiquetage de l'arbre de la CAH des chapitres, on remarque que j22, réduite au mot απορρασις=négation, est la plus écartée sur (F1<0); et c'est pourquoi on l'a placée en dehors du cadre des graphiques, ce que signale une flèche. Viennent ensuite 512, {corps, âme}; 543, qui contient les marques d'une discussion; 542, {nature, cause, principe}.

Du côté (F1>0), la classe la plus écartée est 418: {dis, roi}; vient ensuite 524, qui comprend les cas obliques du pronom de la 2-ème personne du singulier ('te' ou 'à toi'; mais non 'tu'); puis 515, vocabulaire des récits de l'Évangile; et 546, {je, me, moi suis...}.

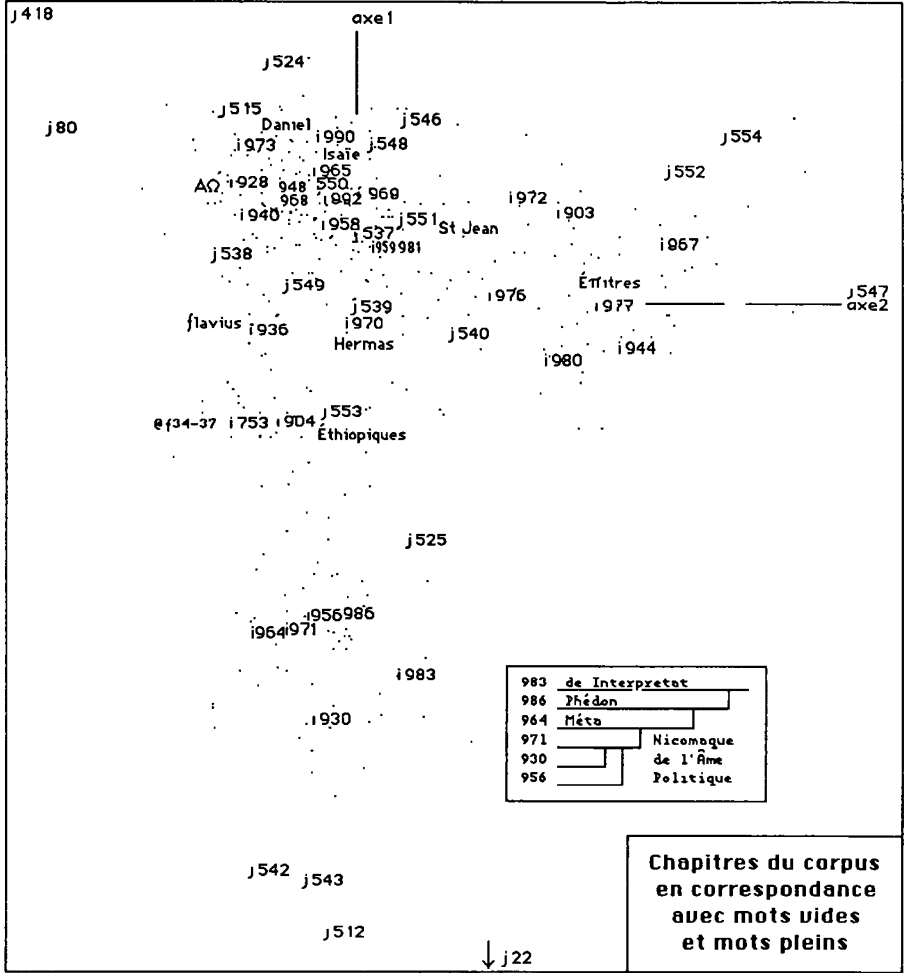
5.1.2 Le deuxième axe et le plan (1,2)

Sur (F2>0) les classes de chapitres les plus écartées sont 967, 944, 977; suivies de 903, 980: il s'agit exclusivement des Épîtres; ce que marque sur le plan (1,2) l'accumulation des lettres 'π' à droite du graphique; les lettres 'π'



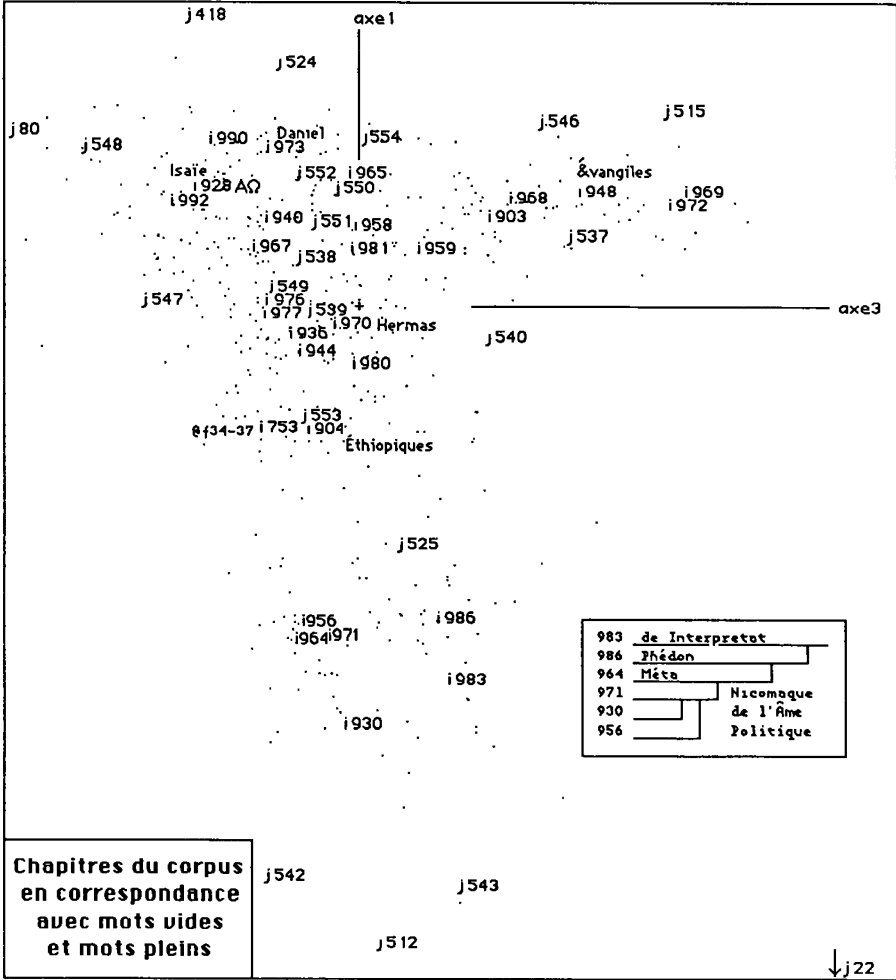
qu'on peut voir à gauche correspondent à des chapitres de l'Apocalypse, $\pi\Lambda\Omega$, et non des Épîtres.

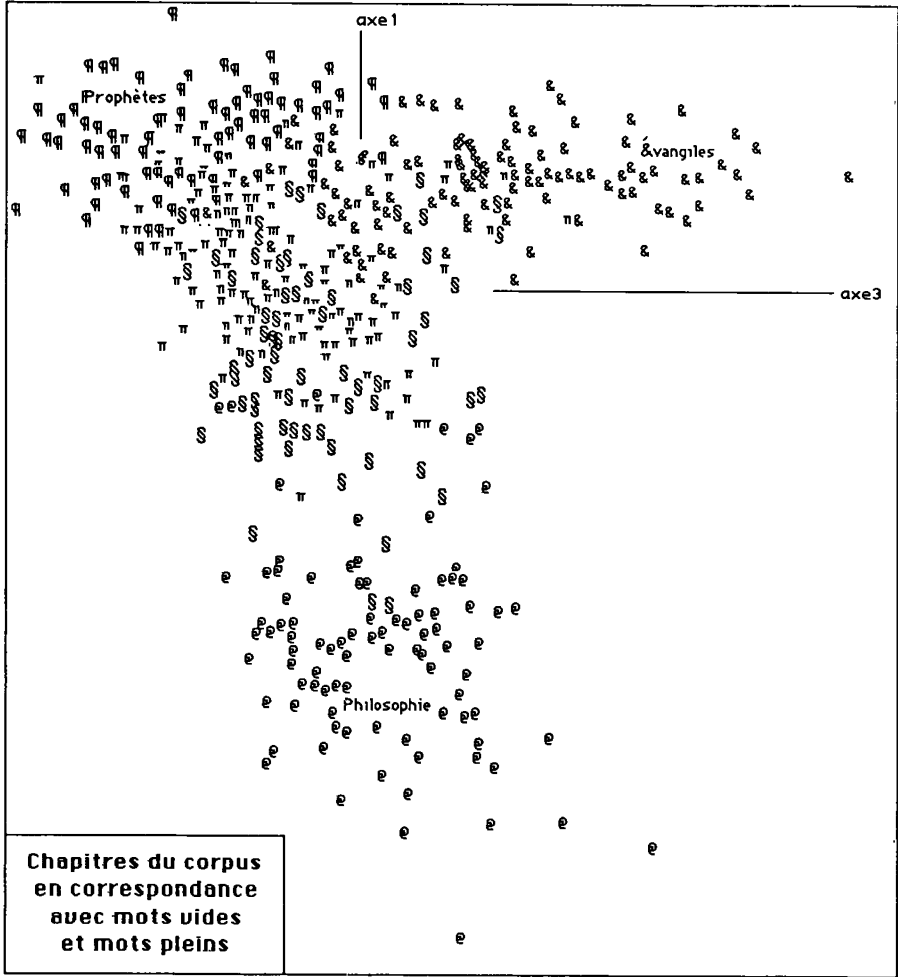
Dans le quadrant ($F1 > 0$; $F2 < 0$), se superposent au sein du nuage des caractères ¶ et &: ¶, Prophètes, {973 992 990}; et &, Évangiles synoptiques (St Matthieu, St Marc, St Luc), {948 968 965}, ou Actes des Apôtres, 958. De l'Évangile selon St Jean, les chapitres de discours, 972, ont, au contraire, un $F2$ nettement positif; le reste, 969, étant, sur l'axe 2, voisin de l'origine.



Dans le quadrant ($F1 < 0$; $F2 < 0$), on trouve la littérature hellénistique, FLAVIUS et les Éthiopiennes, 936, 904; et les textes philosophiques de PLATON et d'ARISTOTE (à l'exception du 'de Interpretatione', 983, qui est dans le quadrant ($F1 < 0$; $F2 > 0$), mais peu écarté de l'axe 1); HERMAS est proche de l'origine.

Quant aux mots, s'écartent le plus sur ($F2 > 0$) les classes 547 et 552, où l'on reconnaît des thèmes des Épîtres. Du côté ($F2 < 0$), on trouve 418, {dis roi}, particulièrement employé dans Daniel et les chapitres ¶136-39 d'Isaïe; puis j80, θηρον, la bête, vue dans certains chapitres du Pasteur ou de l'Apocalypse.





facteurs 2 et 3 pour les séparer: on a, en bref: Épîtres ($F2 \gg 0$), Évangile ($F3 \gg 0$) et prophéties ($F2 < 0$; $F3 < 0$).

Quant aux classes de formes, se détachent dans ($F1 > 0$; $F3 > 0$): {515 537 546}, dont les deux premières sont employées dans les Évangiles; et la troisième, 'Je' et sa suite, est employée dans les discours en St Jean, mais aussi dans Isaïe, dans HERMAS. Dans ($F1 < 0$; $F3 > 0$) on remarque j22, 'négarion'; et, dans ($F1 > 0$; $F3 < 0$), j80, ἀρνιον, bête, ainsi que 548: où l'on croit lire un oracle du prophète Isaïe.

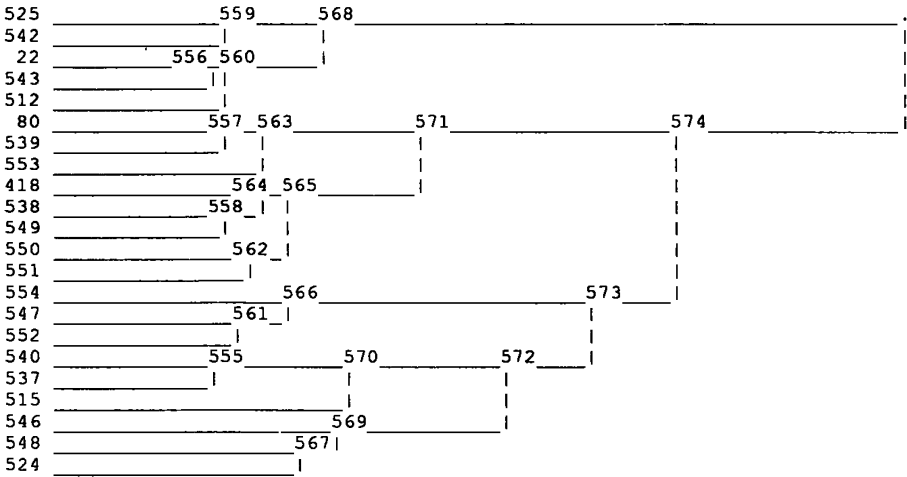
5.1.4 Au-delà du troisième axe

Nous ne citons que les classes pour lesquelles sont élevées non seulement les valeurs absolues de F4 (ou F5), mais aussi les corrélations (au carré), COR4 (ou COR5), portées sur le listage Facor.

Du côté (F4<0) sont particulièrement corrélées à l'axe 4 les classes de chapitres du Prophète Isaïe, 992 et, surtout, 990; les accompagnent les classes de mots { 524 548 546}.

Du côté (F4>0) les plus fortes corrélations sont celles de 936, FLAVIUS, et de la classe de mots 553.

Du côté (F5>0) on note 983, le traité de *Interpretatione*, dont l'association j22, αποφασις négation, crée l'axe5; mais, ayant mis en supplément {@<L, @<M} on ne trouve plus d'association entre le de *Interpretatione* et αληθεις, vrai (à la différence de ce que donnait l'analyse préliminaire rapportée au §5.0).



§5.2: CAH d'après 93 facteurs: arbre de la partition des mots en 22 classes

5.2 Classification des formes de mots pleins et vides

Nous soumettons au lecteur l'arbre de la partition en 22 classes définie par les 21 nœuds les plus hauts; et le tableau du contenu des classes. Les formes grecques sont accompagnées de traductions françaises. Pour les emplois des cas des noms et pronoms et, plus encore, ceux des temps des verbes, il faudrait un commentaire. Or les notions générales sont connues de qui a étudié quelque peu le grec; mais seraient mal comprises de qui ignore totalement cette langue. Quant aux nuances, le statisticien philhellène ne les saisit pas toutes! Plutôt qu'un commentaire, nous insérons donc, dans notre exposé, de brèves remarques.

- 80| θηριον bête
539| κυρια ουτοις εισιν Dame, ce sont
553| δε τελος ενεκεν κατα το μερος τροπον μην περι ανηρ προσωπον
| λαβειν μικρον παρα των τοιο δια της υποκ την ταισ αις τας
| fin en vue de sur part mode homme face prendre petit par sous (article..)
-
- 418| ειπε βασιλευ βασιλευσ βασιλευσ dis, roi
538| γυναικος γυναικα γυνη εκει γη αυτη, αυτην αυτησ femme là elle
549| πυρ υδωρ γη, γην ουρανου γησ feu eau terre ciel
| επ προσωπου οροσ εφ επι εκ ης ωσ sur devant (face) vers la comme
| παντες παντα ακ εργα τα και tous que (relat neutre) oeuvres et
-
- 550| ετη αρτον αυτοι καρδιασ ημεραν ακ ημερασ ακ μετ μετα απο θανατου
| ans pain eux coeur jour après de mort dont jusqu il ουκ εωσ αυτοσ
| εγενετο συν ανδρες ονοματι χειρασ αγιου ακ αυτοουσ πληθοσ
| advint avec hommes nom mains saint multitude faire peuple qui d'eux
| ποιησαι λαον οικ αυτων τουσ πολλιν νεκρων θεου κυριον εισ
| προσ υιου ακ τον les ville morts Dieu Seigneur vers fils le
551| αυτου του ερημω, γυνη γυνην ειδον αγγελοσ ονομα οιδα
| (de) lui de (au) désert voix virent ange nom sais
| εχων ενωπιον οσ ακ τω, ανθρωπον παν ανθρωπων τη, εν
| ayant devant qui au homme(s) tout (à) la dans
-
- 554| πατροσ πατερα πατηρ υμεισ ακ υμιν ακ αδελφοι κυριω, κυριου
| υμασ ακ υμων père (... , à, de...) vous frères (au, du) Seigneur
547| εργα πιστει πιστιν πιστεωσ νομω, νομον σαρκι νομου
| (des) oeuvres foi loi chair
552| ημων ακ κοσμου γεγραπται καθωσ ινα ακ θανατου αμαρτιασ ακ ζωησ ζωη
| εργα χαριν χαρισ θεω, θεου θεουσ πνευμα πνευματοσ πνευματι
| nous monde (est) écrit comme afin mort péché vie oeuvre grâce Dieu esprit
-
- 540| σημειον ρημα χρονον μη καλον εαν ουχι λεγω ανθρωπου βασιλεια
| βασιλειαν αληθεσ ανθρωποσ εστιν αδελφον ακ αρχασ ουκ οτι ακ
| signe parole temps ne (pas) beau si dis homme royauté vérité est frères
| de principes que (conj.)
537| συ ουρανον ουκ ουτοσ ακ οκ ουν εκεινοσ ειπον απεκριθη ερχεται
| οπου ακ δυναται ελεγεν ελεγον τοπον παλιν tu ciel que (relat) donc
| celui-ci dirent répondit vient où peut di (ren) t (de) nouveau
-
- 515| αυτω, ην αρχιερευσ αυτον ηλθεν εξηλθεν οτε ακ ησαν αυτοισ
| (à) lui, eux étai (en) t pontifes alla sortit quand plorou μαθηται
| κυριε ιδωσ λεγοντες λεγων αποκριθεισ υιοσ ειπαν ειπεν
| barque disciples Seigneur! voyant disant répondant fils di (ren) t
-
- 546| μου λεγει γησιν μοι εμου εμε με ειμι εγω je moi suis di (sen) t
-
- 548| ημερα, εκεινη, εσονται εσται σπερμα δικαιοσυνην σαρκα τεκνα
| (en) ce jour seront sera race chair enfants justice
| δοξησ γωσ δοξαν δοξα εθνων εθνη χρισιν λαοσ πολισ κυριοσ
| lumière gloire nations peuple jugement Cité Seigneur
524| σοι σε σου à toi, te, de toi

Le contenu de certaines classes de formes a déjà été considéré dans ses rapports avec les textes, lors de l'interprétation des facteurs. L'étiquetage de la classification des chapitres permettra de rendre raison de la constitution de la plupart des classes de formes. On pourrait également étiqueter la classification des formes en terme de classes de chapitres: mais nous nous en abstenons,

c		Partition en 22 classes : mots de la classe numéro c
525		αγαθον κακον καλωσ ευ ποιειν ποιει γαρ βει πως εξ τισ ει λογουσ λογων λογω, λογου ανθρωποισ λογον λογοσ λεγειν τι επι ταυτ οταν< ουδε ουτωσ νυν εναντιον τουτο αυτο αλλα ου ουχ αλλα bon mauvais bien faire fait faut comment de(ab) un parole homme dire encore quand ni ainsi maintenant contraire ce même mais non(pas)
542		ψυσει ψυσιν ουτε αιτιαν αρχασ αιτιασ nature ni cause principe
22		αποφασισ negation
543		ομοιωσ δηλον αναγκαιον αναγκη ειπειν η αν αληθη λεγεισ de même évident nécessaire dire ou vrai (tu)dis
		οιμαι ψαινεται δοκει (je)crois (il)semble
512		σωματοσ ψυχη ψυχησ ψυχην corps âme

Tableau des classes de formes de mots pleins et de mots outil

Au sommet de la hiérarchie, l'ensemble des 288 formes conservées en principal dans l'analyse du §5.1, se subdivise en deux branches, 568 et 574.

Du fait de l'étendue du tableau, on a placé ci-dessus, le contenu de la branche 568, et ci-contre, à gauche, le contenu de la branche 574.

Pour chacune des 22 subdivisions retenues on a donné les formes grecques, en suggérant une traduction française qui nous paraît correspondre à l'emploi des formes dans le corpus. Les différences de cas ne sont pas toujours indiquées.

Quant à l'orthographe du Grec, on a suivi les conventions appliquées dans nos programmes de tri, comme on l'explique dans [LING TRI 2], §1.2, (in *CAD*, Vol.XV, n°2, 1991). En particulier, les accents et les esprits doux sont totalement supprimés; l'esprit rude est rejeté à la fin du mot, et noté '<'; le iota souscrit n'est conservé qu'en fin de mot, et noté comme un signe séparé 'ι'.

parce que, pour commenter cet étiquetage, on devrait répéter ce qui doit être dit à propos de celui de la CAH des chapitres.

Avant de passer à la classification des chapitres, nous ferons deux remarques. D'une part, à l'exception de quelques unes d'entre elles réduites à un très petit nombre de formes, les subdivisions retenues contiennent à la fois des mots vides et des mots pleins. D'autre part, les formes rentrant dans le paradigme d'un même mot (déclinaison d'un nom, conjugaison d'un verbe...), sont souvent réunies à un bas niveau de la hiérarchie, particulièrement s'il s'agit d'un mot plein ou d'un pronom des 1-ère et 2-ème personnes: ce qu'on expliquera par le fait que de tels mots sont, tout au long de leur paradigme, liés à un même thème ou à un même genre. Une exception notable étant celle du pronom 'tu', dont les cas obliques constituent la classe 524; tandis que le nominatif est dans 537.

5.3 Classification des chapitres

On a d'abord considéré une partition en 22 classes définie par les 21 nœuds les plus hauts; puis, au vu du contenu des classes ainsi retenues et de l'arborescence complète des chapitres, on a voulu subdiviser certaines classes; et, finalement, on a construit une partition en 29 classes définie par des nœuds spécifiés.

983	@<B	@<E	@<D	@<H	@<F	@<I	@<C	@<G	@<J	@<A	@<K	@<N	22+++	540+++	
986	@f39	\$ÉT8	\$ÉT _w	\$ÉT _q	@f01	@f26	@f18	@f40	@f30	@f38	@f02	@f04	@f03	@f05	
	@f14	@f13	@f20	@f33	@f16	@f15	@f17	@f06	@f07	@f22	\$ÉT1	@f21	@f25	@f19	
	@f23	@f24	@f32	@f12	@f11	@f29	@f31	@f10	@f09				525++	543+++	512+
964	@MA7	@f28	@f27	@MA4	@MA5	@MAX	@MA8	@MA3					542++++	553+	
971	@N10	@f08	@N06	@MA9	@N01	@N12	@N11	@N09	@N02	@N05	@N07	@A2	@N04	@N03	
	@N08	@P11	@MA6	\$ÉT _k	@MA2	@MA1				525+	543+		553++		
930	@N13	@A1	@A3	@A5	@A4					525++	512++++		553+		
956	@P03	@P01	@P12	@P13	@P05	@P06	@P07	@P04	@P02	@P10	@P08	@P09	553+	542+++	

Tableau des classes de chapitres

Au sommet de la hiérarchie, l'ensemble des 507 chapitres conservés en principal dans l'analyse du §5.1, se subdivise en deux branches, 1012 et 1009.

Du fait de l'étendue du tableau, on a placé ci-dessus, le contenu de la branche 1009, et ci-contre, à gauche, le contenu de la branche 1012.

Ainsi, se trouvent distinguées la classe 753, formée des chapitres {@f34... 37} du Phédon; la classe 904, {\$ÉT...} formée de 26 chapitres des Éthiopiens; la classe 958, {&@...}, 25 chapitres des Actes des Apôtres; 948, qui comprend les récits de la Passion dans les 3 évangiles synoptiques (avec &M17 &M20, deux chapitres de St Matthieu qui, d'ailleurs, contiennent l'annonce de la Passion); et, enfin, au sein des chapitres d'ARISTOTE, 930, qui comprend, avec le chapitre 13 de l'Éthique à Nicomaque, 4 sur 5 des chapitres du Livre I du Traité de l'Âme: 930 = {@N13 @A1 @A3 @A5 @A4}.

Le listage iVacorjq nous a servi à étiqueter la CAH des chapitres en termes de classes de mots: les plus fortes contributions sont notées sur le tableau des classes; peu sont sur l'arbre, où la place libre a servi à résumer le contenu des classes.

Il n'est pas surprenant qu'au sommet de l'arbre se détache la classe 1009 qui coïncide presque avec l'ensemble des textes philosophiques du corpus; la seule exception notable étant les 4 chapitres {@f34... 37}=753 du Phédon, où Platon expose le sort des âmes après la mort. L'étiquetage des textes philosophiques par les classes de mots s'accorde avec ce que l'on sait du vocabulaire de la dialectique. Le mot n° 22, ἀποφασισ, caractérise le *de Interpretatione*, où sont toutes ses occurrences; et la classe 512, qui comprend σωματος, génitif de corps, et trois cas de ψυχη, âme, culmine dans le traité de l'Âme, 930, puis dans le Phédon, 986.

Se détache ensuite la classe 1005, qui comprend tous les chapitres des Épîtres, exception faite des première et deuxième de St Jean et de 8 chapitres divers; à quoi s'associent les 2/3 des chapitres du Pasteur d'HERMAS, écrit allégorique, d'inspiration chrétienne, daté du début du deuxième siècle; (avec, en particulier, dans 970, une vision expliquée par une Dame, d'où l'étiquetage par 539 = {κυρια ουτοκι ειςιν : Dame ce sont}).

965	Nat_985_&_:	03_:	07_:	10_:	11_:	12_:
958	@ct_					
936	_fl_984_					
753	_962_		@f34-37			
904		Ethiopiennes				
928	Apoc_994_:	04_				
940						
973	Daniel_		+ Is hist			
992			:02_ Isaïe_			
990						
959		997_syno_:	08_&vangiles_			
948	_975_		Passion_			
968						
969			:00_		&v St Jean	
903	_996_				Épîtres 1,2	
972						
981	H989_999:	01:05			Épîtres	
976					+ 14 Hermas	in 981
970	Hermas_				+ 7 Hermas	in 970
944					(8+2 Hébreux	in 976, 981)
980	_987_993_					
977						
967						
983	_de Interpretat_:	09_	543++_	525+_		Philosophie_
986	Phédon_		:06_			
964	Méta_998_		553+			
971	_991_995_		Nicomaque			
930			de l'Ame			
956			Politique			

CAH d'après 93 facteurs: arbre de la partition des chapitres en 29 classes
NB sur les noeuds de l'arbre les numéros (1000... 1012) sont écrits (:00... :12)

L'étiquetage des subdivisions de 1005 n'est pas uniforme: mais dans les classes de mots {554 547 552} un lecteur familier de St Paul, croit entendre la voix de l'Apôtre des Gentils.

Après 1005 se sépare la classe 1008, qui est celle des Évangiles: St Jean dans 1000; et les trois autres (dits synoptiques du fait de leur parallélisme) dans 997. Entre 1000 et 997, la similitude procède du thème, non de la langue: chaque fois que l'on prend en compte les formes de mots outil, se manifeste l'originalité, la singularité, du style de l'Évangile selon St Jean. (Toutefois dans la présente analyse, comme dans beaucoup d'autres, &J02 rejoint les synoptiques).

Dans la classe 515 de formes, associée à plusieurs subdivisions de 1005, on a le schéma d'un récit évangélique: {(à)lui,eux étai(en)t pontifes alla sortit quand barque disciples Seigneur! voyant disant répondant fils di(ren)t}. De même 537, {tu ciel que(relat) donc celui-ci dirent répondit vient où peut di(ren)t (de)nouveau} avec 969, puis 972 et 948.

La classe 546, qui est celle de la première personne, culmine dans 972, les discours du Seigneur rapportés en St Jean (particulièrement, les discours après

la Cène, avant la Passion: &J14-17). La 1-ère personne est, à un moindre degré, fréquente dans 981: Hermas dit ce qu'il a vu; et dans 990: Dieu parle en Isaïe.

La classe 554 {père (...à,de...)vous frères (au,du)Seigneur} dont le sommet a déjà été vu avec St Paul dans 967, est également fortement associée à 972, discours en St Jean.

Le caractère linguistique de l'Évangile selon St Jean ne réside sans doute pas dans le contenu; lequel, d'une part, ainsi que l'atteste le partage en deux subdivisions {969 972}, n'est pas uniforme sur l'ensemble des chapitres; et, d'autre part, est proche de celui des trois Évangiles synoptiques, ce qui entraîne la constitution d'une classe, 1008, comprenant les quatre Évangiles. Nous avons remarqué l'usage parcimonieux de mots outils renfermés dans 553.

La classe 553 (qui comprend 23 formes) est le plus associée d'une part aux classes de chapitres {936 753 904}, subdivisions de 964, (littérature hellénistique: FLAVIUS, Éthiopiennes et, à un moindre degré, la fable géologique du Phédon); et, d'autre part, aux subdivisions retenues de 998, {964 971 930 956}, (œuvres d'ARISTOTE, de *Interpretatione* excepté).

À l'opposé, la classe 553 est le moins associée aux subdivisions de la classe 1000, {969 903 972}: i.e. à l'Évangile selon St Jean et à la 1-ère Épître de St Jean.

Dans l'Apocalypse (928), où ce même Apôtre St Jean écrit, en son nom propre, le récit de révélations qui lui ont été faites, les formes de 553 ont un niveau de fréquence moyen; et il en est de même dans les Évangiles synoptiques (997 = {959 948 968}).

Reprenons la lecture de la partition des 507 chapitres. Reste à considérer la classe 1010, scindée en 1007 (123 cap.) et 1002 (73 cap.).

La classe 1002 comprend 59 des 66 chapitres du Livre d'Isaïe; (font seuls exception ¶I50 et ¶I53, rangés dans 989, au sein des Épîtres; ¶I06 avec Daniel dans 973; et ¶I36-39, qui contient des épisodes de la vie d'Ézéchias, avec des oracles transmis à ce roi).

Le thème de tout le Livre est renfermé dans 548 {(en) ce jour seront sera race chair enfants justice lumière gloire nations peuple jugement Cité Seigneur}; mais la subdivision 990, dont 3 chapitres seulement ont un rang inférieur à 40, se distingue par l'emploi fréquent de 'je' (546: cf. *supra*) et des cas obliques du pronom 'tu' (524: te, toi; mais non su=tu nominatif ou 'toi', vocatif, qui est dans 537: on a de nombreuses occurrences de cette forme dans des dialogues des récits évangéliques: tels que la deuxième partie de &J01, &J08...). Nous l'avons dit au §2.2: on attend du dénombrement des outils personnels qu'il saisisse quelque chose de la vivacité du jeu des paroles, quand le texte montre,

avec les mots, ceux qui les disent ou les reçoivent. Il apparaît que le début et la fin du Livre diffèrent sur ce point.

La classe 1007 se scinde en 1003 et 1004.

Dans 1004 sont distinguées les subdivisions {928 940 973}. On a déjà dit que 973 comprend, outre Daniel, ¶I06 et ¶I36-39 (épisodes de la vie d'Ézéchias, avec des oracles transmis par Isaïe à ce roi): d'où l'étiquette 418+++++ (418 = dis, roi). La classe 928, étiquetée 549+++ , est formée de 16 des 28 chapitres de l'Apocalypse; deux autres chapitres, 13 et 17, {ΑΩδ ΑΩι} sont dans 940 ainsi que 3 chapitres du Pasteur avec pour étiquette 80+++++ , 'la bête'. Ainsi la classe 1004 apparaît cohérente quant à son contenu.

La classe 1003 se partage en 985 et 984. La classe 985 contient, outre le chapitre ¶θ13 de Daniel (histoire de Susanne), 40 chapitres du Nouveau Testament, dont 26 des 28 chapitres des Actes des Apôtres (&@...); 24 de ces chapitres constituent la subdivision 958, étiquetée 550++++ : le poids de cette classe est particulièrement faible dans l'ensemble des œuvres philosophiques d'ARISTOTE et de PLATON (1009 et 753).

L'autre subdivision de 985, la classe 965, comprend principalement des chapitres des Évangiles synoptiques: notamment des récits de la Nativité et de la Résurrection. On remarquera, que compte tenu de la préparation méticuleuse du tableau analysé, la CAH ne se ressent pas de la présence des chapitres &M01 et &L03, qui contiennent des généalogies avec répétition de formules (cf. *supra* §2.1).

Nous terminerons en considérant la classe 984. On y a distingué trois subdivisions: {936 753 904}: 753 est ce passage du Phédon où Socrate décrit les lieux que hantent les âmes après la mort; 936 renferme l'ensemble des chapitres saisis dans les Livres X et XVIII des Antiquités Judaïques de FLAVIUS, ainsi que deux chapitres des Éthiopiennes; 904 est formé de 26 des 33 chapitres que nous avons distingués dans le Livre I de cet aimable fiction.

Que 984 s'agrège à 985 pour constituer la classe 1003, suggère que, dans le Nouveau Testament, la langue des récits, et, notamment, celle de St Luc dans les Actes, n'est pas étrangère à l'usage commun de la littérature hellénistique. À l'analyse factorielle, la classe 1003 se caractérise par (F4>>0), caractère que possèdent 985 et 984; 1003 se scinde suivant l'axe 1: pour 985 (Actes etc.), on a (F1>0); tandis que 984 (notamment ses subdivisions: 753, du Phédon, et 904, Éthiopiennes) va sur (F1<0) vers les textes philosophiques de la classe 1009.

§6 Éléments supplémentaires et discrimination

Dans les études de corpus de textes, nous avons souvent eu recours à l'analyse discriminante (notamment au §3 de [TEXTES GRECS]). En bref, on analyse un tableau croisant un ensemble d'œuvres avec un lexique; la ligne

afférente à une œuvre n'étant autre que le cumul des lignes afférentes aux chapitres de celle-ci. Les chapitres individuels, adjoints en supplément à l'analyse, sont affecté chacun à l'œuvre du centre de laquelle il est le plus proche.

Afin d'apprécier la valeur potentielle de cette méthode pour affecter des chapitres non pris en compte dans l'analyse principale, on a coutume de réserver un 'échantillon d'épreuve', *le*: c'est-à-dire des chapitres (e.g. ceux dont le rang est divisible par 4) non comptés dans les cumuls définissant les centres des œuvres. L'affectation des chapitres de cet échantillon est, *a priori*, plus difficile que celle des autres chapitres; et son taux de succès mesure la valeur potentielle de la méthode.

Dans la présente étude, nous n'avons pas fait d'analyse discriminante parce que la qualité de la classification ascendante présentée au §5.3 est telle que, sans prendre aucunement en compte la répartition des chapitres en œuvres, on retrouve celle-ci avec une approximation que (compte tenu de l'hétérogénéité inhérente aux textes) on n'a pas de raison de prétendre dépasser.

Mais il reste à apprécier la stabilité potentielle des résultats obtenus.

```
380 chapitres x 291 mots de {ΔμΔα, θν} ; écrêtage pour les mots pleins;
en suppl: {evh, ejnaj, te} ; et (1/4) des chapitres dont {@<L, @<M} ;
trace : 2.193e+0
rang : 1 2 3 4 5 6 7 8 9 10
lambda : 1500 918 860 686 467 410 385 371 329 307 e-4
taux : 684 419 392 313 213 187 176 169 150 140 e-4
cumul : 684 1102 1494 1807 2020 2207 2382 2551 2701 2842 e-4
```

À cette fin, on reprend le principe de l'échantillon d'épreuve en l'appliquant à l'analyse factorielle. De façon précise, on analyse un tableau (I-Ie) × ∂Δ qui ne diffère de celui considéré au § 5.2 qu'en ce que, de l'ensemble, I, des lignes (chapitres), un quart, *Ie*, comprenant les chapitres {@<L, @<M}, a été retiré; les formes {εψη, ερναρ, τε} étant, comme précédemment, mises en supplément.

```
c 1017 1016 1015 1014 1013 1012 1011 1010 1009 1008 1007 1006 1005 1004
car 509 416 299 224 93 198 79 102 67 117 35 75 96 46
T 788 410 319 256 228 200 174 167 128 119 112 107 106 97
A 1016 1015 1014 1012 994 1005 989 1009 995 947 998 992 987 985
B 1013 1008 1006 1003 1011 1010 1004 1007 990 1002 956 997 1001 1000
```

§6 : CAH des 509 chapitres d'après leurs coordonnées sur 93 axes issus de l'analyse d'un tableau principal croisant 380 chapitres avec 288 mots :

la somme des indices de niveau est 2167 e-3 ; les taux T sont comptés en e-4 .

Sur les axes issus de cette analyse, tous les chapitres ont des coordonnées: ceux de *Ie* comme supplémentaires, les autres comme principaux. On a pu ainsi effectuer une CAH de I; laquelle, dans ses grandes lignes, diffère peu de celle présentée au §5.2; et, jusque dans les détails, ne le cède en rien à celle-ci.

c	Partition en 28 classes : Sigles des chapitres de la classe numéro c
969	εM01 εM02 εa12 εL02 εL01 ϖd13
963	εa20 εa27 εa21 εa16 εa08 εa19 εa11 εa18 εa10 εa09 εm16 εM28 εL16 εL08 εL24 εL09 εa05 εa04 εa14 εa17 εa15 εa13 εa03 εa01 εa02 εa25 εa24 εa23 εa28 εM25 εa26 εa22
984	πHb1 εa07 εL15 εM03 εM04 εL04 εm01 εL14 εL13 εL19 εL10 εL07 εM22 εm12 εL20 εL18 εm10 εM19 εm03 εm09 εm11 εM21 εM08 εM14 εm08 εm05 εm06 εM15 εm07 εL05 εM09 εm02 εL23 εM27 εm15 εL22 εM20 εM17 εM26 εm14
971	εM06 εM18 εM05 εL17 εM16 εM12 εL11 εL06 εM07 εm04 εM13 εL12 εM23 εM11 εM10 εL21 εM24 εm13
940	πIoΔ πPb2 \$HrF \$HrE \$HrD
-957	\$HrB \$HrG \$HrC \$HrI \$Hr1 \$Hr8 \$HrP \$Hr2 \$Hr5 \$Hr4 \$Hr9
975	εL03 \$F2 \$F3 \$F1 \$F4 \$f7 \$f3 \$ÉT5 \$f6 \$f8 πJk3 \$HrA \$f5 \$f1 \$f4 \$f2
771	@f36 @f35 @f37 @f34
896	\$ÉTw \$ÉTq \$ÉT9 \$ÉTh \$ÉTc \$ÉT6 \$ÉT4 \$ÉTi \$ÉTo \$ÉTv \$ÉT1 \$ÉT2 \$ÉTd \$ÉTx \$ÉTtu \$ÉT3 \$ÉT7 \$ÉTs \$ÉTm \$ÉTn \$ÉTt \$ÉTr \$ÉTj @f08 \$ÉT8 \$ÉTa \$ÉTf \$ÉTp \$ÉTb \$ÉTg \$ÉTe
952	\$HrM \$HrO \$HrN ϖd07 πΑΩd πΑΩi
979	πΑΩ2 πΑΩ3 πΑΩ7 πΑΩb πΑΩg πΑΩh πΑΩm εa06 πΑΩ4 πΑΩk πΑΩl πΑΩ5 πΑΩ6 πΑΩa πΑΩ9 πΑΩ8 πΑΩf
956	ϖd03 ϖI37 ϖI36 ϖd01 ϖI39 ϖd06 ϖd14 ϖd02 ϖd05 ϖd04 ϖd08 ϖd11
981	εJ18 εJ19 εJ20 εJ02 εJ11 εJ04 εJ21 εJ13 εJ12 εJ06 εJ07 εJ03 εJ01 εJ09
884	πIa4 πIa5 πIa1 πIa2 πIa3
978	εJ17 εJ05 εJ10 εJ08 εJ15 εJ16 εJ14
992	ϖI26 ϖI63 ϖI64 ϖd09 ϖI58 \$Hr7 ϖI47 ϖI54 ϖI60 ϖI38 ϖI57 ϖd10 ϖI49 ϖI44 ϖI48 ϖI41 ϖI43 ϖI46 ϖI45
967	ϖI55 ϖI51 ϖI40 ϖI42 ϖI59 ϖI03 ϖI33 ϖI01 πHb8 ϖI56 ϖI65 ϖI61 ϖI66 ϖI50 ϖMa2 ϖMa1 ϖMa4 ϖMa3
993	ϖI15 ϖI35 ϖI23 ϖI34 πΑΩc πΑΩj πΑΩn ϖI25 ϖI02 ϖI17 ϖI20 ϖI07 ϖI19 ϖI27 ϖI13 ϖI05 ϖI24 ϖI18 ϖI14 ϖd12 πΑΩo ϖI06 ϖI32 ϖI09 ϖI21 ϖI31 ϖI08 ϖI30 ϖI04 ϖI11 ϖI16 ϖI28 ϖI10 ϖI29 ϖI62 ϖI52 ϖI12 ϖI22
947	πJk2 πGL2 πRm3 πGL3 πRm2 πRm7
972	πCaB πCa7 πCaD πCa9 πCaE πCaC πCa3 πCa2 πCb3 πCb4 πCaA πCa5 πCa1 πCa6 πRm8 πGL6 πGL5 πGL4
973	πRm9 πRm4 πR10 πHb7 πHb2 πHb9 πHbA πHb6 πHb4 πHbB πTa6 πTa4 πTa2 πJk4 πJk1 πCa8 πR11 πHbC πPa2 πHbD πCaF πGL1 πTa5 πTt1 πTb4 πTb3 πTb2 \$Hr3 πIHb πHb5 πIHc πPHm πR13 πR12 πR14 πTt2 πTt3 ϖI53 πRm6 πRm5 πTa1 πTb1
988	πCb7 πPb3 πCb8 πθa2 πθa3 πθb2 πCb1 πθb1 πEp3 πCL1 πEp1 πEp2 πPa4 πfπ3 πPb3 πHb3 \$Hr6 \$HrH \$HrJ \$HrL \$HrK πJk5 πCb5 πCbA πEp5 πCL3 πEp6 πfπ4 πR16 πCL4 πCaG πθb3 πθa4 πθa5 πCb6 πCbB πCa4 πCbC πPa3 πCb2 πCbD πR15 πfπ1 πRm1 πfπ2 πEp4 πCL2 πCb9 πPa5 πPa1

L'ensemble des chapitres des textes de philosophie constitue une classe 1013 qui se distingue de tout le reste; comme au §5.3, font exception les chapitres {@f34-37} du Phédon.

7 Conclusions et perspectives

7.1 Taxinomie des textes du corpus

De l'ensemble des classifications effectuées, sort une taxinomie que nous décrivons en distinguant les segments les plus stables.

Dans [TEXTES GRECS] (in *CAD* vol.XVI n°1) nous remarquions 'l'unité de la langue classique vis-à-vis de celle du Nouveau Testament ou des Septantes'. En ajoutant au corpus de nombreux chapitres d'ARISTOTE et de PLATON, on voit se séparer nettement une branche de la littérature philosophique, où manque seulement le bloc @f34-37 qui expose le mythe géologique du Phédon.

Les autres passages non bibliques viennent du roman des Éthiopiens, des Antiquités Judaïques de FLAVIUS, et du *Pasteur* d'HERMAS. Les deux premiers textes sont liés dans la plupart des CAH; le troisième a, avec la Bible, une analogie de ton qui en rend la place incertaine, au sein de notre corpus.

Les Épîtres du Nouveau Testament (dont la plupart sont explicitement écrites au nom de St Paul) forment un ensemble compact bien reconnu par les CAH. On sait que l'Épître aux Hébreux, peut-être traduite d'un original sémitique, se distingue des autres Épîtres: la distribution de ses 13 chapitres varie d'une analyse à l'autre. Au §3.1, comme dans [TEXTES GRECS], cette Épître se montre, par la distribution des outils non personnels, à la fois en rapport avec les Actes et avec des textes hellénistiques, particulièrement le Pasteur. Mais si l'on prend en compte les outils des personnes ou les mots pleins, l'Épître aux Hébreux, de par son contenu, va dans la branche des autres Épîtres où ses chapitres sont assez bien groupés.

Tout confirme que l'Évangile selon St Jean et la 1-ère Épître de cet Apôtre forment un bloc bien caractérisé et auquel rien ne ressemble: dès 1979, Mgr. B. de SOLAGES et l'Abbé J.-M. VACHEROT avaient montré, par l'analyse des correspondances, que le chapitre 21 de Jean ne se sépare pas des autres et que le tout se distingue des Évangiles synoptiques (cf. *Prat. Ling.*, LC1 n°3). Nous n'avons toujours pas trouvé de texte grec qui soit écrit comme l'Évangile selon St Jean. En revanche dans une analyse basée exclusivement sur les mots pleins, les chapitres des quatre Évangiles se rejoignent (cf. §4); et, si l'on compte à la fois outils et mots pleins, l'Évangile selon St Jean peut former une seule branche avec les synoptiques: mais cette figure, vue au §5.3, disparaît au §6 du fait d'un simple changement d'échantillonnage.

Le Livre du Prophète Isaïe (auquel se joint Malachie) est généralement reconnu par la CAH: les blocs ¶I01-35 et ¶I40-66 sont inégalement distribués dans les subdivisions taxinomiques; et ¶I36-39 est à part, avec Daniel.

L'Apocalypse s'intitule: 'Révélation de Jésus-Christ... qu'il a fait connaître, en l'envoyant par son ange à Jean, son serviteur...'. Mais la langue

de l'Apocalypse n'est aucunement celle de l'Évangile selon St Jean: par la langue, comme par le thème, l'Apocalypse confine aux prophéties de l'Ancien Testament; mais la place n'en est pas fixée.

Les Évangiles synoptiques s'identifient généralement à une branche de la CAH; mais les récits, notamment ceux de la Nativité et de la Résurrection peuvent se séparer du reste et rejoindre, e.g., les Actes des Apôtres dont l'ensemble est souvent très bien reconnu.

7.2 Style et contenu

D'une classification de I, l'étiquetage de l'arbre en termes de classes de formes de J, permet de rendre raison: en bref, les mots outil marquent le style; et les mots pleins, le contenu. Plus précisément, on distingue au sein du lexique trois segments: outils, outils personnels et mots pleins. Un mot plein est choisi comme caractéristique si, pour plusieurs des chapitres, il se place dans les 10 ou 15 premiers rangs (cf. §2). Et le tableau des occurrences des mots pleins est écrêté pour réduire l'effet des répétitions en grappes.

Ainsi, comme on l'a vérifié au §6, la classification hiérarchique de l'ensemble des chapitres d'un corpus peut atteindre la précision et la stabilité d'une analyse discriminante; ce, bien que la répartition des chapitres en œuvres ne soit pas prise en compte pour créer les tableaux analysés.

Avec quelque 250000 occurrences, le corpus dépasse de beaucoup ce qu'on a coutume de traiter dans les expériences d'analyse de données. Mais on ne peut prétendre avoir embrassé ne fût-ce qu'un domaine autonome de la littérature grecque: par exemple, les Actes des Apôtres et les Épîtres du Nouveau Testament se ressentent des usages rhétoriques de l'époque hellénistique; usages non attestés ici. Il semble difficile de rassembler un corpus adéquat; lequel semblerait, *a priori*, devoir être au moins dix fois plus étendu que celui traité.

Il nous paraît du moins satisfaisant qu'en étendant simultanément le corpus et les faits linguistiques pris en compte, on découvre, de plus en plus clairement, l'ordre dans la complexité.

Références bibliographiques

[CAD XII-XVII (1-2-3)]: "Analyse du vocabulaire et recherche du thème dans les articles des volumes XII à XVII de CAD, (1) Le corpus et les résumés, (2) Lexiques et grappes; (3) Typologie et discrimination", CAD, Vol XVIII, n°1, pp. 47-60, 61-74, 75-96; (1993).

[LING. TRI]: "Programmes de statistique linguistique fondés sur le tri par fusion de fichiers de texte", CAD, Vol XV, n°1, pp. 59-82; (1990).

[LING. TRI 2]: "Programmes de statistique linguistique: mise à jour et applications à l'étude du contenu de textes bibliques en grec", *CAD*, Vol XVI, n°2, pp. 133-160; (1991).

[Prat. Ling]: *Pratique de l'Analyse des Données en Linguistique et Lexicologie*; Dunod, Paris; (1981).

[TEXTES GRECS]: "Typologie de textes grecs d'après les occurrences des formes des mots outils", *CAD*, Vol XVI, n°1, pp. 61-86; (1991).