

## INTRODUCTION

En 1984, après la publication aux éditions G. Olms, de l'*Index verborum* et des listes de fréquence de la *Métaphysique* d'Aristote<sup>1</sup>, des philosophes et des philologues classiques de l'Université de Liège décidaient d'enregistrer en mémoire d'ordinateur la totalité des œuvres du Stagirite.

La transcription des textes anciens sur bandes ou sur disques magnétiques n'est pas une fin en soi. Elle ne trouve sa justification que si elle débouche sur des travaux qui ouvrent au chercheur des voies d'investigation qui lui étaient jusque-là peu accessibles : les procédures d'indexation des langues grecque et latine qui ont été développées au Laboratoire d'Analyse statistique des Langues anciennes (L.A.S.L.A.) ont notamment pour objectif de fournir des dénombrements exhaustifs et des relevés propres à résoudre les nombreux problèmes que pose la littérature de l'antiquité gréco-romaine.

Lorsque l'on applique à de tels relevés les méthodes de la linguistique quantitative, il est possible de comparer l'une à l'autre des œuvres ou des parties d'œuvres, en vue d'apporter des éléments de réponse notamment à des questions de chronologie relative, lesquelles sont capitales pour étudier l'évolution de la pensée d'un philosophe. C'est ce que montre, par ses travaux, un des auteurs de l'index de la *Métaphysique*<sup>2</sup>, qui, grâce à un usage prudent de la stylométrie, a

---

<sup>1</sup> L. DELATTE, Chr. RUTTEN, S. GOVAERTS, J. DENOZ, *Aristoteles, Metaphysica. Index verborum, Listes de fréquence*, Hildesheim, Olms-Weidmann, 1984.

<sup>2</sup> Chr. RUTTEN est l'auteur d'une dizaine d'études stylométriques sur le corpus aristotélicien. Le lecteur consultera notamment :

- *Aristote, Ethique à Nicomaque X et Métaphysique, AB et ZHΘ. Essai de chronologie relative*, dans *Actes du Congrès international "Informatique et Sciences humaines"* (L.A.S.L.A., nov. 1981), Liège, L.A.S.L.A., 1982, pp. 839-844.
- *Aristote, Métaphysique Z. Essai de stylométrie*, dans *Revue de l'Organisation internationale pour l'Etude des Langues anciennes par Ordinateur*, 1982, n° 1-4, pp. 163-192.
- *Note sur quelques conjoncteurs chez Aristote, Métaphysique, Γ, Z et H*, dans *Revue de Philosophie Ancienne*, 2 (1983), pp. 197-201.
- *Aristote, Métaphysique Z, 3. 7. 8. 9 et 17. Essai de chronologie relative*, dans *Revue, Informatique et statistique dans les sciences humaines*, 1983, n° 1-4, pp. 175-188.
- *Stylométrie des Catégories*, dans *Aristotelica, Mélanges offerts à Marcel De Corte*, Bruxelles-Liège, 1985, pp. 314-336.

publié plusieurs essais de chronologie relative de cette œuvre, en se fondant sur la répartition des parties du discours ou sur la distribution des conjoncteurs les plus fréquents.

Ces études statistiques se fondent sur des relevés établis avec la plus grande rigueur, car la finesse et l'exactitude des observations auxquelles on aboutit dépendent d'abord de la qualité des données sur lesquelles repose l'enquête.

## La préparation d'un index

### I.- Le contenu du fichier.

Pour nos recherches, nous avons choisi d'accompagner chaque mot qui figure dans les textes que nous analysons, d'informations normalisées et codifiées se rapportant au vocabulaire et à la grammaire. Dès lors, les enregistrements que nous constituons et que nous conservons sur bandes magnétiques portent en regard de chaque forme du texte six informations :

- 1.- le lemme, c'est-à-dire la forme canonique qui sert de rubrique dans le dictionnaire choisi comme ouvrage de référence. Etant donné que nous utilisons Liddell et Scott<sup>3</sup>, les nominaux seront généralement rangés sous la forme du nominatif singulier, sauf dans les cas où elle fait défaut; pour les verbes, c'est la première personne du singulier de l'indicatif présent actif ou moyen qui est utilisée pour le regroupement des différentes formes et pour le classement alphabétique.
- 2.- les indices de lemme sont des codes numériques qui permettent de distinguer les mots homographes. Ainsi, le lemme ἀπόφασις a deux significations

---

- *Métaphysique B et Γ : essai de chronologie relative dans Aristoteles Werk und Wirkung*, t. I (*Aristoteles und seine Schule*), éd. par J. Wiesner, Berlin, De Gruyter, 1985, pp. 273-287.

- *Aristote, Métaphysique, Lambda 8. Essai de stylométrie*, dans *Méthodes quantitatives et informatiques dans l'étude des textes* (Actes du Colloque de Nice, 5-8 juin 1985), Paris, Slatkine, 1986, pp. 747-758.

- *Métaphysique M et N. Essai de chronologie relative*, dans *Mathematik und Metaphysik bei Aristoteles (X. Symposium Aristotelicum)*, Berne, Haupt, 1987, pp. 33-44.

- *Analyse comparative des chapitres de la Métaphysique d'Aristote fondée sur les fréquences d'emploi des parties du discours; confrontation entre l'ordre du textus receptus, les références internes et l'ordre du premier facteur*, à paraître dans *Les Cahiers de l'Analyse des données*, (12), 1988 (en collaboration avec J.-P. Benzécri).

<sup>3</sup> H.G. LIDDELL, R. SCOTT, H.S. JONES, *A Greek-English Lexicon*, Oxford, Clarendon, 1940 et *A Greek-English Lexicon, A Supplement*, Oxford, 1968.

principales distinguées par les chiffres 1 et 2 qui sont équivalents aux rubriques A et B que propose le dictionnaire. En principe, pour l'attribution des indices numériques, nous suivons l'ordre de succession de Liddell et Scott. La seule exception concerne les composés de *-ειμι* pour lesquels nous attribuons systématiquement l'indice 1 aux verbes dérivés de *εἰμί* et l'indice 2, à ceux qui sont formés à partir de *εἰ̃μι*.

- 3.- la référence qui accompagne chaque mot est conforme à l'édition de Bekker; elle mentionne la page, la colonne et la ligne. A ces éléments traditionnels, nous en avons ajouté trois qui n'apparaissent pas dans l'index : le numéro du mot dans la ligne, le numéro du mot dans l'œuvre et enfin, le numéro du mot dans la phrase. Cette dernière information est plus spécialement destinée à d'éventuelles recherches sur la longueur des phrases.
- 4.- un code représentant la catégorie grammaticale du lemme. Ce code qui affecte chaque mot, est le plus souvent le même pour les diverses formes d'un lemme. Ainsi, toutes les occurrences de *ἀγαθός* sont classées dans la catégorie grammaticale *I*, qui est celle des adjectifs, même si certaines formes sont employées substantivement. Nous n'avons pas cru souhaitable, en effet, de faire apparaître dans nos *indices* qui ont d'abord une portée lexicographique, des indications stylistiques ou sémantiques. Par ailleurs, décider qu'un mot a valeur de substantif suppose un choix parfois subjectif de la part du philologue qui réalise l'indexation. Ce choix peut avoir pour conséquence de tromper l'utilisateur de l'ouvrage.  
Par contre, pour les mots grammaticaux, nous indiquons la catégorie de la forme et nous distinguons, par exemple, les emplois de *καί*, conjonction de coordination et adverbe. De même, nous codifions différemment *οὐδέ*, selon qu'il est adverbe ou conjonction. Les distinctions que nous opérons ainsi influent sur le classement des occurrences dans l'index. Nous reviendrons à cette question ultérieurement.
- 5.- des notes relatives à l'apparat critique. A cet égard nous nous bornons à l'essentiel, c'est pourquoi nous ne signalons que :
  - a) des conjectures de l'éditeur que nous avons choisi;
  - b) des *crucés*;
  - c) des mots qui, présents dans certains manuscrits, sont rejetés par l'éditeur.

En outre, nous affectons d'un code particulier les mots qui font partie d'une citation.

- 6.- des codes destinés au calcul automatique de la référence. Ils servent à marquer les mots qui se trouvent en fin de ligne, en fin de phrase, en fin de colonne ou en fin de page.

En ce qui concerne la référence, on retiendra que les mots coupés en raison de leur place en fin de ligne, sont rejetés entièrement sur la ligne suivante. Cette convention permet d'obtenir un nombre de lignes par page toujours conforme à celui de l'édition.

## II.- Les procédures de lemmatisation et d'analyse.

On se rend aisément compte de la somme de travail qui incomberait au philologue s'il devait porter en regard de chaque mot d'un texte l'ensemble des informations qui viennent d'être décrites.

Pour réaliser cette phase essentielle de la recherche, nous avons développé une méthode semi-automatique qui permet d'accélérer la préparation d'un *index* en la confiant à l'ordinateur. La machine dispose, pour effectuer cette tâche, d'un dictionnaire grec informatisé dans lequel chaque mot-forme est accompagné de son lemme et de sa catégorie grammaticale.

Au moment où une œuvre est enregistrée en ordinateur, celui-ci cherche dans son lexique un enregistrement qui contienne une forme identique à celle qu'il traite et, s'il en repère un, il reproduit en regard de la forme le lemme et le code morphologique qu'il trouve dans son dictionnaire.

Le lexique informatisé contient plus de soixante mille formes et il s'enrichit de nouvelles unités à chaque utilisation du système. On comprend dès lors que notre programme identifie correctement plus de 95 % des occurrences d'un texte d'Aristote. Pour d'autres auteurs, les résultats se situent aux environs de 90 %<sup>4</sup>.

Cependant, qu'on ne s'y trompe pas, la machine n'effectue qu'un travail préparatoire; en fait, elle se borne à suggérer des solutions au philologue. Ce dernier doit ensuite vérifier les propositions d'indexation, les corriger ou les compléter. Il sera particulièrement attentif aux mots pour lesquels l'ordinateur a produit plusieurs lemmes et plusieurs analyses.

A l'issue de la vérification faite par le philologue et après le report des corrections sur support informatique, le fichier est prêt pour les phases purement automatiques du traitement : la mise en ordre alphabétique, l'établissement des relevés fréquentiels lexicaux et morphologiques, la constitution de concordances complètes ou sélectives, l'impression de l'*index verborum* ou encore la production d'un lexique inverse des formes contenues dans une œuvre.

---

<sup>4</sup> Le lecteur trouvera une description plus complète des méthodes automatiques de lemmatisation et d'analyse des langues classiques dans J. DENOZ, *Le traitement des textes latins, grecs et français au Laboratoire d'Analyse statistique des Langues anciennes* dans *Revista de la Universidad complutense*, 25 (1976), n° 102, pp. 143-167.

Cette liste n'est pas limitative; en effet, les fichiers informatisés offrent de nombreuses possibilités d'exploitation. C'est ainsi que l'on peut facilement obtenir les occurrences de telle ou telle conjonction de subordination dont on veut examiner la syntaxe, ou encore faire apparaître les suites de mots introduites par une préposition, etc.

## L'index de la *Poétique*

### I.— Le choix de l'œuvre.

On pourrait s'étonner de la publication d'un nouvel index de la *Poétique* et se demander si cet ouvrage n'est pas inutile puisqu'existe déjà, en plus de l'*Index Aristotelicus* de Bonitz<sup>5</sup>, le remarquable *Lexique de la "Poétique" d'Aristote* publié récemment par A. Wartelle<sup>6</sup>.

Néanmoins, notre travail trouve sa raison d'être dans le fait qu'il s'intègre à l'entreprise plus large d'indexation et d'analyse du *Corpus Aristotelicum* dont nous avons parlé au début de la présente introduction. Par ailleurs, notre documentation est plus riche que celle de nos prédécesseurs puisqu'elle comporte des indications précises sur la fréquence des mots, sur la distribution du vocabulaire et sur la répartition des catégories grammaticales, indications qui sont indispensables lorsque l'on veut entreprendre des recherches stylométriques.

Dans cette perspective, le fait que le texte de la *Poétique* est traité selon les mêmes procédés que celui des autres œuvres d'Aristote<sup>7</sup> autorise des comparaisons fondées sur des données parfaitement homogènes, ce qui est rarement possible avec les lexiques réalisés sans l'aide des méthodes informatiques.

De telles comparaisons aideront peut-être à résoudre les problèmes que pose la *Poétique*. J'en mentionnerai deux à titre d'exemple, le premier concerne la composition de l'œuvre, le second sa chronologie relative dans le *corpus* aristotélicien.

Selon les critiques, l'œuvre répond, dans son ensemble, à un plan bien ordonné, même si sa composition présente quelques irrégularités. Ainsi, de

---

<sup>5</sup> H. BONITZ, *Index Aristotelicus*, dans *Aristotelis opera*, V, Berlin, Académie de Berlin, 1870.

<sup>6</sup> A. WARTELLE, *Lexique de la "Poétique" d'Aristote*, Paris, Les Belles Lettres, 1985.

<sup>7</sup> Indépendamment de l'*Index de la Métaphysique* déjà publié, un volume consacré au traité *De l'âme* sort de presses en même temps que le présent volume. Par ailleurs, des *indices* de la *Physique*, des *Parties des animaux* et du traité *De l'Interprétation* sont actuellement en préparation.

1452 a 14 à 1452 a 27, Aristote arrête son exposé sur la fable pour parler très brièvement des divisions de la tragédie, puis il reprend en 1452 a 28 le développement interrompu avant 1452 a 14. Les exégètes de la *Poétique* considèrent la digression sur les parties de la tragédie comme un passage interpolé, voire apocryphe.

De même, en 1454 b 19, Aristote propose une typologie de la reconnaissance en commençant par se référer à ce qu'il a dit antérieurement : ἀναγνώρισις δὲ τί μὲν ἔστιν, εἴρηται πρότερον. S'agit-il, ici encore, d'un passage interpolé?

Certains considèrent ces deux passages comme des développements nouveaux et plus nuancés sur des questions traitées trop rapidement. Ils expliquent leur hypothèse en disant :

La *Poétique* est le cours professé par Aristote sur la poésie, et l'auteur a pris soin de l'opposer lui-même aux ἐκδεδομένοι λόγοι, aux traités publiés, dans l'espèce le dialogue *Sur les Poètes*. Avec ses digressions, ses retours en arrière, ses longues parenthèses, ses hésitations parfois, la *Poétique* a l'allure et aussi tout l'intérêt d'un exposé verbal. Cet exposé, soyons assurés qu'Aristote l'a fait plus d'une fois, ajoutant au canevas primitif des parties nouvelles, sans prendre la peine, assez vaine pour un maître qui ne vise qu'à l'instruction des ses auditeurs, de remanier le tout en vue d'une composition irréprochable<sup>8</sup>.

Cette conception est séduisante dans la mesure où elle semble répondre à une certaine logique, mais la phrase εἴρηται δὲ περὶ αὐτῶν ἐν τοῖς ἐκδεδομένοις λόγοις ἱκανῶς<sup>9</sup>, qui lui sert d'argument, permet-elle d'affirmer que la *Poétique* est un traité qu'Aristote n'a pas publié; ἐκδεδομένοις ne sert-il pas plutôt à rappeler ce qui a été dit dans les traités publiés antérieurement? Dans cette hypothèse, 1454 b 17-18 peut se trouver aussi bien dans une œuvre écrite que dans un exposé oral.

En outre, il me semble difficile d'admettre que les irrégularités observées dans le plan de l'œuvre s'expliquent par des ajouts à un exposé primitif. Admettons qu'Aristote ait développé progressivement et recomposé certaines parties de la *Poétique*; cela n'explique pas pourquoi il ne les a pas présentées à la place qu'elles auraient dû occuper logiquement même dans un exposé oral.

Au reste, est-il essentiel de savoir si le texte de la *Poétique* tel qu'il nous est parvenu, est issu d'un exposé oral ou s'il est d'abord une œuvre publiée? Il serait beaucoup plus important, pour étudier l'évolution de la

---

<sup>8</sup> J. HARDY, *Aristote, Poétique*, Coll. des Universités de France, 3<sup>e</sup> éd., Paris, Les Belles Lettres, 1961, p. 8.

<sup>9</sup> *Poétique*, 1454 b 17-18.

pensée du Stagirite, de pouvoir dire à quels traités celui-ci fait allusion lorsqu'il écrit *ἐν τοῖς ἐκδεδομένοις λόγοις*. Ceci pose le problème de la chronologie relative de la *Poétique* et de la place qu'elle occupe dans l'ensemble du *Corpus Aristotelicum*. Cette question, on le sait, n'a reçu à ce jour aucune réponse satisfaisante.

Les documents et les relevés que nous publions dans ce volume devraient permettre d'aborder sous un éclairage nouveau les problèmes de composition et de chronologie que nous venons d'évoquer. En effet, en partant d'informations objectives et exhaustives, on peut comparer les unes aux autres les différentes parties de la *Poétique* et on peut aussi étudier les rapports qui existent entre cette œuvre et les autres textes d'Aristote.

## II.— L'édition.

Pour la préparation de notre volume, le texte de base que nous avons choisi est celui que J. Hardy a publié dans la Collection des Universités de France<sup>10</sup>.

La référencement suit de très près celle de l'édition de Bekker, dont elle diffère dans quelques cas qui sont sans incidence sur le repérage des mots. En fait, on constate pour trois ou quatre pages, un simple décalage d'une ligne résultant de contraintes typographiques.

## III.— L'ordonnance de l'ouvrage.

Ce volume contient des relevés lexicaux et grammaticaux un *index verborum* et une liste des mots employés dans la *Poétique*.

### LES RELEVÉS.

A la suite de l'introduction, figurent deux tableaux. Le premier donne la distribution des occurrences par catégories grammaticales et le second la distribution du vocabulaire.

---

<sup>10</sup> J. HARDY, *Aristote, Poétique*, Coll. des Universités de France, 3e éd., Paris, Les Belles Lettres, 1961.

## L'INDEX VERBORUM.

Cette première partie de l'ouvrage fournit les renseignements suivants : la fréquence du mot, le lemme et éventuellement son indice, la forme, qui peut être accompagnée d'indications relatives à la tradition manuscrite, et enfin la référence.

La fréquence apparaît devant le lemme dont elle indique le nombre d'emplois dans la *Poétique*. Il faut noter que les leçons rejetées par l'éditeur, ainsi que les mots appartenant à une citation, ne sont pas pris en compte dans le calcul du nombre d'occurrences.

A la suite de chaque lemme figurent ses différentes formes, rangées en ordre alphabétique, telles qu'elles apparaissent dans l'édition à laquelle nous nous référons.

On remarquera cependant que pour des raisons de classement et de présentation, nous supprimons les accents graves et les remplaçons par des accents aigus. En outre, nous avons dû, pour des raisons typographiques, éliminer le tréma lorsqu'il apparaît sur une lettre accentuée.

Pour consulter avec profit l'*index verborum*, le chercheur sera attentif au fait que les formes d'un même lemme sont rangées selon trois critères cumulatifs.

- 1.- L'ordre alphabétique; ainsi, sous le lemme  $\tau\varepsilon$ , on trouvera successivement les formes  $\vartheta'$ ,  $\tau'$   $\tau\varepsilon'$  et  $\tau\varepsilon$ .
- 2.- La référence; les formes identiques sont rangées selon l'ordre d'apparition dans le texte.
- 3.- La catégorie grammaticale de la forme; ce critère ne s'applique en principe qu'aux mots grammaticaux dont la nature varie en fonction du contexte. C'est le cas, par exemple de  $\kappa\alpha\acute{\iota}$  qui est tantôt adverbe, tantôt conjonction de coordination.

Les formes qui portent des indications relatives à la tradition manuscrite et celles qui appartiennent à des citations sont rangées en fin d'article; elles sont signalées de la façon suivante :

[ ]	mot rejeté par l'éditeur;
< >	conjecture;
†	<i>cruz</i> ;
“ ”	citation.

L'ordre que nous avons adopté pour le classement des catégories est le suivant : article, substantif, adjectif, numéral, adjectif-pronom, verbe, adverbe, préposition, conjonction de coordination, conjonction de subordination, interjection et particule.

On notera que pour alléger quelque peu la présentation du volume, nous n'avons imprimé ni les références de  $\kappa\alpha\acute{\iota}$ , adverbe et conjonction, ni celles de  $\acute{\omicron}$ .

Ces deux lemmes figurent simplement dans l'*index verborum* avec leur fréquence globale<sup>11</sup>.

#### LA LISTE DES MOTS.

Il s'agit d'une liste dans laquelle les mots sont ordonnés selon leur fréquence d'emploi décroissante dans la *Poétique*. Les lemmes les plus utilisés dans l'œuvre sont imprimés en premier lieu et les plus rares apparaissent à la fin. Lorsque plusieurs vocables ont la même fréquence, ils sont rangés en ordre alphabétique. Ce relevé porte parfois le nom d'**index hiérarchisé**.

On trouve aux deux dernières pages de cette liste soixante-quinze lemmes de fréquence nulle (zéro); ce sont des mots qui n'apparaissent que dans des leçons rejetées par l'éditeur ou dans des citations.

---

<sup>11</sup> Parmi les 588 emplois de *καί*, on dénombre 493 conjonctions de coordination et 95 adverbes. En ce qui concerne *ὁ*, sur ses 1453 emplois, 1411 sont des articles et 42 des pronoms.