

LA RICHESSE DU VOCABULAIRE PATRISTIQUE COMME VALEUR MESURABLE

Lorsque nous avons établi, grâce au concours du Laboratoire d'Analyse statistique des Langues anciennes, à Liège, le premier volume des *Indices Chrysostomici*, paru en Allemagne chez Olms, nous savions que le bénéfice d'un travail de ce genre doit aller d'abord à l'éditeur de l'oeuvre.

Le chercheur, lui, découvre bien d'autres emplois. Tout d'abord, notre index facilite une enquête sur les sources du vocabulaire de Chrysostome. Sur cent des mots signifiants commençant par *α*, on n'en trouve pas moins de trente-huit appartenant au vocabulaire d'Homère, de Sophocle, d'Euripide, contre quinze venus de la Septante. Ces *Indices* fournissent donc un point de départ commode et solide pour des recherches de lexicographie.

En outre, la liste de fréquence des lemmes permet de constater combien le vocabulaire est étroitement lié à la psychologie des personnages, à la situation dans laquelle ils se trouvent. Θεός est le mot le plus fréquemment employé dans le premier volume des *Indices Chrysostomici* : 190 fois sur 58 152 occurrences; il n'est dépassé que par des mots-outils ou par des mots-thèmes aussi courants que λέγω, ὀράω, γίγνομαι, ποιέω. Aussitôt après viennent χαλεπός, 122 fois, et πάσχω, 116 fois. Ce sont vraiment les mots qui conviennent à la situation de Jean et de ses amis. Quant au mot ἀρετή, 83 fois, il représente le but auquel ils tendent de toutes leurs forces, tandis que le mot ἀθυμία, 76 fois, traduit le mal contre lequel ils doivent se défendre.

Avant de pousser plus loin l'analyse, il faut tenir compte du fait que notre premier volume regroupe trois oeuvres contemporaines, mais de longueur différente : les *Lettres à Olympias* (29 333 mots), la *Lettre d'exil* (8 460 mots) et *Sur la Providence de Dieu* (20 359 mots).

Si cet Index invite déjà à comparer entre elles les trois dernières oeuvres de Jean, sa valeur augmentera à mesure que paraîtront les volumes suivants qui fourniront des points de comparaison avec d'autres oeuvres du même auteur. Par exemple, le volume qui donnera l'index des homélies *De incomprehensibili* permettra de savoir si le vocabulaire de Jean a varié, sur ce thème, entre sa prédication à Antioche et à Constantinople, puisque ces homélies s'étalent, dans le temps, sur les deux périodes de son apostolat.

Comparaison de Jean avec lui-même, mais aussi avec d'autres auteurs chrétiens. Quand nous aurons l'index des homélies de Jean *Sur la Genèse*, nous pourrons comparer son vocabulaire avec celui de Basile *Sur l'origine de l'homme*, établi par les mêmes procédés et figurant dans le volume 160 de la collection "Sources Chrétiennes". Un index des homélies d'Hésychius réalisé au L.A.S.L.A. et imprimé chez Olms va nous donner la possibilité de comparer le prédicateur d'Antioche et celui de Jérusalem, en particulier dans leurs homélies respectives sur le personnage de Lazare.

Les Pères de l'Eglise sont, en ce domaine, un sujet d'étude privilégié puisque les uns et les autres reprennent, dans leur enseignement, le commentaire des mêmes passages de la Bible. La continuité et tout à la fois la distance que chaque théologien établit avec ses sources scripturaires et avec les Pères apostoliques apparaissent bien dans un index complet de son vocabulaire, que l'on peut comparer à des instruments de travail classiques (Concordance du Nouveau Testament et de la Septante, *Index apostolicus* de Goodspeed). Le volume de nos *Indices* consacré au dialogue *Sur le sacerdoce* de Jean Chrysostome et qui est en cours d'élaboration devrait en apporter une nouvelle preuve.

Comparaison d'auteurs chrétiens entre eux, mais aussi comparaison avec des auteurs païens contemporains, pour Jean avec Thémistios et Libanios; ce dernier, par exemple, dans la monodie sur la sédition d'Antioche, peut être mis en parallèle avec Jean dans les homélies *Sur les statues*. La confrontation de leur vocabulaire sur des points voisins serait du plus haut intérêt pour des travaux de sémantique.

Il faut donc multiplier ces *Indices* composés grâce au traitement automatique des textes. Seuls ils sont capables de fournir, sur

un grand nombre de mots, les données précises et sûres qui permettent au chercheur d'entreprendre de nouvelles enquêtes.

Ainsi, la liste de fréquence du premier volume des *Indices Chrysostomici* met en évidence la diversité du vocabulaire de Chrysostome puisque, sur un total (après élimination des noms propres) de 3990 lemmes, 1423 ne sont employés qu'une fois et 635 ne sont employés que deux fois.

Partant de cette observation, Marie-Louise Guillaumin a tenté de cerner la notion de richesse du vocabulaire patristique et de lui donner un contenu objectif. Elle expose ici la méthode qu'elle a suivie et les premières conclusions qu'elle a tirées de cet exercice de "statistique linguistique".

De toute évidence, il fallait procéder par comparaison, l'analyse d'autres textes pouvant conforter ou controuver une impression initiale. D'où l'idée de mettre à profit les études de vocabulaire latin classique menées au L.A.S.L.A. et publiées dans sa *Revue*, en 1974 et 1978, par J. Denooz et S. Govaerts. Pour pallier ce que la mise en parallèle de ces ouvrages d'époques différentes, les uns latins et les autres grecs, avait d'arbitraire, j'ai choisi de compléter le dossier par un texte *latin* contemporain des écrits de Chrysostome. Grâce au *Catalogus verborum* des *Tractatus in Evangelium Iohannis* de saint Augustin, édité en 1976 par le "Thesaurus Linguae Augustinianae" d'Eindhoven (Pays-Bas), j'ai pu rassembler des indications chiffrées sur le groupe des sermons I-XIV; ces prédications d'un autre évêque de l'âge d'or patristique, par leur volume, par leur contenu, et par la date de 406-407 que A.-M. La Bonnardière leur attribue, peuvent être utilement introduites dans nos évaluations.

En un premier temps, j'ai voulu voir si la part des mots de fréquence 1 variait beaucoup d'une oeuvre à l'autre, et s'il y avait là un critère de "richesse du vocabulaire". Parmi les ouvrages traités au L.A.S.L.A., j'ai retenu ceux dont la longueur était voisine de nos textes patristiques, puis j'ai combiné en un seul tableau des données puisées dans les précieux articles de la *Revue* et des nombres déterminés par moi :

OEUVRES	formes	lemmes	lemmes fréquence 1	en % de l'oeuvre
AUGUSTIN, Tract. Ev. lo. I-III	± 8400	1011	444	5.27
CHRYSOSTOME, Lettre d'exil	8417	1578	807	9.58
OVIDE	9852	2327	1051	10.66
SALLUSTE, Catilina	10094	1773	751	7.44
TITE LIVE	16665	2581	1047	6.28
AUGUSTIN, Tract. Ev. lo. I-V	±20100	1600	642	3.18
CHRYSOSTOME, Sur la Providence	20138	2441	1093	5.42
SALLUSTE, Jugurtha	20282	2260	773	3.81
VIRGILE, Enéide	30059	3494	1181	3.92
CESAR, De Bello Gallico	42826	2798	868	2.02
AUGUSTIN, Tract. Ev. lo. I-XIV	±57600	2853	957	1.66
CHRYSOSTOME, Olympias. Ex. Prov.	57646	3990	1423	2.46

Au regard de ces chiffres, on peut faire trois remarques : les mots de fréquence 1 représentent des fractions très comparables si l'on considère des écrits de longueur voisine; ces mots jouent un rôle décroissant à mesure que les oeuvres s'étendent, y compris dans les cas où elles sont du même auteur; enfin, Chrysostome occupe régulièrement une place intermédiaire entre poètes et prosateurs, alors qu'Augustin n'atteint pas les chiffres de ces derniers. Ainsi serait confirmée l'observation que la richesse du vocabulaire chrysostomien est lié à son caractère poétique.

Poursuivant dans cette voie, j'ai, en un second temps, porté sur un même graphique, pour chaque oeuvre (ou groupe d'oeuvres), en abscisse, le nombre total d'occurrences, en ordonnée, celui des lemmes. A l'intersection des deux coordonnées, un point caractérise cette oeuvre. Si l'on rejoint par une courbe les points relatifs à Chrysostome, on constate, d'une part, que le vocabulaire d'un auteur s'accroît, à mesure que s'amplifie son ouvrage, selon une progression logarithmique. On constate, d'autre part, que les poètes se situent tous au-dessus de la courbe, que les prosateurs s'en rapprochent, à l'exception de César (sans doute, le récit de la Guerre des Gaules est-il rédigé avec une économie voulue de moyens) et d'Augustin, ce qui s'accorde avec une remarque précédente.

Après avoir ainsi éprouvé l'intérêt des tableaux de distribution du vocabulaire classique établis par le L.A.S.L.A., selon une méthode très clairement exposée par les auteurs, restait à dresser un tableau rigoureusement semblable pour les oeuvres d'exil de Chrysostome, à partir de la liste de fréquence qui termine le volume I des *Indices Chrysostomici*.

Je l'ai fait, en respectant les mêmes règles de présentation : sur chacune des lignes (numérotées dans la col. 1), est portée, en ordre décroissant, une des fréquences rencontrées (col. 2), puis vient le nombre de mots de même fréquence (col. 3), ensuite le nombre d'occurrences correspondant à cette fréquence (col. 4 = col. 2 x col. 3), le nombre de lemmes ainsi rencontrés (col. 5), le total des occurrences déjà atteint (col. 6); enfin, le pourcentage du texte occupé par la fréquence concernée (col. 7), et en cumulant ces derniers résultats, le pourcentage du texte composé des seuls mots relevés (col. 8).

TABLEAU DE DISTRIBUTION DU VOCABULAIRE DE JEAN CHRYSOSTOME

1	2	3	4	5	6	7	8
1	8301	1	8301	1	8301	14.3999	14.3999
2	5068	1	5068	2	13369	8.7915	23.1915
3	1236	1	1236	3	14605	2.1441	25.3356
4	1144	1	1144	4	15749	1.9845	27.3201
5	908	1	908	5	16657	1.5751	28.8953
6	775	1	775	6	17432	1.3441	30.2397
7	702	1	702	7	18134	1.2177	31.4575
8	699	1	699	8	18833	1.2125	32.6700
9	665	1	665	9	19498	1.1535	33.8236
10	564	1	564	10	20062	0.9783	34.8020
11	461	1	461	11	20523	0.7997	35.6017
12	434	1	434	12	20957	0.7528	36.3546
13	402	1	402	13	21359	0.6973	37.0520
14	390	1	390	14	21749	0.6765	37.7285
15	364	1	364	15	22113	0.6314	38.3599
16	355	1	355	16	22468	0.6158	38.9758
17	349	1	349	17	22817	0.6054	39.5812
18	339	1	339	18	23156	0.5880	40.1693
19	321	1	321	19	23477	0.5668	40.7261
20	310	1	310	20	23787	0.5377	41.2639
21	304	1	304	21	24091	0.5273	41.7912
22	303	1	303	22	24394	0.5256	42.3168
23	288	1	288	23	24682	0.4996	42.8164
24	287	1	287	24	24969	0.4978	43.3143
25	275	1	275	25	25244	0.4770	43.7914
26	274	1	274	26	25518	0.4753	44.2667
27	269	1	269	27	25787	0.4666	44.7333
28	268	1	268	28	26055	0.4649	45.1982
29	251	1	251	29	26306	0.4354	45.6336
30	238	1	238	30	26544	0.4128	46.0465
31	234	1	234	31	26778	0.4059	46.4524
32	232	1	232	32	27010	0.4024	46.8549
33	229	1	229	33	27239	0.3972	47.2521
34	225	1	225	34	27464	0.3903	47.6425
35	223	1	223	35	27687	0.3868	48.0293
36	216	2	432	37	28119	0.7494	48.7787
37	198	2	396	39	28515	0.6869	49.4657

