

J.-P. BENZÉCRI

F. BENZÉCRI

Typologie de textes latins d'après les occurrences des formes des mots outil

Les cahiers de l'analyse des données, tome 16, n° 4 (1991),
p. 439-466

http://www.numdam.org/item?id=CAD_1991__16_4_439_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 LATINS D'APRÈS LES OCCURRENCES DES FORMES DES MOTS OUTIL

[TEXTES LATINS]

J.-P. & F. BENZÉCRI

1 Thème de la recherche et préparation des données

1.1 Le Trésor de la Langue Latine

Partant du texte sacré du Nouveau Testament, nous avons déjà analysé l'emploi des formes de mots outil dans des textes écrits en Grec classique. Marquée par des influences dialectales complexes, embrassant un héritage immense, cette langue offre matière à des entreprises statistiques dont le programme même ne nous apparaît pas clairement!

Le Trésor de la Langue Latine suffirait, certes, quant à lui, à nourrir la méditation de plusieurs vies de savant, mais il est au moins possible de l'embrasser du regard: déjà, J. DENOZ a analysé la fréquence des catégories grammaticales en Latin, d'après un ensemble de textes où il nous semble que sont représentés la plupart des genres classiques; la seule exception majeure étant celle du Théâtre Comique (cf. [FRÉQ. CAT. LATIN]).

Encouragé par cet exemple, nous avons entrepris d'analyser l'emploi des formes de mots outil en Latin; en sortant du champ des auteurs classiques, pour explorer l'œuvre des Pères de l'Église, sans omettre les textes chrétiens les plus récents. Nous pensons, en effet, que, même si la littérature de la fin de la République Romaine ou du début de l'Époque Impériale a édifié, sous l'influence omniprésente des modèles grecs, certains monuments dont aucune autre langue n'offre d'exemple, c'est par l'Église que la voix de Rome a retenti pour l'Éternité jusqu'aux confins de l'univers.

On s'en convaincra d'après l'antithèse de ses deux prosateurs les plus féconds: Cicéron et Saint Augustin. Imbu de culture grecque, comme l'étaient ses contemporains romains du dernier siècle avant l'Ère Chrétienne, Cicéron a pour principal mérite, devant l'histoire de la Pensée universelle, d'avoir, par de nombreux néologismes, mis le latin à même de répéter les leçons des orateurs et des philosophes grecs.

Depuis lors, il n'est pas exagéré de dire que le monde entier l'a suivi dans cette voie! car, (ainsi que nous l'avons affirmé, in *CAD*, Vol XV, p.7): il n'est pas possible à un étranger de parler de science sans quelque trois cents mots grecs (empruntés, comme "analyse" à ἀνάλυσις; ou calqués, comme "qualité" sur ποιότης; cf. *infra*, §1.2.3).

Mais nous ne nous souvenons pas d'avoir rien rencontré dans Cicéron qui atteste un effort personnel pour se placer seul devant ce qui est, afin de tenter de le comprendre: noble effort qu'on appelle en arabe 'ijtihâd: اجتهاد!

Tout autre est Saint Augustin. L'indocile Africain confesse son impatience d'écolier dans les exercices de la langue grecque (in *Confessions*, Livre I, XIV); devenu Évêque d'Hippone (l'Annaba de l'Algérie d'aujourd'hui), s'il rend hommage aux Pères grecs, c'est sans prétendre être capable de les lire avec aisance: *græcæ autem linguæ non sit nobis tantus habitus...* "nous ne sachons pas si bien le grec..." (in *De Trinitate*, III, præmium); et il s'avance dans la confrontation des vérités reçues par la Foi: *fides quærens intellectum*: "La Foi entreprend de comprendre"...

C'est dans cette perspective qu'ont été choisis les textes saisis par nous pour être analysés: Monuments, en prose et en vers, des Lettres Classiques; Méditations et Discours du Docteur berbère et des auteurs chrétiens qui l'ont suivi; référence à la prose cicéronienne, inévitable pour moins que nous y prenions de goût!

1.2 Le corpus des textes

1.2.1 Latinitatis Monumenta

Le premier chant de l'Énéide

Sous la conduite d'Énée, quelques Troyens échappés à la destruction de leur cité, voguent joyeux, vers l'Italie qu'ils croient être le siège promis d'un empire à venir. Une tempête les bouscule vers Carthage, où règne la Phénicienne Didon. Quand se termine le premier chant de l'épopée de Virgile, le soir tombe sur un banquet: Énée entreprend de faire à Didon le récit des infortunes des Troyens:

Infelix Dido, longumque bibebat amorem.

Malheureuse Didon! elle boit l'amour à longs traits.

Finalement, Énée abandonnera Didon. Selon Virgile, la reine périt dans les flammes. Mais on peut imaginer une autre issue: nous attendons qu'un chef africain, se donnant pour le fils de la Phénicienne et du Troyen, reçoive d'un poète le droit légitime de régner sur tout le bassin de la Méditerranée...

Comme Énée, Virgile est pieux: non qu'il croit à son propre récit payen peuplé de dieux et de déesses; mais témoin des guerres civiles, Virgile croit que le poids de la misère humaine n'est pas vain, étant promis à une mystérieuse grandeur:

Tantæ molis erat Romanam condere gentem.

Tant dut coûter de peine,
Ce long enfantement de la grandeur romaine.

Ainsi traduit Delille, non sans élégance, dans une langue qui n'a pas l'impérieuse concision du Latin.

L'Art Poétique d'Horace

Les Français ont chez Boileau un écho du modèle latin. Mais Horace n'est classique que par la sobriété du ton: il sacrifie l'unité du plan à la vivacité du détail. Et son entretien didactique s'interrompt brusquement sur cette image d'un mauvais poète, acharné à se faire entendre:

Non missura cutem, nisi plena cruoris hirudo.

comme une sangsue, qui ne lâchera prise qu'après s'être gorgée de sang.

Le chant premier des métamorphoses d'Ovide

Comme Musset, Ovide est gentilhomme; comme La Fontaine, il a plus écrit que ce qu'il est permis aux honnêtes gens de lire! Mais Ovide a de la délicatesse, il n'est pas de ces élégiaques vulgaires qui semblent n'avoir vécu que pour monnayer, en vers, le scabreux récit de leur vie.

Le premier chant des Métamorphoses propose une cosmogonie qui débute au cahos de la matière naissante, s'illumine à l'âge d'Or, dégénère jusqu'à l'âge de Fer, dont nous déplorons de ne pas voir l'issue... Ovide sait prendre tous les tons et traite avec aisance tous les sujets; mais il ne peut longtemps se retenir d'être plaisant. Io, quelque temps métamorphosée en génisse et redevenue fille, s'abstient de parler, par crainte de mugir: Ovide philosophe n'a de cesse qu'il se reprenne à badiner.

De Rerum Natura: la physique selon Lucrèce; chant premier

Lucrèce a en commun avec Ovide qu'exposant à partir des atomes la composition des choses, il prend Vénus pour patronne:

Te sociam studeo scribendis versibus esse.

C'est toi, lui dit-il, que je veux pour compagne en écrivant ces vers! Mais le farouche disciple d'Épicure fulmine contre toute religion. On s'étonnera que l'éditeur de son orageux poème, ne soit autre que Cicéron, l'orateur aux amples périodes.

Le sixième chant de la Pharsale de Lucain: la bataille

Mort avant l'âge de trente ans, Lucain lègue à la postérité le récit épique d'une guerre civile où périt ce que, par un faux sens qu'autorise l'usage, on appelle l'antique Liberté de la République Romaine — Mais y-eut-il jamais un sens qui ne soit faux, quand ceux qui se partagent le pouvoir veulent faire accroire à d'autres hommes qu'ils en sont les mandataires...

Les poètes romantiques ont traité de tels sujets dans toutes les langues de l'Europe moderne: aucun d'eux n'a, comme Lucain, montré, sans fausses couleurs, l'issue d'un combat se décidant sur l'âpre vigilance ou la lassitude des chefs et des militants.

Le livre I des Commentaires de la Guerre des Gaules

Tous les Français savent, d'après César, que la Gaule était divisée en trois parties; et la gloire de Vercingétorix survit au naufrage des études classiques dans maints personnages colorés dont les noms riment avec le sien.

Qui relit, sans chauvinisme, les Commentaires, découvre entre des peuples frères, mais dissociés, des combinaisons politiques fragiles, avec des exploits d'héroïsme qu'amplifie jusqu'à l'absurde l'éloquente fierté des chefs Gaulois ou Germains; tandis que César, qui a tout vu, use de tout selon son projet; confiant dans la patience du légionnaire romain, tantôt robuste terrassier, tantôt combattant raidi par la parole du chef; et toujours endurent, parce que, prosaïquement, il a l'espoir d'une retraite.

Il n'y a plus de César: mais il y a toujours des Gaules en certains lieux.

Trente cinq chapitres de l'Histoire de la Guerre de Jugurtha, par Salluste

Après la chute de Carthage, Rome administra directement, comme une province, le territoire de l'actuelle Tunisie; et son pouvoir s'étendit graduellement vers l'Est, par voie de protectorat et d'alliance. Vers 160 avant notre ère, dans la lignée de Masinissa, roi berbère allié des Romains, naquit Jugurtha; qui, par sa bravoure, son intelligence, son charme auquel des libéralités princières rendaient sensibles les héritiers déjà corrompus de ces consuls qu'a magnifiés Tite-Live, crut pouvoir fonder un empire.

On a saisi le texte jusqu'à l'apostrophe fameuse de Jugurtha, contraint de quitter Rome, où ses intrigues ont dégénéré en violence ouverte:

Urbem venalem et mature perituram, si emptorem invenerit!

Ville vénale, et qui périrait bientôt, si elle trouvait un acheteur!

La Vie d'Agricola, par Tacite

À la fin du premier siècle de notre ère, l'île qui porte aujourd'hui le nom de Grande Bretagne n'est qu'en partie comprise dans l'Empire: Agricola, beau-père

de Tacite, s'illustre en affermissant l'emprise de Rome. Face au chef, dont la piété filiale lui commande d'exalter la conduite, l'historien ne se fait pas faute de figurer les *Bretons*, non comme des bêtes féroces, mais comme des héros épris de liberté. Rome fut grande parce qu'elle ne dénia jamais la qualité d'hommes à ceux qu'elle vainquit, soumit et, finalement reçut au nombre de ses citoyens.

1.2.2 O Roma felix

Le livre I des Confessions de Saint Augustin

C'est comme l'auteur des confessions que Saint Augustin est connu de tous: ce saint eut la jeunesse d'un pécheur, et d'aucuns imaginent qu'il a, de ses péchés, laissé le récit complaisant. Mais dans la trame de sa vie, Augustin suit le fil de sa conscience; sa découverte du libre arbitre de l'homme, connu de toute éternité dans la préscience divine. Quand s'achève le livre I, le futur évêque pratique encore, dans sa ville natale, des exercices de rhétorique dont l'Énéide fournit le thème.

Le livre V de La Cité de Dieu

La Cité de Dieu accomplie, n'est autre que le Ciel même habité par les saints: mais, dans le cours du temps, la Cité de Dieu, qui vit de la foi, voyage entre les impies:

inter impios peregrinatur ex fide vivens;

ce voyage c'est l'histoire, que Dieu, selon les raisons de sa providence, conduit jusqu'au terme. Au livre V, Augustin remémore la grandeur terrestre de Rome, rétribuée par Dieu sur terre de ce qu'elle a fait pour gagner la terre. Dans la vie d'un empire, comme dans une seule vie d'homme, la volonté humaine déroule librement ce qui, pour Dieu, tient en un point:

Non ergo propterea nihil est in nostra voluntate, quia Deus præscivit quid futurum esset in nostra voluntate.

Que Dieu sache, par avance, ce qui sera dans notre volonté n'implique pas qu'il n'y ait rien dans notre volonté.

Sancti Aurelii Augustini, hipponensis episcopi, De TRINATE Liber primus

Toute religion entend conduire les hommes à Dieu, par l'obéissance à des préceptes. La religion catholique, pour distribuer des sacrements qu'elle affirme être divins, doit enseigner des mystères: dire, en quelque manière, aux hommes, des choses qu'ils ne peuvent comprendre, mais doivent savoir... Voilà pourquoi, entre tant d'œuvres de Saint Augustin, le Traité de la Trinité est celui, par excellence, qu'il s'est préparé à écrire toute sa vie: il n'y a qu'un seul Dieu, et il y a trois personnes divines; sans ces deux assertions à la fois, le christianisme ne serait qu'une velléité de Salut.

Sancti Gregorii Magni, romani pontificis, *Moralium Liber primus*

Mort en 603, Saint Grégoire le Grand, moine et pontife, est le pape de l'an 600. Connu de tous pour être le législateur du chant grégorien, Grégoire I a laissé, dans son commentaire du Livre de Job, l'exposé par excellence de la Morale Chrétienne:

Nam primum quidem fundamenta historiae ponimus; deinde per significationem typicam in arcem fidei fabricam mentis erigimus; ad extremum quoque per moralitatis gratiam, quasi superducto ædificium colore vestimus.

Le Commentateur donne d'abord le sens littéral d'un passage du Livre sacré, en tant qu'il décrit la vie du patriarche Job; il relit ensuite ce même passage comme une prophétie appliquée au Christ et à l'Église; il en tire enfin une morale par une analogie qui n'est pas étrangère à celle des paraboles de l'Évangile.

Il nous a paru que cette multiplicité d'interprétation pourrait engendrer une multiplicité de style. Quant à la langue, le pontife romain n'a pas la rudesse de son homonyme et contemporain, l'évêque de Tours; mais, dans l'épître d'introduction aux *Moralia*, il confesse, non sans affectation, n'avoir point visé à une correction grammaticale, que, dit-il, n'ont pas observée les traducteurs eux-mêmes de la Bible en latin!

De consideratione ad Eugenium Papam, de Saint Bernard, (præfatio et Liber primus)

Le Pape qui, vers 1150, règne à Rome, sous le nom d'Eugène III, a été, naguère, à Clairvaux, moine et disciple de Saint Bernard. Celui-ci lui écrit, avec le respect d'un fils et l'affection jalouse d'un père, pour l'avertir de préférer, en son cœur, à l'action, la *considération*, c'est-à-dire la méditation; et de ne point se laisser absorber par la charge du gouvernement :

Magna virtus patientiæ; sed minime hanc tibi ad ista optaverim. Interdum impatientem esse, probabilius.

C'est une grande vertu que la patience; mais je ne te souhaite pas de l'avoir en cette matière: parfois, être impatient est plus digne de louange.

Nam fructus horum quid, nisi aranearum telæ?

Car le fruit de ces occupations qu'est-ce, sinon des toiles d'araignées?

L'encyclique *Jucunda sane* de Saint Pie X

Treize cents ans après le départ de Saint Grégoire le Grand, son successeur, nouveau venu sur le siège de saint Pierre, adresse à ses frères dans l'épiscopat une encyclique *sur la responsabilité de ceux qui gouvernent l'Église*. Souvent cité, et admirablement peint, l'auteur des *Moralia* nous est rendu par cette lettre, dont l'éloquence est cependant toute classique.

1.2.3 Eloquentiæ flumen

La première Catilinaire de Cicéron

Quousque tandem abutere, Catilina, patientia nostra?
Jusques à quand, Catilina, abuseras-tu de notre patience?

Ceux mêmes qui d'un long apprentissage du latin n'ont retenu qu'une phrase, reconnaîtront ces premiers mots de l'exorde cicéronien prononcé devant les Sénateurs, le huit Novembre de l'an 63 avant l'incarnation. Moins de deux mois plus tard, Catilina, parti de Rome, était battu et tué à Pistoia; mais le système politique que Cicéron croyait avoir sauvé ne survécut guères à Catilina.

Huit lettres de Cicéron au proconsul Lentulus

Les lettres qui nous restent de Cicéron ont été groupées en 47 livres. Le premier Livre des *Lettres familières* comprend dix lettres, dont huit adressées au Proconsul Lentulus Spinther alors gouverneur de Cilicie (56-53), destiné à périr, après la bataille de Pharsale, avec d'autres chefs du parti pompéien vaincu. La première phrase de la quatrième lettre suffit à donner le ton:

Au 16-ème jour des calendes de février notre position dans le sénat était très solide: nous avons, la veille, brisé la proposition de Bibulus sur les trois députés, et n'avions plus qu'un combat à livrer, contre la proposition de Volcacijs; mais nos adversaires, par toutes sortes de chicanes, traînèrent les choses en longueur...

Les intrigues des assemblées n'étaient pas plus édifiantes il y a vingt siècles qu'elles ne le sont aujourd'hui; et il ne nous paraît pas que Cicéron, dans ces comptes rendus, s'élève très au-dessus de la mêlée.

Le songe de Scipion

Cum in Africa venissem, M'. Manilio consuli ad quartam legionem tribunus (ut scitis) militum; nihil mihi potius fuit, quam ut Masinissam convenirem, regem familiæ nostræ iustis de causis amicissimum...

Arrivé en Afrique, sous le consul M. Manilius, comme tribun de la quatrième légion (ainsi que vous le savez), je n'eus rien de plus cher que de visiter Masinissa, ce roi, pour de bien justes raisons, très cher à notre famille...

Comblé par la chaleur de l'hospitalité berbère, le petit fils de Scipion l'Africain quitte enfin la table. Un profond sommeil l'envahit; son grand-père, puis son père, lui apparaissent en songe; il entend une leçon de philosophie cosmologique et de morale. Telle est la fiction de Cicéron!

de Finibus Bonorum et Malorum ad Brutum Liber Primus

Non eram nescius, Brute, cum, quæ summis ingeniis exquisitaque doctrina philosophi Græco sermone tractavissent, ea Latinis Litteris mandarem, fore, ut hic noster labor in varias reprehensiones incurreret.

En confiant aux lettres Latines ce qu'avaient dit en Grec des philosophes dotés d'esprits sublimes et à la science achevée, je savais bien, Brutus, que notre entreprise encourrait divers reproches...

Cicéron n'entreprend pas moins l'exposé des doctrines du Bien et du Mal. À sa critique de ce qu'Épicure a fait, notamment, de l'atomisme de Démocrite, Torquatus répond en défendant en tout l'Épicurisme. La forme du livre serait celle d'un dialogue, si Torquatus, prenant la parole, n'avait exprimé cette préférence, tout autre que maïeutique:

... *uti oratione perpetua malo, quam interrogare aut interrogari.*

...j'aime mieux parler sans m'arrêter que de poser des questions et d'en recevoir.

Marci Tullii Ciceronis Academicorum Posteriorum Liber Primus

Le nom d'Académie est donné à une suite d'Écoles philosophiques, depuis l'Ancienne Académie de Platon, jusqu'à la Moyenne et à la Nouvelle École d'Arcésilas, de Carnéade et de leurs successeurs; lesquels, de l'aptitude dialectique à tout prouver et tout combattre, passèrent au scepticisme déjà contenu en germe chez Socrate.

Dans le fragment qui nous reste du dialogue de l'*Ancienne Académie*, Cicéron prône d'abord la traduction en Latin de la philosophie des Grecs: il explique qu'il a calqué *qualitas* sur ποῦότης; dérivant celui-là de *qualis*, "quel", comme celui-ci l'est de ποῦός. Puis, abordant les doctrines des Stoïciens et des Académiques, il va jusqu'à la thèse du scepticisme.

Saint Augustin a écrit un dialogue *Contre les Académiques*; et quand il revendique pour l'homme la faculté de connaître ou de vouloir, sans déroger à la prescience divine, notamment dans la première partie du Livre V de *la Cité de Dieu*, il emprunte à Cicéron l'exposé des doctrines classiques.

1.3 De la saisie du texte aux tableaux de correspondance

1.3.1 Division en fragments chapitres et versets

La méthode suivie ayant déjà été exposée dans [LING. TRI], [TEXTES GRECS] et [LING. TRI 2], on se bornera ici à de brefs rappels et à des compléments.

Comme dans la saisie du texte grec du Nouveau Testament, chacun des textes a été subdivisé en chapitres et "versets" ou paragraphes, fixés ici par nous afin de retrouver aisément les occurrences des formes et de diviser, au besoin, les textes en fragments ne coïncidant pas nécessairement avec les chapitres des éditions usuelles (cf. [TEXTES GRECS], §1.3). Les programmes de mise en

forme et de tri, forlat§5 et trilat§5, reconnaissent les subdivisions d'après les numéros portés en tête des versets et vérifient que le numérotage se fait sans omission, reprenant seulement à 1 au seuil de chaque chapitre (cf. [LING. TRI 2], §1.2).

L'ensemble des textes de notre corpus est ici traité comme un texte unique, noté ici 'Qnz', dont des versions successives de plus en plus longues ont été créées, au fur et à mesure que nous progressions dans la saisie. Les résultats des analyses ne s'étant jamais contredits, nous n'en décrivons pas l'enchaînement. Pour subdiviser 'Qnz' en chapitres ou fragments, désignés par des sigles en évoquant le contenu, on a créé un fichier de subdivision 'Qnz§Ax', (cf. [TEXTES GRECS], §1.3). Nous ne publions pas ce fichier, nous bornant à donner les règles observées dans le choix des subdivisions.

Il faut d'abord que les fragments soient assez longs pour que leur profil lexical ne soit pas si fortement marqué par les fluctuations du style (notamment la répétition rhétorique), qu'un ou deux fragments ne créent un axe factoriel... Il est rare qu'un fragment descende en dessous de 500 occurrences; et pourtant nous rencontrerons quelques cas de profils aberrants.

On préférera, d'autre part, que les limites entre fragments coïncident avec certaines des limites entre chapitres adoptées dans les éditions usuelles. Et, surtout, puisque notre but est de découvrir une influence du genre ou du style sur le choix des formes, on évitera de faire enjamber un fragment sur le passage du récit au discours dans un texte historique, ou de l'exposé littéral à l'allégorie dans un commentaire tel que celui de Saint Grégoire le Grand sur JOB.

Ces deux principes étant suivis, ou, du moins, conciliés, on s'appliquera à créer autant de fragments qu'il est possible, afin d'éprouver l'homogénéité de chaque œuvre ou de trouver des convergences de style entre passages homologues d'œuvres différentes.

Dans le présent travail, on considère un corpus de dix-neuf œuvres divisées en 234 fragments; dont le nombre, pour une seule œuvre, va de quatre, pour le *Songe de Scipion*, à trente et un, pour l'épître d'envoi, la préface et le Livre I des *Moralia in Job*.

1.3.2 Le lexique Δ des formes de mots outil

On sait qu'une fois fixés le texte 'Qnz' et le fichier 'Qnz§Ax' qui en délimite les fragments, il reste à choisir un dictionnaire 'Δ', ou liste des formes dont les occurrences seront dénombrées, par fragment, afin de construire le tableau de correspondance de base entre ensemble I des fragments et ensembles J des formes.

Le choix de 'Δ' repose sur la liste, ordonnée par fréquence croissante, de toutes les formes se rencontrant dans le corpus. Il est clair que les hapax et

prépositions: a, ab, ad, ante, cum, de, e, ex, in, inter, per, post, pro, propter, sine, sub, secundum;

formes du pronom relatif ou interrogatif: cui, cuius, qua, quae, quam, quem, qui, quibus, quid, quis, quo, quod, quorum, quos;

formes d'autres pronoms non personnels et adverbess pronominaux: ea, eius, eo, eorum, eos, eum, hac, haec, hic, hinc, his, hoc, id, illa, ille, illi, illo, illud, ipsa, ipse, ipsum, ista, multa, omnes, omnia, omnibus, omnium, se, sibi, sua, suis, suo;

conjonctions de coordination: ac, atque, aut, ergo, et, igitur, nam, sed, tamen, vel, vero;

conjonctions de subordination: an, cum, dum, nisi, si, sicut, ut;

adverbes (de liaison, à valeur conjonctive, etc.): autem, contra, etiam, iam, ita, magis, nunc, quam, quasi, quidem, quippe, quoque, sic, tam, tum, ubi, unde, ut;

formes de la troisième personne ou formes impersonnelles, des verbes être, pouvoir, ...: erat, erit, esse, esset, est, fuit, posse, potest, sit, sunt;

négations: ne, nec, neque, nihil, non, nulla;

LE LEXIQUE Δ DES FORMES DE MOTS OUTIL

même les formes de faible fréquence n'ont pas de place dans Δ ; mais, une fois choisi un seuil de fréquence minima, il reste à choisir parmi les mots situés au-dessus du seuil. Ici, (à la différence de ce qu'on a fait dans [LING. TRI 2],) on a retenu exclusivement des formes de mots outil: éliminant non seulement les mots pleins proprement dits: *deus*, 'Dieu', *hominum*, 'des hommes' (génitif)...; mais aussi *res*, 'chose', *ait*, 'dit-il', ainsi que toute forme, verbale ou pronominale, renvoyant à la première ou à la deuxième personne.

Après avoir effectué de nombreuses expériences, nous considérerons ici un lexique ' Δ ' de 119 formes; dont la moins fréquente, *erit*, 'il sera', se rencontre 80 fois dans le corpus; ce qui correspond à un seuil de fréquence fixé au tiers du nombre des fragments (qui est de 234). Ainsi, se trouve dans ' Δ ' la préposition *secundum*, 'suivant', 'selon',..., dont la dialectique fait grand usage et qui est un maître mot de la scholastique.

1.3.3 Formes et morphèmes

Il importe de noter les difficultés rencontrées du fait des variations de l'orthographe, des polysémies; et l'élargissement apporté à la notion de forme.

Tous les étudiants savent que les éditeurs de Salluste, adoptent une orthographe dite archaïque; laquelle se distingue, principalement par le timbre de quelques voyelles: *existumet*, pour *existimet*, 'qu'il estime'; orthographe qu'on abandonne d'ailleurs quand l'historien interrompt son récit pour rapporter un discours, comme si l'archaïsme n'était pas de mise sur les lèvres d'un orateur public. Nous avons rétabli l'orthographe usuelle, bien que les variantes n'affectent guère que des mots pleins, non dénombrés ici.

Une autre divergence naît de l'orthographe étymologique qui, notamment, élimine des doubles lettres: *inlustris*, pour *illustris*. Nous avons toujours suivi le lexique de Gœlzer, qui fait peu de part à l'étymologie.

Orthographe et polysémie sont liées. Ainsi, pour le datif-ablatif pluriel du pronom relatif, la forme la plus répandue est *quibus*; mais on rencontre *quis*, ou *queis*. Or *quis* est également le nominatif masculin du pronom interrogatif. Nous avons décidé de rétablir partout *quibus*, en y adjoignant, s'il y a lieu, entre parenthèses, *queis*, dont les occurrences ne sont pas dénombrées. La forme *cum*, a de nombreuses fonctions: celles de la préposition française 'avec'; celles de la conjonction de subordination 'comme', 'puisque'...; et, en association avec *tum*, *cum* sert à marquer le parallélisme. Or nous croyons que la variante *quum* n'est jamais usitée dans le sens d' 'avec'. Cependant, présentement nous avons mis partout *cum*, avec un accent *cùm* (que les programmes d'analyse ne prennent pas en compte) pour signaler que l'original est *quum*.

Nous avons borné l'élaboration préalable des textes au minimum requis pour affranchir l'analyse statistique des artéfacts orthographiques; non seulement afin de préserver l'automatisme de la méthode; mais aussi parce que la résolution des polysémies entraîne dans des ramifications infinies, auxquelles on hésite à assigner un terme; comme l'atteste le dénombrement des sens dans les grands dictionnaires...

Certains mots latins ont deux formes; dont les dictionnaires affirment qu'elles ont strictement le même sens, leur emploi étant déterminé phonétiquement par le contexte (comme pour les prépositions *a* et *ab*; *e* et *ex*): on peut, dans ce cas, dénombrer séparément les deux formes. Mais parfois, une liberté de choix est laissée à l'auteur: ainsi, pour *ac*, on spécifie "devant consonne seulement", tandis que *atque* s'emploie dans tout contexte. Or il se trouve que Tacite, dans sa vie d'Agricola, manifeste une prédilection pour *ac*. Il nous a paru bon de vérifier que la forte originalité du texte devait peu à cette prédilection: dans les analyses rapportées ici, les occurrences de *ac* et *atque* sont cumulées sous le sigle 'atqc'.

Outre les formes *strito sensu*, on souhaiterait dénombrer des morphèmes, la distinction entre formes et morphèmes dépendant, d'ailleurs, parfois, de la tradition orthographique, basée sur l'accentuation. Ainsi, le latin a une particule de coordination postposée, *que*, qui a le sens de 'et', comme dans *cælo terraque*, 'dans le ciel et sur la terre'. On a dénombré sous le sigle 'qve' les syllabes *que* se trouvant en fin de mots; en retranchant du compte les occurrences de {*atque*, *quoque*, *neque*}; ce qui laisse passer *quisque*, 'chacun', *æque*, 'également', etc..., où la syllabe finale *que* n'a pas valeur de coordination.

Nous avons également compté certaines désinences bien caractéristiques: {*ibus*, *orum*, *arum*}, du datif-ablatif ou du génitif pluriel; en définitive, seul *ibus* nous a servi à caractériser le style; plus précisément, sous le sigle 'ibu]', on dénombre les finales en *ibus*, en retranchant *quibus* et *omnibus*, comptés séparément.

Il nous est rapidement apparu que la rareté des mots outil est l'un des caractères les plus manifestes de la langue des poètes, ou, à un moindre degré, des historiens. Même si cette rareté, toute relative, laisse encore matière à nos dénombrements, on a voulu la prendre en compte pour elle-même. Dénombrer globalement toutes les occurrences de formes non comprises dans le lexique 'Δ', aurait donné à la colonne '≠Δ', ainsi créée, un poids triple environ de celui de toutes les autres colonnes prises ensemble. On a choisi d'introduire les deux variables 'πmot', nombre de mots divisé par 10, et 'πltr', nombre de lettres divisé par 100.

On appréciera, d'après leur corrélation avec le premier axe, l'importance des diverses variables ainsi relevées, concurremment avec les formes de 'Δ'.

2 Discrimination entre œuvres et classification des fragments

2.0 Les méthodes taxinomiques

Ici encore, on se borne à un bref rappel, renvoyant pour plus de détails au §2 de [TEXTES GRECS], dont on reprend ici les termes.

Procédant par pure induction, la Classification Ascendante Hiérarchique (CAH) constitue d'abord, à partir de l'ensemble des fragments, 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 peut retrouver dans ces agrégats des textes ou des groupes de textes relevant du même genre; voire éclairer et perfectionner la taxinomie stylistique.

Si l'on ne reconnaît pas de groupe acceptable, le statisticien pourra, dans son analyse, tenir compte explicitement de la répartition en œuvres de la totalité, ou seulement de la moitié (des deux tiers,...) des fragments; et, d'après ces données, définir pour chaque œuvre un centre, ou point moyen des fragments qu'on sait y être compris; l'affectation d'un fragment à une œuvre se faisant en déterminant le centre qui en est le plus proche. C'est l'Analyse Discriminante.

Si l'œuvre n'a été prise en compte que pour une moitié des fragments, le taux d'erreur résultant de la règle d'affectation sera généralement plus élevé pour l'autre moitié. Or c'est précisément sur cette autre moitié (ou ce tiers restant; appelé, pour cette raison: échantillon d'épreuve) que l'analyse discriminante peut faire la preuve de son efficacité; laisser espérer qu'à l'avenir, il soit possible, sur des fragments isolés, de se fier au statisticien pour reconnaître le style, voire l'auteur; et, plus encore, donner du style, du genre, une définition générale, en termes d'emploi des formes, sans recourir au contenu des textes ni au sens littéraire des humanistes.

Dans la présente étude, la CAH donne de l'ensemble des fragments, une vue cohérente: agrégeant les fragments en classes qui correspondent, à peu près, à des œuvres, et ordonnant celles-ci selon leurs genres. Mais, pour plus de clarté, nous débiterons l'exposé par l'analyse discriminante.

19 oeuvres x 119 formes et autres variables

trace :	2.706e-1									
rang :	1	2	3	4	5	6	7	8	9	10
lambda :	688	411	374	210	163	144	131	104	90	77 e-4
taux :	2543	1519	1383	775	602	532	482	385	333	283 e-4
cumul :	2543	4062	5446	6220	6822	7354	7837	8221	8555	8838 e-4

2.1 Correspondance entre œuvres et formes

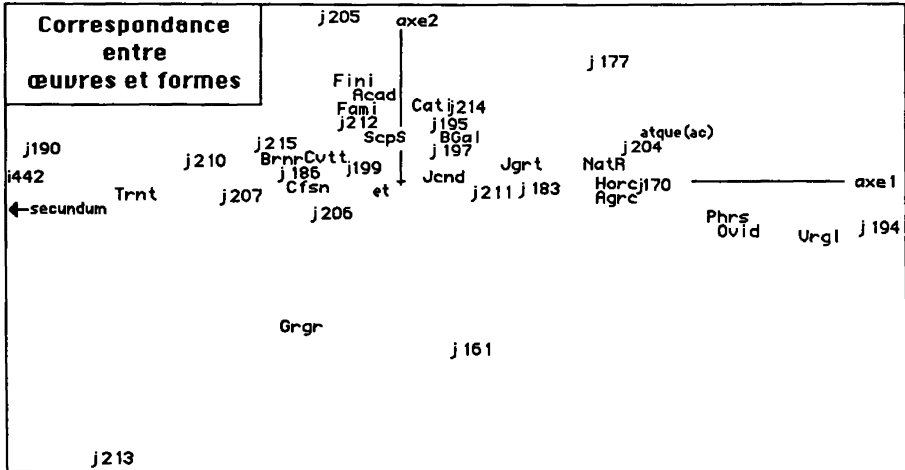
SIGI	QLT	PDS	INR	F 1	COR	CTR	F 2	COR	CTR	F 3	COR	CTR	F 4	COR	CTR
Trnt	790	105	113	-414	589	263	-37	5	4	216	160	131	102	36	53
Phrs	699	35	53	465	532	111	-125	39	13	218	117	45	66	11	7
Ovid	759	37	60	480	519	123	-132	40	16	296	198	86	30	2	2
Vrgl	824	33	65	601	687	175	-171	56	24	204	79	37	27	1	1

SIGJ	QLT	PDS	INR	F 1	COR	CTR	F 2	COR	CTR	F 3	COR	CTR	F 4	COR	CTR
enim	853	8	14	-540	619	34	10	0	0	218	100	10	-250	132	24
quia	967	8	37	-682	362	52	-857	572	138	-93	7	2	-183	26	12
scnd	640	2	26	-1297	384	39	-69	1	0	889	180	34	572	75	25
πltr	883	112	8	114	677	21	-48	123	6	-24	32	2	31	51	5
πmot	941	196	8	94	768	25	-34	100	6	6	3	0	26	57	6
qve	932	43	96	703	824	312	-117	23	15	218	79	55	-58	6	7

À partir du tableau 234 × 119, croisant les 234 fragments avec les formes, on construit, par cumul des blocs de lignes afférents à chaque œuvre, un tableau 19 × 119, croisant les 19 œuvres avec les formes. On analyse le tableau des œuvres 19 × 119 comme principal, avec, en supplément, le tableau 234 × 119, qui décrit les fragments. L'analyse fournit 18 axes factoriels, et le tableau des valeurs propres montre la prépondérance de l'axe 1, puis des axes 2 et 3 sur ceux qui les suivent.

mots dans les oeuvres et dans les cap. du de <i>Trinitate</i>															
48	Trnt	Brnr	Horc	Vrgl	Grgr	Jgrt	Cfsn	Phrs	Cvtt	Agre					
	Ovid	Jcnd	NatR	BGal	Acad	Cati	Fami	Fini	ScpS						
	Tr01	Tr02	Tr03	Tr04	Tr05	Tr06	Tr07	Tr08	Tr09	Tr10					
	Tr11	Tr12	Tr13	Tr14	Tr15	Tr16	Tr17	Tr18	Tr19	Tr20					
	Tr21	Tr22	Tr23	Tr24	Tr25	Tr26	Tr27	Tr28	Tr29						
scnd	70	5	0	0	2	1	3	0	10	0					
	0	0	0	1	2	0	0	1	0						
	4	0	0	1	0	0	1	1	0	1					
	1	0	2	2	3	0	0	0	0	1					
	1	14	10	1	5	6	9	1	6						

Sur le demi-axe (F1<0), scnd, *secundum*, se détache très à l'écart, associé au de *Trinitate* (Trnt) de Saint Augustin. Employée surtout dans certains chapitres de Trnt, la préposition scnd se rencontre dans les autres œuvres de Saint Augustin, de *Civitate Dei*, Cvtt, *Confessions*, Cnfs; mais aussi ailleurs, notamment dans l'*Academicorum liber* de Cicéron où figure deux fois l'expression *secundum naturam*, 'conforme à la nature', opposée à *natura contraria*, 'contraire à la nature', pour exposer une distinction faite par Zénon. D'ailleurs, scnd, ne crée aucunement l'axe 1 (CTR1=39) et n'est pas la variable la mieux corrélée avec cet axe: COR1(scnd) = 384 ; COR1(qve) = 824.



Nous étant ainsi justifié d'avoir conservé un terme qui mûrit pour l'usage scholastique, interprétons l'axe 1: de Trnt à Vrgl, de la dialectique théologique à la poésie classique, les textes marquent une gradation en passant par la rhétorique de Saint Augustin et de Cicéron: {Cfns, Cvtt}, {Fini, Fami, Acad, ScpS}, puis Cati; et l'histoire. La première Catilinaire est la seule œuvre proprement rhétorique du corpus; tandis que les dialogues philosophiques de Cicéron n'ont que partiellement la forme dialectique.

La concision de nos historiens préférés les rapproche des poètes! on se plaît à trouver l'*Agricola* de Tacite un peu au-delà de la poésie didactique du *de Natura Rerum*, dans la direction de Virgile. Quant à la place d'Horace au même niveau, il faut souligner que le grand lyrique est ici représenté par l'Art Poétique, dont le ton ne peut aussi manquer d'être didactique.

L'encyclique *Iucunda sane*, ne s'écarte pas du centre de gravité: cette œuvre dont la diversité de ton est sensible à la lecture, suit avec élégance les modèles classiques, sans adopter de style propre.

De toutes les variables, la plus corrélée à l'axe 1 est la coordination postposée *qve*, dont, au §1.3.3, on a expliqué le dénombrement. Il faut noter que, dans des analyses préliminaires, la syllabe finale 'que' s'est montrée moins caractéristique de la concision poétique que ne l'est 'qve', d'où sont retranchés {quoise, neque, atque}.

Après *qve*, on trouve le plus corrélés à l'axe 1 { π ltr, π mot} qui n'apportent cependant pas de forte contribution à l'axe, s'écartant peu de l'origine. Nous avons dit que la prépondérance des mots pleins sur les outils est de règle chez le poète; ce qui nous a incité à créer les variables { π ltr, π mot}.

sans échantillon d'épreuve	échantillon d'épreuve d'un tiers
-> Trnt : 23/29	-> Trnt : 17/20 ; 5/ 9 ;
-> Brnr : 12/12 + Tr03 Tr04 Tr06	-> Brnr : 8/ 8 ; 3/ 4 ; Tr04 ; Tr06
-> Horc : 7/ 7 + Phxv	-> Horc : 4/ 4 ; 3/ 3 ; ; Phxv
-> Vrgl : 8/ 8	-> Vrgl : 6/ 6 ; 2/ 2 ; ; Phr6 Phrx
-> Grgr : 31/31	-> Grgr : 20/20 ; 11/11 ;
-> Jgrt : 16/16	-> Jgrt : 11/11 ; 5/ 5 ; ; Tr12
-> Cfsn : 12/12 + Tr08	-> Cfsn : 8/ 8 ; 4/ 4 ; Tr08 ; Tr03
-> Phrs : 10/11	-> Phrs : 7/ 7 ; 1/ 4 ;
-> Cvtt : 20/20 + Tr02 Tr05	-> Cvtt : 14/14 ; 6/ 6 ; Tr05 ; Tr09 Br07
-> Agrc : 10/10	-> Agrc : 6/ 6 ; 4/ 4 ;
-> Ovid : 7/ 7	-> Ovid : 5/ 5 ; 2/ 2 ;
-> Jcnd : 7/ 7	-> Jcnd : 5/ 5 ; 2/ 2 ;
-> NatR : 13/13	-> NatR : 8/ 8 ; 5/ 5 ;
-> BGal : 15/15	-> BGal : 10/10 ; 5/ 5 ;
-> Acad : 8/ 8	-> Acad : 6/ 6 ; 2/ 2 ;
-> Cati : 6/ 6	-> Cati : 4/ 4 ; 2/ 2 ;
-> Fami : 6/ 6	-> Fami : 4/ 4 ; 2/ 2 ;
-> Fini : 12/12	-> Fini : 8/ 4 ; 4/ 4 ; ; Scp4
-> ScpS : 4/ 4	-> ScpS : 2/ 2 ; 1/ 2 ;

2.2 Analyse discriminante et classification des fragments

Dans l'espace de dimension 18 où l'analyse du §2.1 a tracé les axes factoriels, il a y un nuage principal, I, des œuvres (représentée chacune par son profil moyen); ainsi qu'un nuage, Is, des fragments, mis en supplément, et aussi le nuage J des variables (formes outil et autres cumuls). Ainsi qu'on l'a dit au §2.0, l'analyse discriminante rattache chaque fragment *is* au centre *i* dont il est le plus proche. Le tableau de gauche fait le bilan de ce qui est ainsi attaché à chaque centre. Par exemple, sur la 1-ère ligne, on lit: -> Trnt : 23/29; ce qui signifie qu'au centre Trnt sont seulement attachés 23 des 29 chapitres du Livre I du *de Trinitate*. Sur la ligne 2, -> Brnr : 12/12 + Tr03 Tr04 Tr06, signifie qu'au centre défini par le *de Consideratione* de Saint Bernard, sont attachés, outre les 12 chapitres de l'œuvre, les chapitres {3, 4, 6} du *de Trinitate*; etc...

Afin de confirmer l'excellente discrimination dont rend compte le tableau de gauche on a fait une analyse, semblable à celle du §2.1, mais en définissant le centre de chaque œuvre d'après les deux tiers de ses fragments; un tiers restant en échantillon d'épreuve. Puis on a rattaché les chapitres aux centres. Le bilan, donné par le tableau de droite, est complexe; les postes en sont séparés par des signes ';'. Prenons la ligne 2: -> Brnr : 8/ 8 ; 3/ 4 ; Tr04 ; Tr06. Au centre Brnr sont attachés: les 8 chapitres, (8/8), constituant l'échantillon de base du *de Consideratione*; 3 des 4 chapitres, (3/4), constituant l'échantillon d'épreuve; et les chapitres Tr04 et Tr06, appartenant, respectivement, à l'échantillon de base et à l'échantillon d'épreuve du *de Trinitate*.

Si l'on met à part le *de Trinitate*, œuvre complexe, le taux d'erreur, pour l'échantillon d'épreuve lui-même, est inférieur à 10%.

Revenant à l'analyse du §2.1, nous trouvons, dans la CAH, une autre voie pour saisir la structure d'ensemble des œuvres et en apprécier l'homogénéité.

Trnt: secundum++ j210+++ j194-

Trnt				34		36	
Brnr	194-	25	30	32		Brnr: 215++++	214+
Cfsn	et++++						
Cvtt	215++	28				Cvtt 207+	
Cati	215++	26				Cati Atq+ 204++	
Jcnd	206++					Jcnd = CdG	
ScpS	190+	24				ScpS 183+++ 212+	
Acad	190+					Acad 205++++ 212+	
Fami	186+++					Fami 205++	
Fini	214++					Fini 205+++	
Grgr	213+++++	206+	207+				
BGal	199++++			33		35	
NatR	194	177++	197++	31		NatR 204+	
Jgrt	195++	197+	29		Jgrt 170++		
Agrc	Atq+++						
Horc	214+++	27	194++		Horc et+		
Phrs		23	194++++				
Ovid			Ovid 194++++				
Vrgl			183+	194++++			CAH des oeuvres

La CAH des œuvres sépare d'abord, de tout le reste, la classe 35 des historiens et des poètes. Entre ceux-ci, le partage serait exact si le *de Natura Rerum* ne suivait les historiens, ainsi que le laissait prévoir l'examen de l'axe 1.

Saint Grégoire se sépare ensuite de la classe 34, qui réunit rhétorique et dialectique. Le *de Trinitate* étant à part, il reste la classe 32; dont Cicéron et Saint Augustin occupent respectivement les deux branches 24 et 30.

La CAH des fragments retrouve presque identiquement les résultats de la CAH des œuvres: les fragments s'agrègent en classes qu'on peut désigner par le nom d'une œuvre, en signalant seulement de rares exceptions: {Tr03 Tr04} sont dans Brnr; Tr08 dans Cfsn (confessions), Tr05 dans Cvtt (Cité de Dieu); un chapitre des Acad. suit le *de finibus*, du même Cicéron. Mais il vaut la peine de noter une subdivision du *de Trinitate*, i442, qui en comprend quatre chapitres.

Trnt: 210+++ 194- ; 442: secundum+++++ 210+++

Trnt		451			464		466
442			442 = {Tr23 Tr26 Tr27 Tr22}				
Brnr		454	460	462		Tr03 Tr04 dans Brnr	
Cfsn					Tr08 < Cfsn		
Cvtt		458			Tr05 < Cvtt		
Jcnd							
Cati							
ScpS		455					
Acad							
Fami							
Fini			Acad2 < Fini				
Grgr							
BGal				463		465	
NatR			461				
Agrc		459					
Jgrt							
Horc		457					
Phrs		453					
Ovid							
Vrgl							CAH des fragments

c	Partition en 23 classes : Sigles des formes de la classe c
212	a omnium quibus quam haec quae Acad+
186	sine ne ut de cum quod Fami+ Vrgl-
207	sed ipsa nisi sunt etiam vero cuius pro Grgr++ Cvtt++ Vrgl-
206	qui suo eius ad inter in Grgr++
161	post per Grgr++++

199	his sibi se eum suis eo eos eorum BGal+++++

205	ea quidem nihil esset id esse Vrgl- Acad++ Fami+ Fini++

170	erat ubi fuit Jgrt++++ Ovid+
195	omnibus ex neque BGal+ Jgrt+++ Cvtt-
7	atque(ac) Cati+ Agrc++++
177	nulla posse e Fini+ Acad+ NatR++++ Cvtt-
197	multa contra igitur quoniam magis tamen nam Jgrt++ NatR++++

183	hic quem sic ille Vrgl++ Ovid+ ScpS++
204	tum iam quorum omnes ante nunc at Vrgl++ NatR++ Cati+ Cvtt-
211	ibu ab quos sua quo πmot πltr =CdG Fl+++++
=====	
213	quasi unde hinc dum quippe ergo quia Grgr+++++
=====	
215	quis ista illa vel non tam hac cui si Brnr+++ Cvtt++
214	aut nec sit quid an Horc+++ Fini++
29	et Cfsn++

95	secundum Trnt+++++
190	enim autem Cfsn++ Acad+ Fini+
210	erit omnia ita illud ipsum qua illi ipse propter Trnt+++++
	potest hoc est illo sicut Trnt+++++
=====	
194	sub quoque qve Vrgl++ Phrs++ Ovid++

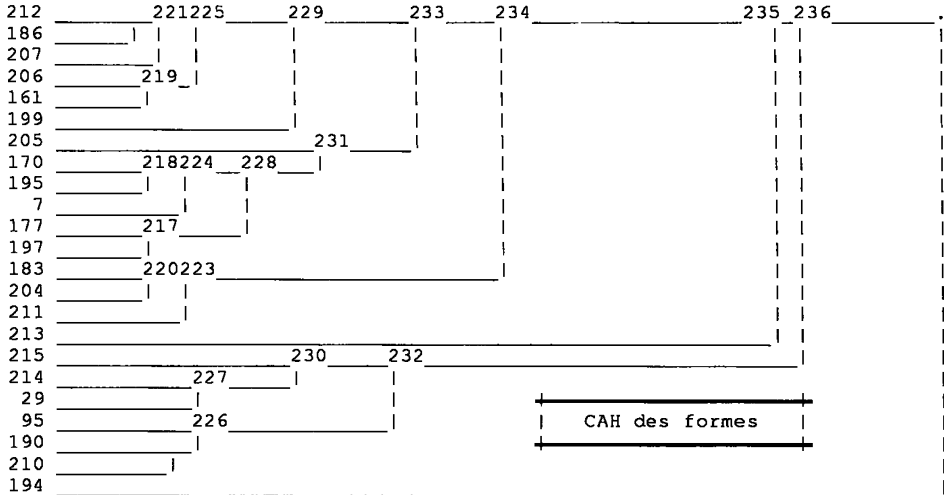
2.3 Classification des formes d'après les œuvres

La CAH des œuvres est étiquetée d'après les classes de formes; et réciproquement pour la CAH des formes. Les deux étiquetages concordent entre eux, sans toutefois qu'ils 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.

Mais tandis que la CAH des œuvres s'interprète facilement en termes de style et de genres littéraires, la CAH des mots outil ne répond exactement à aucune division des mots en parties du discours; et sa structure globale n'est elle-même pas simple.

Se détache d'abord la classe j194, {sub quoque qve}, associée aux poètes: on retrouvera au §3.3, une classe 180, {at nunc quoque sub qve}, étiquetée de même; et, présentement {at nunc} sont dans j204, liée à Virgile.

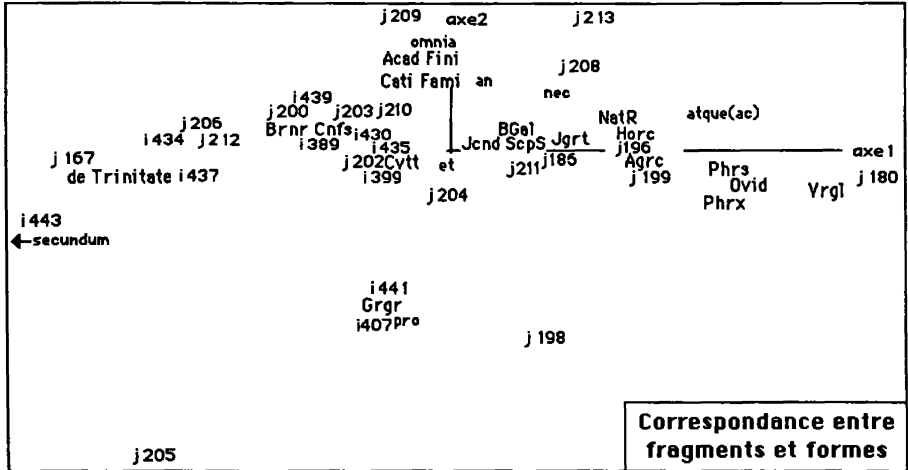
Le reste constitue la classe 236; qui comprend, outre deux branches principales j232 et j234, la classe j213 associée à Grgr, *Moralia in Job*.



Hormis une référence à Horace donnée pour j214, les subdivisions de la branche j232 n'ont aucune étiquette qui renvoie à un poète ou à un historien. Les prépositions comprises dans j232 sont *secundum* et *propter*, 'à cause de'. Ce ne sont pas des prépositions usuelles, comme *sub in ab ex*, qui expriment des rapports de position ou de mouvement. Les formes de j232 nous paraissent plutôt propres à conduire un raisonnement qu'un récit.

La branche restante, j234, n'offre prise à l'interprétation que si on la subdivise: j223 est plutôt propre aux poètes ainsi que l'atteste, outre l'étiquetage, la place des centres de ses subdivisions {j183, j204, j211} du côté ($F1 > 0$) sur le plan (1,2) qui illustre le §2.1. Le cas de j211 doit être noté: cette classe a, pour son poids, une inertie très faible; car son centre est proche de l'origine, ou Centre de Gravité (CdG). Mais j211 s'écarte de l'origine presque exactement dans la direction de l'axe 1: $COR1(j211) = 776$; et j211 contient { π mot π ltr} dont le dénombrement a été introduit pour prendre en compte la rareté des mots outil dans certains textes, cf. §1.3.3, *in fine*.

Reste enfin j233; seule l'étiquette Ovid+ pour j170 signale un lien positif avec les poètes; tandis qu'on lit plusieurs fois Vrgl-. On trouve dans j233 des classes associées aux historiens; particulièrement j199, déjà notée avec BGal, *de Bello Gallico* de César, vers ($F3 < 0$). La classe j205 {ea quidem nihil esset id esse}, étrangère à Virgile, caractérise plusieurs œuvres de Cicéron, d'où les mentions: {Acad++ Fami+ Fini++}. On soulignera ici un fait déjà apparu à la CAH des œuvres: quant à l'usage des mots outil, les *Lettres familières* sont proches du *de Finibus*, ainsi que d'autres dialogues philosophiques; bien que le thème des lettres, nullement philosophique, soit confiné dans l'actualité politique.



3 Correspondance entre fragments et formes et classification des fragments et des formes

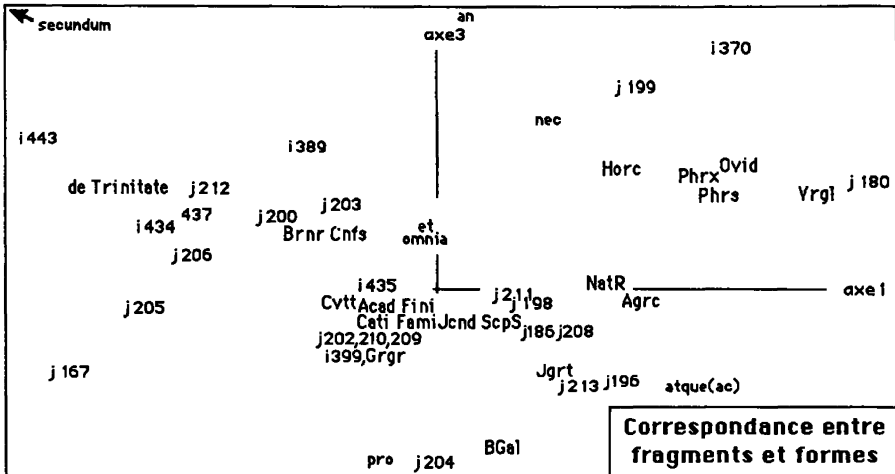
Fragments x formes de Δ ; Tr12 et μI1+ en supplémentaire

trace :	8.645e-1																		
rang :	1	2	3	4	5	6	7	8	...	10	...	20							
lambda :	776	496	449	322	285	219	201	185	173	114	e-4								
taux :	898	574	520	373	330	253	233	215	200	131	e-4								
cumul :	898	1472	1992	2365	2695	2948	3181	3395	3797	5302	e-4								

3.1 Correspondance entre fragments et formes

SIGJ	QLT	PDS	INR	F 1 CO2 CTR			F 2 CO2 CTR			F 3 CO2 CTR			F 4 CO2 CTR		
dum	550	3	12	37	0	0	-1284	407	85	-245	15	3	-718	127	41
esse	502	9	13	-99	8	1	680	387	85	-294	72	18	-203	35	12
quia	697	8	20	-723	234	52	-971	422	146	42	1	0	-297	40	21
scnd	653	2	28	-1760	206	64	-273	5	2	1156	89	48	2307	353	264
πltr	731	112	3	121	613	21	-50	108	6	-9	4	0	12	6	1
πmot	752	196	3	95	635	23	-33	79	4	17	21	1	15	16	1
qve	838	44	35	699	712	275	-76	8	5	281	115	77	46	3	3

Dans l'analyse présentée ici, deux des 234 fragments sont mis en supplément: Tr12 et μI1+. Le chapitre 12 du Livre I du *de Trinitate*, Tr12, argumente sur le verset du prologue de l'Évangile selon Saint Jean: *Omnia per ipsum facta sunt*, 'tout par lui a été fait': Tr12 s'en trouve lié aux trois formes {per, ipsum, omnia} assez fortement pour déterminer l'axe 5. Au début du Livre I des *Moralia in Job*, après l'épître d'envoi et la préface, Saint Grégoire le Grand entreprend de louer Job, d'avoir été bon parmi ceux qui étaient mauvais: grand mérite, d'où procèdent des louanges que l'Écriture décerne en maints passages: et c'est pourquoi 'd'où', *hinc*, se trouve répété 8 fois dans ce fragment noté μI1+ (l'initiale μ étant mise pour *Moralia*),.



Quoiqu'il en soit de ces détails, la présente analyse concorde exactement avec celle présentée au §2.1. Afin de le vérifier, on a calculé les corrélations, sur l'ensemble des 232 fragments (Tr12 et μ 11+ non compris), entre facteurs de même rang issus des deux analyses: la corrélation n'est véritablement détruite

corr (axe1, axe1) = 9.962 e-1	corr (axe7, axe7) = 7.737 e-1
corr (axe2, axe2) = 9.848 e-1	corr (axe8, axe8) = -6.838 e-1
corr (axe3, axe3) = 9.808 e-1	corr (axe9, axe9) = -2.387 e-1
corr (axe4, axe4) = 8.171 e-1	corr (axe9, axe10) = -2.427 e-1
corr (axe5, axe5) = 8.173 e-1	corr (axe10, axe10) = -6.592 e-1
corr (axe6, axe6) = -9.212 e-1	

qu'au rang 9. Et de façon précise, on voit, dans le plan (9,10) le fragment X du

```

121| Phrx      : 180++  199+++++
|CLAS|QLT PDS INR| F 1 CO2 CTR| F 3 CO2 CTR| F 9 CO2 CTR| F10 CO2 CTR|
|i121| 678  5 13| 418 132 10| 302 69  9| -455 157 54| -492 183 63|

```

```

199| hic quis  : Phrx+++++
|CLAS|QLT PDS INR| F 3 CO2 CTR| F 8 CO2 CTR| F 9 CO2 CTR| F10 CO2 CTR|
|j199| 632  4 21| 541 120 26| 422 73 38| -626 161 89| -633 164 92|

```

Livre VII de la Pharsale, Phrx, associé, par une répétition rhétorique, à la classe j199 des formes {hic, quis}: *hic patriæ perit omne decus...* 'ici tombent ceux qui faisaient l'honneur de la Patrie'; je ne vous dirai point *quis corruat ictu, quis steterit...* 'qui tombe en recevant le coup, qui reste debout...'

Pour aider à comparer les plans (1,2) et (1,3) du §3.1 à ceux du §2.1, on a, dans ceux-là, remplacé les numéros de plusieurs classes de fragments par les sigles des œuvres qui en constituent le fond. Ainsi, dans le plan (1,3), la classe i442, formée des 15 fragments du *de Bello Gallico*, et donc désignée par le sigle BGal, est proche de j204 {sine suo sibi ne eius eo suis eos se}; comme, au §2.1, BGal est proche de {his sibi se eum suis eo eos eorum}.

c	Partition en 22 classes : Sigles des fragments de la classe c																						
434	Tr01	Br10	Tr15	Tr29	Tr26															secundum+++	212++		
437	Tr28	Tr17	Tr19	Tr10	Tr07	Tr11	Tr13	Tr18	Tr02	Tr16											206+	212++	200+
443	Tr20	Tr21	Tr25	Tr24																secundum+++++	212++		
439	Cvtx	Tr04	Br12	Br09	Fam1	Br03	Tr09	Br08	Cv9a	Cv9b											212+	200++	
430	Br06	Br05	Cf13	Cfx8	Br01	Br02	Br04	Cnf7	Tr06	Tr05	Cfxx	Cfx6	Cnf6	Cfxv					et+	203+	200++		
399	Cv17	CV18	Cv18																		pro+++++	200+	
435	Cvt1	Cvt3	Br07	Cvt6	Cvt5	Tr14	Cv16	Cv12	CV12	Cv19	Cv20	Juc6	Cv24	Br11					200++				
429	Scp2	Scp1	Scp3	Juc4	Juc2	Juc5	Juc3	Juc7	Juc1	Jg8+	Jgxx											=CdG	
444	Cat2	Cat3	Cat6	Cat1	Cat4	Cat5	Fam2	Fam6	Fam3	Fam5	Fam4	Fin2	Acd2	Fin4					210+	209++			
440	Acd4	Acd5	Acd7	Acd3	Acd6	Fnx1	Acd8	Fin8	Finx	Fndz	Fin9	Scp4	Fin7	Fin3					206+	209+++++			
407	μI6a	μP7+	μI8+	μPr5	μI18	μI4+	μP9+											206+	205++++	pro++	198++		
441	μPr3	μI6d	μI14	μI31	μPr1	μI20	μI15	μEp5	μI13	μPr4	μP6-	μI5@	μI11	μI6c					205++++				
442	BG13	BGqz	BG16	BGqZ	BG11	BG17	BG12	BG18	BG15	BG14	BGxi	BG1x	BG19	BGdz					210+	204++++			
436	¶Cvt8	Nat2	Ntdz	Nat9	Ntxi	Natx	Nat4	Nat6	Nat7	Nat3											180+	208++++	
438	Jg1+	Jg14	Jg24	Jg31	Jg18	Jg35	Jgx+	Jg28	Jg4+	Jgxx											213++	196+	208+
401	Jg6+	Jg32	Jg25	Jg12	Agr7	Agr8	Agr6	Agr3	Agr2	Agrx	Agr1	Agr9	Agr4	Agr5					atque(ac)++++				
389	¶Cf1+	¶Cf3+																		omnia+	et++	an++++	199+
412	Hr07	Hr02	Hr03	Hr05	Hr04	Phxv	Hr06	Hr01											et+	an++	180+		
121	Phrx																				180++	199+++++	
384	Phr2	Phr5	Ph12	Ph14	Phr6	¶Vr07	Phr7	Phr8											180++++				
370	Mut1	¶Phr1	Mut7	Mut5	Mut3	Mut6	Mut4	Mut2											180+++++				
385	¶Vr05	¶Vr03	¶Vr02	¶Vr08	¶Vr01	¶Phr3	¶Vr04	¶Vr06											180++++				

3.2 Classification des fragments d'après leur correspondance avec les formes

La classification présentée ici a été effectuée dans l'espace engendré par les 20 premiers axes factoriels issus de l'analyse du §3.1. Des résultats semblables ont été obtenus avec 30 axes, 93 axes, ... ainsi qu'avec d'autres lexiques; avec 60 axes, les œuvres de Cicéron, entraînant la majorité des chapitres de l'encyclique *jucunda sane*, se sont agrégées au bloc des historiens, laissant toutes les œuvres de Saint Augustin constituer une seule classe, où est compris le *de Consideratione* de Saint Bernard. Nous avons déjà commenté cette variante taxinomique, au §2.1, d'après le plan (1,3) issu de la correspondance entre œuvres et formes de mots outil.

434	_____	449	_____	456	_____	Trnt	_____	460	_____	462	_____
437	_____		_____		456:scnd+++		212+++		_____		_____
443	_____		_____		_____		_____		_____		_____
439	_____	448	_____	452	_____	200+++	_____	458	_____		_____
430	_____		_____		_____		448	Brnr	Cnfs		_____
399	_____	445	_____		_____		445	Cvtt			_____
435	_____		_____		_____		_____		_____		_____
429	_____	451	_____	453	_____	Cicéron	_____	429	Scps	Jcnd	
444	_____		_____		209+++		444	Cati	Fami	Fini	
440	_____		_____		_____		440	Acad	Fini		_____
407	_____	447	_____	_____	_____	Grgr	_____	198+	205+++++	_____	
441	_____		_____		_____		_____		_____		_____
442	BGal	_____	_____	_____	_____	459	_____	Atque+	_____	461	_____
436	NatR	_____	_____	457	_____		_____	208+			_____
438	Jgrt	_____	454	_____		(200	203	205)-			_____
401	Agrc	_____		_____		_____	_____	_____			_____
389	_____	450	_____	455	_____	_____	_____	180+++++	_____		_____
412	_____		_____		412	Horc	_____	_____			_____
121	_____	446	_____		_____		_____	_____			_____
384	_____	431		_____	384	Phrs	_____	_____		CAH	de 232 fragments
370	_____		_____		370	Ovid	_____	_____		d'après	20 facteurs
385	_____		_____		385	Vrgl	_____	_____			_____

La classification du présent § concorde exactement avec celle du §2.2. Certes, les 22 classes retenues ici ne coïncident généralement pas exactement avec une œuvre mais les indications de contenu concises, en terme d'œuvres, qui accompagnent l'arbre, suffisent à résumer le tableau détaillé du contenu des classes: la CAH, agrégeant les fragments, retrouve les œuvres; ou, du moins, crée des classes, dans lesquelles, sauf exceptions (marquées, sur le tableau, du signe '¶'), ne voisinent que des fragments dont la parenté de style était attendue.

On trouvera une exception dans le début du Livre I des Confessions: Saint Augustin, cherchant où et comment invoquer Dieu multiplie les questions telles que celle-ci: *An ubique totus es et res nulla te totum capit*, 'Ou bien es-tu tout entier partout, et aucune chose ne te contient tout entier?' Ainsi se trouve créée la classe i389, {cf1+, cf3+}, associée à *an*, 'ou bien' et *et*.

Quant à l'ensemble, comme au §2.1, une première dichotomie sépare i461, historiens et poètes. Du reste, i462, se détachent les *Moralia*, i447, laissant i460 à Cicéron et Saint Augustin. La dialectique sublime du *de Trinitate* occupe i456. Dans i458, enfin, la rhétorique de Cicéron, i453, fait pendant à la rhétorique chrétienne, i452.

À la base de la hiérarchie des classes, les détails eux-mêmes méritent l'attention. La classe i438, *Jugurtha*, se subdivise en deux parties dont l'une comprend, outre l'introduction de Salluste en forme de discours sur l'histoire, trois fragments où sont respectivement rapportés un discours d'Adherbal, une lettre de celui-ci, et un discours de Memmius au peuple romain. Cette partie rhétorique est reconnue par la plupart des CAH qu'on a effectuées; s'y adjoignent parfois les dernières paroles de Micipsa exhortant Jugurtha à rester uni à Adherbal et Hiempsal.

c	Partition en 25 classes : Sigles des formes de la classe c
210	a ut quam quibus ita omnium hac iam esset cum quod
202	etiam vero eorum ad de quo tamen qui cuius quem ille
74	pro
204	sine suo sibi ne eius eo suis eos se

213	aut neque
7	atque(ac)
29	et
186	quos multa ante omnes inter
196	erat omnibus his ab ubi sua fuit
211	ex in π ltr π mot ibu
=====	
198	hinc post unde per
205	quasi dum ergo quippe quia
=====	
4	an
199	hic quis
180	at nunc quoque sub qve

95	secundum

167	propter eum
212	sicut illo ipse erit illi qua ipsum est
203	cui sed sit illa haec sic ista tam quid
200	vel si illud hoc non

67	omnia
208	quorum contra posse nulla tum e quoniam nam magis igitur
59	nec
206	autem enim potest nisi sunt
209	id ea quidem ipsa quae nihil esse

3.3 Classification des formes d'après leur correspondance avec les fragments

Comme au §2, la CAH des œuvres est étiquetée d'après les classes de formes; et réciproquement pour la CAH des formes. Au §3.2, on a porté sur chaque ligne du tableau donnant le contenu d'une classe de fragments, des numéros de classes de formes (ou parfois des formes isolées) rendant compte de l'agrégation de cette classe de fragments. Cet étiquetage local ne nous paraît pas rendre directement raison des grandes branches – poésie, histoire, rhétorique – que la CAH a reconnues; et c'est pourquoi l'arbre des fragments (au §3.2) porte un étiquetage sommaire où seules figurent des mentions générales de formes.

Au §3.3, l'arbre de la CAH des formes est principalement étiqueté par des numéros de classes de fragments (accompagnés, selon l'usage, de signes exprimant l'intensité de l'association). Mais le contenu de ces classes de fragments est aussi indiqué, en prenant le signe '<' pour symbole de l'inclusion: ainsi, la mention: {399 435} < Cvtt, rappelle que les classes i399 et i435 sont constituées, principalement des chapitres de la Cité de Dieu.

210	441+	216	444+	228	232	235	236
202	435+					202 ≈ CdG	
pro	399+*****	225				pro:407++, 440-	
204	441+	Gal+****+				{399 435} < Cvtt	
213	444+	Jgrt++	222	230		{441} < Ggr	
atqc	Agrc+*****						
et	430+		226				
186	443++	Nat+	223				
196	Gal+	Jgr++					
211	Agrc+			211 ≈ CdG			
198	440-		224	Ggr+++			
205	437+	441+****					
an	430+	Nat+	219	229			
199	Phrx+*****				an:389+****, Horc++	199:Vrgl++	
180	Phrs+	Ovid+++		Vrgl++			
scnd	Trnt+*****				234		
167	435+	214	Trnt++	231	233	167:BGal++	
212	Trnt++++						
203	444+	215	430+				
200	439+					200:435+	
omnia	434++	218	Nat+	227		omnia:439+	
208	NatR+*****						
nec	439+	440+			{434 439 437} < Augustin		
206	437+	440+	221		{440 444} < Cicéron		
209	444+	440++					

Même si les classes de formes trouvées ici (par une CAH dans l'espace des 20 premiers axes issus de l'analyse du §3.1) ne coïncident pas avec les classes du §2.2, nous avons déjà montré sur plusieurs exemples que des groupes analogues se trouvent de part et d'autre. Sans prétendre épuiser la comparaison, considérons encore la classe j208, {quorum contra posse nulla tum e quoniam nam magis igitur}, associée au §3 avec le *de Natura Rerum*, de Lucrèce: on trouve, au §2, associées à cette même œuvre de poésie didactique, deux classes, j177, {nulla posse e}, et j197, {multa contra igitur quoniam magis tamen nam}, dont le contenu reproduit celui de la j208 du §3.

4 Perspectives et conclusions

Au §2, prenant explicitement en compte l'appartenance de chaque fragment à une œuvre déterminée, on a montré une image cohérente de la correspondance entre fragments et formes de mots outil. Au §3, l'analyse directe de cette même correspondance a retrouvé la même image, sans prendre en compte la distribution des fragments en œuvres. Nous en concluons que le dénombrement des mots outil peut, en latin comme en grec, servir à la stylométrie.

Il est vrai que certains passages, où joue la répétition rhétorique, ont des profils singuliers qui perturbent un peu l'analyse: on s'affranchira de cette difficulté, en écartant ces passages, voire en écartant certaines formes. Nous avons, d'autre part, au §1.3, spéculé sur la polysémie, les équivalences entre formes différant quant à la phonétique... Mais de tels tâtonnements nous ont aussi permis d'apprécier la stabilité des conclusions générales; stabilité que devrait confirmer des analyses fondées sur un plus vaste corpus.

Références de statistique linguistique

[FRÉQ. CAT. LAT], par J. DENOZ, “Application des méthodes d'analyse factorielle à la fréquence des catégories grammaticales en latin”, in *CAD*, Vol XIII, n°1, pp. 19-40; (1988).

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

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

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

Sources des textes analysés (cf. §1.2)

N.B. La diversité des éditions sur lesquelles ont été saisis les textes nous a contraint à considérer attentivement les variantes de l'orthographe et des liaisons entre mots (cf. §1.3.3); en sorte que, même si notre transcription n'est ni définitive dans son principe, ni sans erreur dans sa réalisation, le succès des analyses nous assure du moins qu'il est possible d'embrasser dans leur ensemble les textes latins comme relevant d'une unique tradition littéraire.

a Latinitatis Monumenta

Publii Virgilii Maronis Carmina Omnia, perpetuo commentario ad modum Joannis Bond, explicuit Fr. Dubner, Parisiis, ex typographia Firminorum Didot; (1858).

Œuvres complètes d'Horace, de Juvénal, de Perse, de Sulpicia, de Turnus, de Catulle, de Propertius de Gallus et Maximien, de Tibulle, de Phèdre de Syrus, avec la traduction en français, publiées sous la direction de M. Nisard, J.-J. Dubochet et compagne Éditeurs, Paris; (1839).

P. Ovidii Nasonis Opera quæ supersunt, Tomus secundus, sive “Metamorphoseon Libri”, Typis fratrum Barbou, Parisiis; (1793).

Lucrèce: *De la Nature*, Tome Premier, texte établi et traduit par Alfred Ernout, quatrième édition revue et corrigée, Les Belles Lettres, Paris; (1942).

M. Annæi Lucani Pharsalia, cum indice rerum, ..., Editio C. Tauchnitii Stereotypa..., sumptibus succ. Ottonis Holtze, Lipsiæ; (1904).

De Bello Gallico Commentarii, édition à l'usage des classes par M. A. Boué, quinzième édition, J. de Gigord, Paris; (1926).

Traduction de Salluste, avec le texte latin, la vie de cet historien, des notes critiques et des variantes, par J. H. Dotteville, de l'Oratoire, À Paris, chez Lottin l'aîné; (1769).

C. Cornelii Taciti Opera quoad extant, ex fide optimorum codicum accurate recensuit C. H. Weise, Tomus secundus, "Historiarum Libri V. Germania. Agricola. Dial. de oratoribus", editionis stereotypæ C. Tauchnitianæ nova impressio, sumptibus Ottonis Holtze, Lipsiæ; (1874).

b O Roma felix

Saint Augustin: *Confessions*, texte établi et traduit par Pierre de Labriolle, Tome I, Les Belles Lettres, Paris; (1956).

Saint Augustin: *La Cité de Dieu*, texte latin et traduction française, avec une introduction et des notes, par Pierre de Labriolle, Tome I, Classiques Garnier, Paris; (1941).

"Sancti Aurelii Augustini, Hipponensis episcopi, De TRINITATE, libri quindecim", in *Sancti Augustini, Hipponensis episcopi, Opera omnia, post Lovaniensium theologorum recensionem, ...*, accurante J.-P. Migne; sive *Patrologiæ Latinæ Tomus 42*; Paris, (1841).

"Sancti Gregorii Magni, romani pontificis, Moraliū Libri, sive expositio in Librum Beati Job", in *Sancti Gregorii Papæ I, cognomento Magni, Opera omnia*, accurante J.-P. Migne; sive *Patrologiæ Latinæ Tomus 75*; Paris, (1849).

"De Consideratione ad Eugenium Papam", in *Sancti Bernardi Opera, Vol III, Tractatus et Opuscula*, recensuerunt J. Leclerc & H. M. Morchais, Romæ, Editiones Cistercienses; (1963).

"Jucunda sane", lettre encyclique sur la responsabilité de ceux qui gouvernent l'Église, in *Saint Pie X, écrits doctrinaux*, préface de Jean Daujat, Téqui, Paris; (1975).

c Eloquentiæ flumen

Les Catilinaires de Cicéron, présentées par Guy Michaud, Classiques Roma, Hachette, Paris; (1938).

Cicéron: *Lettres Familières*, I, (Livres I–VI), traduction nouvelle avec préface et notes par Édouard Bailly, Librairie Garnier Frères, Paris; (1933).

M. Tullii Ciceronis, De Officiis Libri tres, Cato Maior, Lælius, Paradoxa, Somnium Scipionis, ex recensione Georgii Grævii, cum ejusdem notis, ut & integris animadversionibus Dionysii Lambini, Fulvii Vrsini, Caroli Langii, Francisci Fabritii Marcodurani, Aldi Manutii, necnon selectis aliorum, ..., ex Typographia P. & I. Blaev, Amstelodami; (1688).

M. Tullii Ciceronis Opera Omnia, adiectis indicibus et adnotatione, iterum edidit, Car. Frid. Aug. Nobbe, nova editio Stereotypa C. Tauchnitiana, Tomus VIII, "Academica. De finibus bonorum et malorum. Tusculanæ disputationes", Sumptibus Ottonis Holtze, Lipsiæ; (1874).

