



Le vocabulaire religieux dans trois siècles de littérature française

Étienne Brunet

► To cite this version:

Étienne Brunet. Le vocabulaire religieux dans trois siècles de littérature française. Bible et informatique, 1988, Jérusalem, Israël. pp.147-165. hal-01574298

HAL Id: hal-01574298

<https://hal.univ-cotedazur.fr/hal-01574298>

Submitted on 12 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Le vocabulaire religieux dans trois siècles de littérature française¹

Etienne Brunet

Introduction

Les linguistes et les lexicologues qui s'enhardissent jusqu'à parler de religion sont-ils des sacrilèges ? Et s'ils exercent leur audace à Jérusalem même, ne faut-il pas les bouter hors du temple ? Et si, plus gravement encore, ils s'arment de machines et sacrifient au veau d'or des temps modernes qu'est l'ordinateur, ne méritent-ils pas qu'on les lapide ? Averti du sort réservé ici même à mon illustre patronyme, Saint Etienne, premier martyr, qui, lui, était un juste, c'est à demi rassuré que je me présente devant vous.

Mais je me dis que cette assemblée n'est pas le Sanhédrin et qu'elle compte un certain nombre de « compteux de mots » aussi impies que moi-même. Je me dis aussi que la religion est pour une grande part une affaire de mots et que les guerres de religion ont souvent été une simple querelle de mots. Puisque la divinité échappe à la perception directe, on ne peut l'approcher que par le truchement des mots. La bible est un livre, la prière une parole, et la foi implique la plénitude des mots et du premier de tous les mots, le Verbe par quoi Dieu se désigne lui-même. Et l'on songe ici au célèbre jeu de mots de Hugo :

Car le mot, c'est le verbe, et le Verbe c'est Dieu.

Je me dis enfin que les écrivains – et Hugo en est un bel exemple – ne sont peut-être pas les meilleurs témoins de Jéhovah et qu'on les voit

1. Article paru dans les *Actes du colloque de Jérusalem « Bible et Informatique »*, juin 1988, Slatkine, Genève, p. 147.165. Seule la mise en page a été revue et corrigée.

plus souvent soucieux de diviniser le langage que de donner un langage à la divinité. Mais s'ils sont parfois de piétres témoins de Dieu, ils peuvent être le reflet et l'écho de leur temps, si bien que l'étude du vocabulaire religieux dans la littérature donne un accès point trop indirect à l'histoire du sentiment religieux.

- I -

1 – L'évolution du roman

a – A l'origine de notre enquête, nous voulions isoler un ensemble de textes où le sentiment religieux ne puisse apparaître que par transparence. Et nous voulions écarter la technicité impudique des ouvrages de théologie dont le vocabulaire montre plus de religion que de sentiment. Nous avons donc choisi un genre profane, le roman, où les thèmes religieux sont rarement au premier plan². Le **corpus** a été constitué par recours à la base de données FRANTEXT. Le logiciel STELLA, créé par J. Dendien, permet très aisément de fixer les limites imposées au genre désiré ($g=ROMAN$), à la période choisie ($d=1860-1892$) ou aux auteurs retenus ($a=HUGO$). Bien entendu ces divers critères peuvent être combinés entre eux. Afin d'avoir des sous-ensembles à comparer, on a divisé en cinq tranches le large espace qui va de 1700 à 1964. On a placé les limites intermédiaires de telle sorte que les tranches aient un volume à peu près équivalent; La plus importante a ainsi 13 millions d'occurrences³, quand la plus mince en compte 11 millions. En gros chaque tranche recouvre un demi-siècle, sauf la première qui s'étend sur tout le XVIII^e siècle, jusqu'à la Révolution. Voici le nombre et l'étendue des textes passés en revue :

	<i>tranche 1</i>	<i>tranche 2</i>	<i>tranche 3</i>	<i>tranche 4</i>	<i>tranche 5</i>	<i>total</i>
	1700-1788	1789-1859	1860-1892	1893-1926	1927-1964	1700-1964
<i>textes</i>	204	152	134	147	149	786
<i>occurr.</i>	13.002.038	11.143.710	11.033.320	10.361.700	10.989.320	56.530.088

2. Il n'y a pas d'exclusive cependant, et l'on cite facilement des romans qui inscrivent le thème religieux dans le titre même, comme *La religieuse* de Diderot, et beaucoup de textes de Huysmans (*L'oblat*, *Les foules de Lourdes*) ou de Bernanos (*Sous le soleil de Satan*, *Le journal d'un curé de campagne*, *Dialogues des Carmélites*).

3. Le logiciel STELLA prend la notion d'occurrence dans un sens extensif, qui inclut non seulement les mots mais aussi les signes de ponctuation et de typographie.

Avec 56 millions d'occurrences la surface du texte exploré est considérable, la production romanesque occupant plus du tiers du corpus du T.L.F.

b – Mais reste à établir la part de la **religion** dans cette masse lexicale. On ne saurait demander à la machine d'avoir quelque sens religieux. On doit donc lui donner quelques leçons de catéchisme et dresser une liste des termes qui expriment d'une manière ou d'une autre l'appel de l'au-delà. Ici l'arbitraire est inévitable : la sensibilité religieuse s'exprime de mille manières, dont la plus courante se confond avec l'expression des choses humaines. Le mysticisme emprunte souvent les mêmes voies que le discours amoureux et Thérèse d'Avilla utilise certains des mots brûlants que Racine met dans la bouche de Phèdre. Et dans un domaine où les particularités individuelles jouent un si grand rôle, la charge sémantique du vocabulaire est fort variable, tel mot ayant pour l'un des connotations surnaturelles qui échappent à tel autre. Et plus gravement, suivant qu'on défend ou qu'on attaque la religion, les mêmes mots se remplissent ou se vident de leur contenu. Le mot *Dieu*, qui est si riche pour le fidèle, n'est plus qu'un mot creux pour l'incroyant. Sans sous-estimer ces difficultés – qui se retrouvent chaque fois qu'on aborde le domaine sémantique – nous avons cherché à délimiter un champ, en puisant dans un dictionnaire analogique (le *Petit Robert*) les éléments cités aux articles *religion* et *religieux*. Ces éléments se dispersent dans un espace assez confus où le centre se distingue mal de la circonférence. On y trouve des noms propres, qui désignent la divinité elle-même (*Dieu, Allah, Jéhovah, Jésus, Jupiter, Zeus*), ou ses représentants sur terre (*Confucius, Mahomet, Luther*), ou des lieux saints (*Jérusalem*), mais il n'a pas été possible d'inclure les saints, parce que leur nom ne peut être distingué des prénoms donnés à des personnages souvent fort éloignés des inquiétudes religieuses⁴. Quant au vocabulaire commun, on s'est contenté de la catégorie nominale, car on trouverait difficilement des mots grammaticaux, des adverbes⁵, ou même des verbes qui aient une exclusivité religieuse. Comme parmi les noms propres, le syncrétisme est avoué : les manifestations lexicales du phénomène religieux font

4. On a maintenu toutefois dans la liste *Abraham* et *Judas*, qui servent peu à l'état civil, même si l'on peut citer des exemples célèbres.

5. On a fait exception pour quelques adverbes en *ment*, comme *pieusement* et *religieusement*, dont le sort est lié à celui de l'adjectif correspondant.

nécessairement meilleur accueil, dans une langue de culture chrétienne, aux traditions catholiques. Mais on n'a pas écarté les mots les plus connus qui désignent quelque objet de l'Islam, du judaïsme ou du bouddhisme. Au total 219 formes brutes ont été retenues. Mais certaines n'avaient aucune occurrence dans le corpus, comme *confucianisme*, *manichéiste*, *shintôisme*, *hindouisme*, *coranique*, *druidisme*, *totémisme* ou *védisme*. Et d'autres en avaient si peu que cela interdisait toute exploitation statistique (comme *théisme*, *pharisaïsme*, *mufti*, *panthéisme*, *fétichisme*, *ecclésiastique*). S'agissant de la classe nominale, la lemmatisation a pu être faite à peu de frais, en neutralisant la marque du pluriel. On a réduit ainsi à 127 le nombre de vocables recensés.

Tableau 1. Le vocabulaire religieux dans le roman français.

Effectifs absolus

tranche	1700-1788	1789-1859	1860-1892	1893-1926	1927-1964	1700-1964	1700-1788	1789-1859	1860-1892	1893-1926	1927-1964
textes	roman1	roman2	roman3	roman4	roman5	TOTAL	roman1	roman2	roman3	roman4	roman5
mots	204	152	134	147	149	786	13002038	11143710	11033320	10361700	10989320
	56530874	écarts réduits									
Abraham	39	34	23	17	6	119	-5,37	-4,87	-6,14	-6,50	-8,17
Allah	3	5	4	21	5	38	-3,09	-2,11	-2,38	3,04	-2,07
Babel	3	5	10	15	4	37	-2,87	-1,88	-0,27	1,56	-2,16
Babylone	101	72	52	33	11	269	4,59	2,20	-0,50	-2,76	-6,20
Bouddha	0	4	3	7	1	15	-2,97	-1,14	-1,52	0,25	-2,31
Christ	147	381	405	440	219	1592	-22,08	-9,53	-8,30	-5,42	-16,12
Confucius	2	4	5	2	0	13	-3,66	-2,72	-2,42	-3,12	-3,82
Dieu	2835	4243	2083	3588	2880	15629	-35,62	-4,81	-37,57	-10,32	-1,00
Islam	0	0	0	3	8	11	-4,46	-4,10	-4,07	-3,12	-4,10
Jéhovah	0	5	10	11	2	28	-4,44	-2,76	-1,40	-0,92	-1,94
Jérusalem	38	261	51	70	22	442	-12,68	5,81	-10,15	-8,07	-3,52
Jésus	167	306	476	426	235	1610	-23,18	-14,79	-7,76	-8,37	-12,35
Judas	2	49	28	10	15	104	-5,89	3,50	-0,40	-3,57	-17,37
Jupiter	108	175	49	29	13	374	-7,92	-1,04	-10,49	-11,47	-2,85
Lucifer	18	6	9	6	12	51	0,29	-2,31	-1,46	-2,10	-13,19
Luther	4	15	5	8	6	38	-7,02	-4,66	-6,17	-5,43	-0,62
Mahomet	129	74	14	12	1	230	4,67	-0,10	-7,26	-7,16	-6,00
Neptune	8	22	12	23	0	65	-4,63	-1,32	-3,22	-0,76	-8,81
Satan	60	330	68	79	26	563	-5,25	24,55	-3,00	-1,24	-5,54
Zeus	0	2	1	20	3	26	-4,95	-4,08	-4,29	0,54	-7,42
abbaye	79	86	63	208	27	463	-3,87	-1,68	-3,96	12,00	-3,80
abbé	819	1285	1725	1403	1401	6633	-4,35	17,29	33,80	24,48	-7,65
abbesse	115	124	5	30	6	280	8,30	11,46	-6,40	-2,25	22,07
ange	425	1181	388	545	432	2971	-12,96	21,20	-10,73	-2,80	-6,23
angélique	49	123	46	68	36	322	-1,43	10,49	-0,66	3,10	-8,85
archange	2	34	37	40	35	148	-6,13	0,24	0,85	1,81	-2,09
athée	46	41	29	61	25	202	-5,02	-4,43	-5,77	-1,49	0,51
athéisme	11	21	10	26	9	77	-6,02	-3,75	-5,41	-2,57	-6,22
bible	30	92	91	93	93	399	-9,96	-2,96	-2,95	-2,12	-5,54
bonze	25	4	5	6	8	48	2,36	-2,72	-2,42	-1,96	-2,72
bouddhisme	0	2	9	0	1	12	-4,88	-4,01	-2,28	-4,31	-1,56
bouddhiste	0	1	4	1	2	8	-3,26	-2,64	-1,54	-2,51	-4,21
brahmane	0	14	5	11	2	32	-3,19	2,23	-1,06	1,38	-2,25
cathédrale	11	154	193	200	173	731	-14,47	-2,16	0,99	2,46	-2,17
catholisme	2	27	38	57	20	144	-10,51	-6,88	-5,60	-3,00	-0,50
catholique	140	109	174	235	226	884	-15,80	-15,10	-11,56	-7,33	-7,56
céleste	347	483	161	171	75	1237	-1,77	9,07	-9,34	-7,85	-8,78
cérémonie	557	306	290	299	194	1646	10,31	-0,56	-1,32	0,34	-14,28
chapelle	107	289	393	438	250	1477	-11,27	2,65	9,71	14,09	-6,90
christianisme	104	110	29	52	56	351	-14,10	-11,72	-16,42	-14,36	0,30
etc...											
Total généré	28402	32405	25587	28633	19837	134864					

Le tableau 1 montre un extrait des effectifs ainsi obtenus dans notre corpus romanesque. Comme les cinq sous-ensembles ont à peu près la même étendue la lecture peut être faite directement sur les effectifs bruts. Ainsi l'on voit que le mot *Dieu*, par exemple, a deux fois plus d'occurrences à l'époque romantique (4243 occ.) qu'à l'époque naturaliste (2083 occ.). Cependant la rigueur nous oblige à choisir une mesure probabiliste, non seulement pour ajuster les sous-ensembles à la même échelle, mais aussi pour relativiser les écarts, lorsque ceux-ci peuvent être imputables au hasard – ce qui risque de se produire dans le cas des faibles effectifs. En suivant la tradition établie dans notre discipline, on choisira l'*écart réduit*⁶ comme mesure de pondération. Et le tableau 1 offre dans sa partie droite la conversion en écarts réduits des effectifs réels établis dans la partie gauche.

Certains ne manqueront pas d'observer que les écarts réduits ont le plus souvent le signe négatif. Ainsi le mot *Dieu* a des déficits dans les cinq sous-ensembles, au temps du naturalisme certes ($z = -37,57$), mais aussi dans la période romantique ($z = -4,81$). Or on attendrait que les excédents compensent les déficits. Cela se produirait en effet si l'on s'enfermait dans le corpus romanesque, en ignorant les autres genres, c'est-à-dire si l'on établissait la probabilité p en faisant le rapport de l'étendue d'une tranche à l'ensemble de la production romanesque, soit, pour la période romantique,

$$p = 11143710 / 56530088 = 0,19713$$

On a préféré prendre pour base de calcul l'ensemble du corpus, tous genres réunis, ce qui donne un résultat différent⁷ :

$$p = 11143710 / 147279939 = 0,07566$$

Quelle que soit la période considérée, le vocabulaire religieux apparaît donc sous-représenté dans le roman. C'est que d'autres genres lui font

6. Rappelons la formule de l'écart réduit :

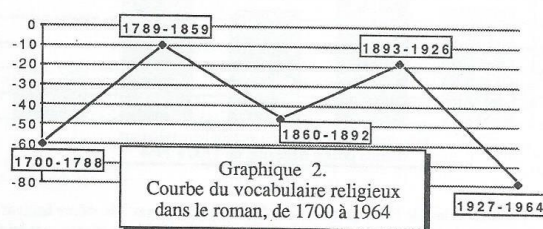
$$z = \frac{\text{effectif réel} - \text{effectif théorique}}{\sqrt{(\text{effectif théorique} \times \text{prob } q)}}$$

(la probabilité q étant le complément à la probabilité p qui mesure le rapport du sous-ensemble à la totalité). Voir Ch. Muller, *Initiation à la statistique linguistique*, Hachette, 1973.

7. Bien entendu la probabilité p s'applique, selon la perspective considérée, au total des occurrences du mot dans le genre romanesque, ou au total du même mot dans le corpus entier, soit, pour le mot *Dieu*, 15629 et 60221 respectivement.

meilleur accueil, comme la poésie et certains essais spécialisés, ce que nous verrons plus loin.

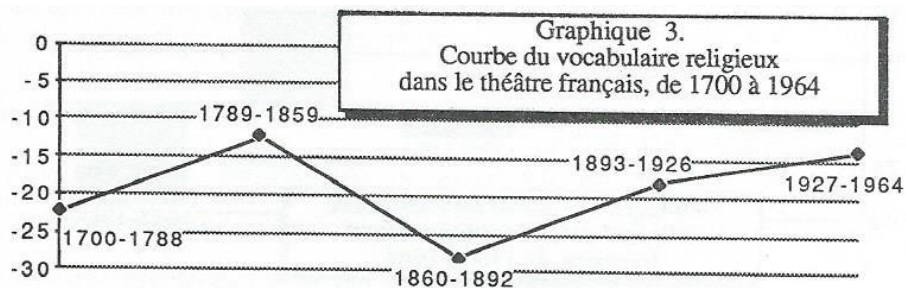
c – Mais dans une première approche, tenons-nous en au genre romanesque, en observant le graphique 2, qui reproduit la distribution dans le temps du vocabulaire religieux. La courbe montre que la présence religieuse n'est pas constante : faible au XVIII^e siècle, jusqu'à la Révolution, elle culmine dans la première moitié du XIX^e, au moment où les valeurs du sentiment prennent le pas sur le rationalisme, ce qui profite à la religion. Le *Génie du christianisme* donne ses lettres de noblesse à la sentimentalité et à l'esthétique religieuses, et abat la séparation implicite que l'âge classique avait établie entre le profane et le sacré, entre la littérature et la religion. Mais à mesure que le romantisme s'efface et que les valeurs de la science prennent le relais, le roman délaisse les thèmes religieux et la courbe s'affaisse. C'est l'époque du scientisme et du naturalisme⁸. Mais une réaction spiritualiste se produit à la fin du XIX^e siècle, qui se développe dans les premières décennies du XX^e siècle et qui donne au vocabulaire religieux un regain d'actualité. Sans doute les grands écrivains spiritualistes comme Claudel, Péguy ou Bergson n'entrent-ils pas ici en ligne de compte, puisqu'ils ne sont pas romanciers. Mais le ciel est dans l'air du temps et les romanciers font écho aux préoccupations spiritualistes de l'époque, ce qui s'observe chez Léon Bloy, Huysmans, Barrès, Romain Rolland, Bourget, et les premiers romans de Mauriac. Le mouvement de bascule se produit encore dans le dernier terme du graphique, qui montre un effondrement. Ni le roman existentialiste, ni le nouveau roman ne semblent faire appel à la conscience religieuse. Mais notre corpus s'arrête en 1964 et le dernier quart de siècle que nous venons de vivre infléchirait peut-être le sens de la courbe.



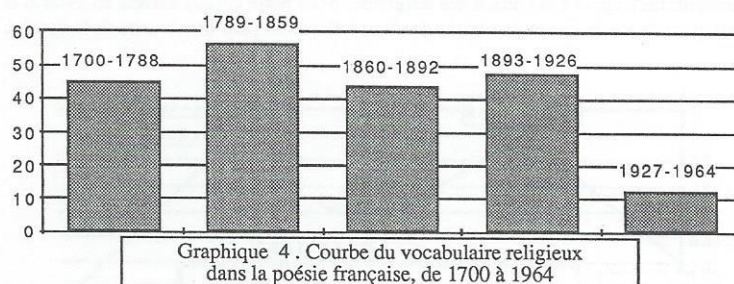
8. Il arrive cependant que le roman naturaliste aborde les milieux où dominent les traditions religieuses. C'est une enquête semblable aux autres que Zola entreprend parfois dans le monde ecclésiastique, par exemple dans *La faute de l'abbé Mouret*, *La conquête de Plassans* ou *Le rêve*. En de tels cas le roman ne se nourrit que d'observations extérieures et non de la religiosité latente qui animait beaucoup des écrivains romantiques.

2 – L'évolution des autres genres

a – Mais il convient de vérifier si les autres genres littéraires confirment ce mouvement. Les procédures qui ont permis d'acquérir les données pour le roman ont été renouvelées pour le **théâtre**. Là encore la religion n'est pas directement sur scène, le théâtre ayant depuis longtemps rompu les attaches qui ont pu à l'origine, dans l'antiquité ou au Moyen Age, relier la représentation théâtrale à la cérémonie religieuse et les tréteaux au parvis des cathédrales. Le résultat (graphique 3) reproduit celui qu'on vient d'observer dans le roman : même sommet au début et à la fin du XIX^e siècle, même fléchissement au début et au centre de la chronologie, et même tendance générale au déficit. La seule différence notable vient de la dernière tranche qui fait la part belle à la religion et à laquelle le théâtre de Claudel n'est pas étranger.

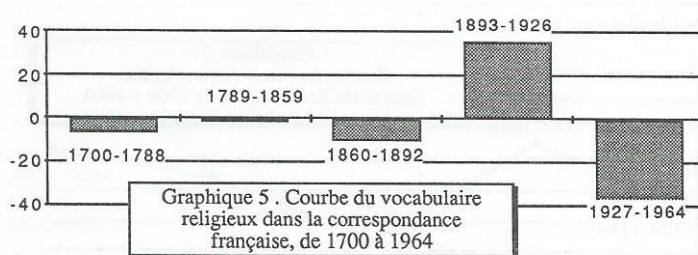


b – Le genre **poétique** diffère radicalement des deux précédents, car il entretient des liens étroits avec la religion. La poésie est comme la religion affaire de sensibilité et toute une frange du vocabulaire est commune aux deux domaines. Quand le poète parle d'*ange*, de *démon*, de *mystère*, de *pasteur*, de *foi*, de *temple* et quand les adjectifs *sacré*, *saint*, *divin*, *céleste*, *mystique*, *pieux* viennent sous sa plume, l'emploi du vocabulaire religieux est-il seulement métaphorique ou bien le *rêveur sacré*, pour reprendre l'expression de Hugo, n'est-il pas un mage, qui communique avec le spirituel ? Quoi qu'il en soit l'interférence de la vision poétique et du sentiment religieux n'est pas douteuse dans le graphique 4, représenté ci-dessous. Cette fois toutes les tranches chronologiques appartiennent à la zone excédentaire.



Les branches de l'histogramme ne s'élèvent pourtant pas à la même hauteur et les mouvements de l'histoire, décelés au théâtre ou dans le roman, se répercutent ici clairement. Les deux sommets s'observent dans les phases de plus grande faveur religieuse : à l'ère romantique et au tournant spiritualiste qui accompagne la transition du XIX^e au XX^e siècle. On ne tirera pas de conclusions excessives du brusque affaissement de la dernière tranche, qui tient surtout au faible nombre de textes poétiques durant cette période (l'étendue n'est plus que de 244.432 mots, quand il s'élève à 2.186.955 dans la période romantique)⁹.

c – Restent deux genres, qui appellent des commentaires voisins. Le premier correspond à la littérature écrite à la première personne. Il s'agit essentiellement de mémoires, de récits de voyages et de lettres. Nous l'appelons **correspondance**.

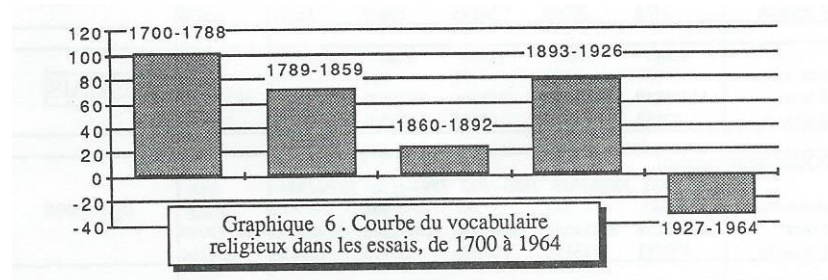


Le sommet y culmine dans la quatrième tranche, ce qui laisse entendre que le renouveau spiritualiste n'est pas une mode de surface, mais un mouvement de fond, qui est ancré au profond de la conscience et qui émerge quand le moi intérieur s'exprime au dehors. Là encore les

9. Ici intervient la flexibilité de l'écart réduit, qui est dépendant non seulement des écarts, mais aussi de la taille des corpus considérés.

déficits s'observent aux deux bouts et au centre de la chaîne chronologique.

Le dernier genre, qui avait provoqué nos réticences initiales, est celui des **essais**. La religion ici n'est plus en rapport avec la sensibilité, mais avec la raison. Elle devient un objet de discussion, soit qu'on la prenne à partie, comme fait Voltaire, soit qu'on l'explique ou qu'on la défende. L'ambiguïté des données ne permet pas qu'on distingue les deux points de vue. Au moins peut-on observer que le fait religieux est de moins en moins l'objet des batailles intellectuelles et que la religion a cessé d'être une question brûlante et même une question d'actualité. C'est l'enjeu capital au XVIII^e siècle et c'est là qu'on relève le sommet de la courbe (graphique 6). Puis après les secousses de la Révolution et l'apaisement du Concordat, la religion, devenant plus individuelle et moins institutionnalisée, suscite de moins en moins de controverses, jusqu'à la flambée de la fin du siècle qui naît avec l'affaire Dreyfus et conduit, au début du siècle suivant, à la séparation de l'Eglise et de l'Etat. L'époque contemporaine semble avoir classé le dossier et peu d'ouvrages raniment la querelle. La religion gagne en liberté ce qu'elle perd en actualité. Elle se dissout dans l'acceptation générale, et peut-être même dans l'indifférence. Et le silence que traduisent les chiffres négatifs est sans doute l'écho sourd de cette indifférence.



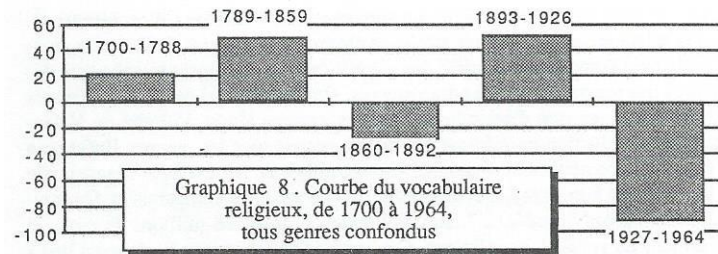
3 – L'évolution générale

Si l'on neutralise le genre littéraire et qu'on observe la tendance générale, tous genres confondus, les observations portent sur un vaste ensemble chronologique qui réunit près de 150 millions de mots, dont un demi-million appartiennent au vocabulaire religieux (voir les effectifs absolus dans le tableau 7). La courbe obtenue (graphique 8) combine les deux mouvements qu'on vient de décrire dans les genres. D'une part on observe un mouvement de balancier qui alterne les moments de ferveur et les crises du sentiment religieux. Les ondulations produisent deux bosses

et trois creux comme dans les courbes précédentes. Mais on note en même temps une évolution linéaire : la tête et la queue du serpent ne sont pas au même niveau. Il y aurait donc une dérive du temps à travers les méandres de l'histoire. Et le vocabulaire religieux, même soutenu par le sentiment poétique, ne se renouvelle guère. Si l'on prolongeait l'observation plus loin dans le passé, il semble bien que ce mouvement de désaffection serait plus net encore. La religion des temps modernes n'est plus l'expression des temps théocratiques et le ciel s'enfonce sous terre. Mais peut-être peut-on soutenir que le fait religieux se manifeste dans des avatars imprévus où le veau d'or de la bible fait peau neuve.

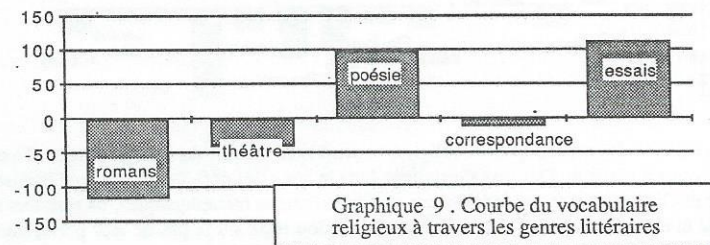
Tableau 7. Les effectifs bruts

tranche	1700-1788	1789-1859	1860-1892	1893-1926	1927-1964			
	roman1	roman2	roman3	roman4	roman5	roman		
nombre textes	204	152	134	147	149	786	ROMANS	
total mots	13002038	11143710	11033320	10361700	10989320	56530088		
total religion	28402	32405	25587	28633	19837	134864		
	théâtre1	théâtre2	théâtre3	théâtre4	théâtre5	théâtre		
nombre textes	106	52	21	42	69	290	THEATRE	
total mots	2357953	1443645	799534	1275650	2447804	8324586		
total religion	5345	3612	1059	2779	6324	19119		
	poésie1	poésie2	poésie3	poésie4	poésie5	poésie		
nombre textes	39	58	37	36	11	181	POESIE	
total mots	782612	2186955	1150298	834729	244432	5199026		
total religion	4610	11318	6124	4973	1099	28124		
	corresp.1	corresp.2	corresp.3	corresp.4	corresp.5	corresp		
nombre textes	53	122	62	62	68	367	LETTRES	
total mots	3353800	9942422	5175898	5049636	5350284	28872040		
total religion	9543	30098	14595	19867	11573	85676		
	essai1	essai2	essai3	essai4	essai5	essais		
nombre textes	181	193	33	111	378	896	ESSAIS	
total mots	11315850	10795850	1743746	5575533	18923220	48354199		
total religion	52883	45743	7087	27361	50277	183351		
EPOQUES (récapitulation)								
	1700-1788	1789-1859	1860-1892	1893-1926	1926-1964	total	époques	
nombre textes	583	577	287	398	675	2560		
total mots	30812253	35512582	19902796	23097248	37955060	147279939		
total religion	100783	123176	54452	83613	89110	451134		
GENRES (récapitulation)								
	roman	théâtre	poésie	corresp	essais	total	genres	
nombre textes	786	290	181	367	896	2520		
total mots	56530088	8324586	5199026	28872040	48354199	147279939		
total religion	134864	19119	28124	85676	183351	451134		
AUTEURS								
	Marivaux	Diderot	Voltaire	Rousseau	Chateaubr.	Hugo	Balzac	Flaubert
nombre textes	25	37	53	21	24	45	37	35
total mots	1102765	1723488	2167960	1290495	1963832	3162133	2586038	2928036
total religion	1811	4661	16942	2775	11873	12642	6212	7712
	Zola	Proust	Valéry	Gide	Giraudoux	Camus	Sartre	TOTAL
nombre textes	45	23	10	45	16	23	10	449
total mots	4144167	1712253	658870	2394911	755582	1055503	532213	28178246
total religion	9329	2056	1145	2056	1926	2818	867	84825



4 – La typologie des genres

De même qu'on vient de neutraliser les genres, ainsi peut-on négliger la chronologie pour observer les oppositions des genres. Ici encore c'est la totalité des 150 millions de mots qui est explorée et qu'on répartit dans les cinq cases relatives au genre. Les effectifs relevés sont transcrits dans le tableau 7. On les convertit immédiatement en écarts réduits, et en graphique, ce dont rend compte le graphique 9.

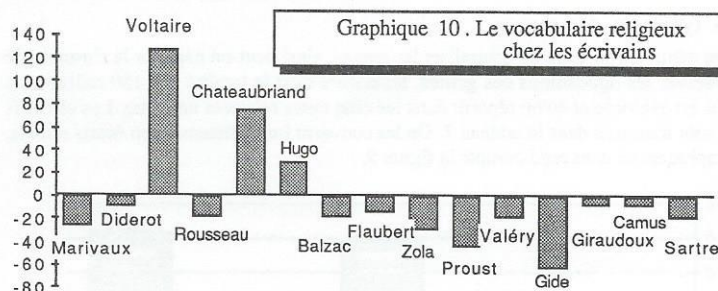


Il est inutile de s'attarder sur cette courbe récapitulative qui confirme ce qu'indiquaient les courbes précédentes. Le roman est très déficitaire, en ce qui concerne le vocabulaire religieux. Le théâtre l'est aussi, mais moins gravement. Et la correspondance se rapproche de la zone médiane. Dans la zone des excédents on relève les textes techniques, pour une raison accidentelle qui tient au choix des sujets. La théologie et les affaires religieuses constituent en effet un thème ou plus exactement une matière, qui se prête aux développements techniques et aux essais. Par contre les accointances de la poésie et de la religion paraissent plus essentielles, plus profondes, et, comme nous l'avons vu, plus résistantes aux aléas de l'histoire.

5 – La spécificité des écrivains

S'agissant du phénomène religieux, la variable essentielle n'est peut-être pas le temps, ni le genre, mais le tempérament de l'écrivain. Car le

sentiment religieux est propre à chacun et dépend peu des modes et des écoles. Il a donc fallu constituer un nouveau corpus qui laisse aux écrivains leur individualité sans les fondre dans des groupes ou des tranches chronologiques. Quatre représentants du XVIII^e siècle ont été choisis : Marivaux, Diderot, Voltaire et Rousseau. Le XIX^e siècle a délégué cinq auteurs : Chateaubriand, Balzac, Hugo, Flaubert et Zola, et le XX^e six autres : Proust, Valéry, Gide, Giraudoux, Camus et Sartre. Plusieurs considérations ont inspiré ce choix : d'une part on n'a voulu retenir que de grandes monographies, qui soient dignement représentées dans la base FRANTEXT, et dont l'étendue atteigne en moyenne deux millions d'occurrences¹⁰. D'autre part la sélection couvre les trois siècles et chacun des moments de l'histoire, mais aussi le dosage tient compte de la variété des genres : certains des noms retenus sont des poètes, d'autres des romanciers ou des dramaturges. On trouve aussi des auteurs d'essais ou de mémoires, et beaucoup d'épistoliers. Certains, comme Hugo, Voltaire ou Valéry ont d'ailleurs une production multiforme qui met en oeuvre tous les genres. Enfin on a écarté les écrivains trop nettement catalogués comme religieux, par exemple Claudel, afin de garder une certaine homogénéité et de donner plus de sens à la comparaison. On trouvera les effectifs au bas du tableau 7, soit 449 textes au total, 28 millions de mots, et 84.825 attestations du vocabulaire religieux. Convertis en écarts réduits ils donnent lieu à la représentation graphique 10.



Le déclin des préoccupations religieuses semble se vérifier dans ce graphique. Tous les écrivains à partir de Flaubert s'installent dans la zone des déficits. Et quoique Proust, Valéry et Gide ne soient pas indifférents

10. Avouons aussi qu'ayant réalisé plusieurs monographies, sur Rousseau, Chateaubriand, Hugo, Zola, Proust et Giraudoux, on a vu ici l'occasion de jeter un pont entre elles, sur un point particulier.

aux spéculations métaphysiques, ils résistent à Claudel et ne se laissent pas convertir¹¹. Et la religion reste sur le pas de leur porte, sans franchir le seuil ni envahir leur vocabulaire. En réalité trois écrivains seulement ont des excédents, il est vrai très larges : Voltaire, Chateaubriand et Hugo. Il y aurait beaucoup à dire sur la religion de Voltaire et nous ne pouvons que renvoyer le lecteur à la thèse magistrale de Marie-Hélène Cotoni. La religion institutionnalisée est sans doute l'ennemi qu'il faut abattre, l'infâme qu'il faut écraser. Mais Voltaire mène ce combat avec trop de passion pour qu'on ne le soupçonne pas de vouloir étouffer une voix qui est en lui et qui ne se tait pas. Quoi qu'il en soit, la statistique ne démêle pas les intentions, elle observe seulement l'omniprésence des termes religieux sous la plume de Voltaire. Les choses sont moins compliquées chez Chateaubriand, qui, d'abord réticent à l'égard de l'Eglise, se convainc sous l'Empire des vertus de la religion, de ses dogmes, de ses commandements et de ses rites et en fait l'apologie dans le *Génie du christianisme*. Quant à Hugo, qui apprécie peu l'Eglise comme puissance temporelle, il croit en Dieu avec ferveur et résistera jusqu'au bout à ses amis républicains qui aimeraient le voir abjurer¹². Et la figure de l'évêque des *Misérables* est plus belle que tout ce qu'on a vu dans les hagiographies.

- II -

On vient de rendre compte de la distribution des mots religieux dans les genres, les périodes et parmi les écrivains sans voir aucun d'entre eux. Seul comptait jusqu'ici l'effectif des sous-ensembles considérés. Mais on peut souhaiter entrer dans les détails et considérer les mots individuels. On n'ira pas jusqu'à examiner pour chaque mot combien d'occurrences sont relevées dans chacune des 25 cases qu'on trouve au croisement de 5 genres et de 5 périodes. On se contentera des tableaux qui rendent compte de la chronologie, ou des genres, sans croiser ces deux facteurs. Il y a deux façons de détailler ces tableaux. Ou bien on s'attache à un mot particulier et l'on trace la courbe de la ligne correspondante. Ou bien on s'intéresse à un genre, ou à une période, ou à un auteur, dont on veut

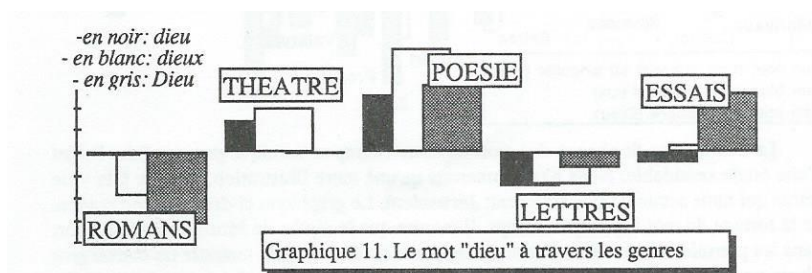
11. On ne laisse pas d'être étonné par l'importance du déficit relevé chez Gide. Car le lecteur de Gide est sans cesse sollicité par des souvenirs et des citations de la bible; Mais l'émergence de la culture religieuse est indirecte, souvent ironique, et presque toujours allusive.

12. Voir sur ce point l'analyse de Henri Guillemin, dans *Victor Hugo par lui-même*, p. 76- 90 et dans *Victor Hugo et l'Eglise* (Genève, 1944).

dresser le profil lexical, et c'est alors la colonne correspondante qui fait l'objet du graphique. Considérons d'abord une ou deux lignes.

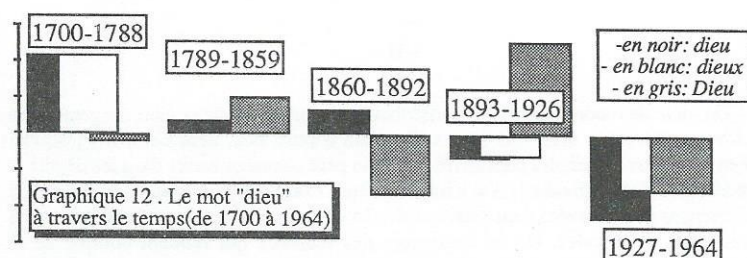
1 – Les mot *dieu* et *Jérusalem*

A tout Seigneur, tout honneur. Commençons par *Dieu*. En réalité derrière ce mot on distingue une trinité lexicale. Si le mot a la majuscule, c'est la divinité en personne qui est désignée. Si l'on a le pluriel, il ne s'agit plus que de dieux païens. Reste le cas du singulier avec la minuscule. Il peut s'agir encore d'un dieu antique ou exotique, ou d'une simple métaphore. Nous n'avons pas voulu confondre ces emplois, qui donnent lieu à trois courbes distinctes et juxtaposées (graphique 11).

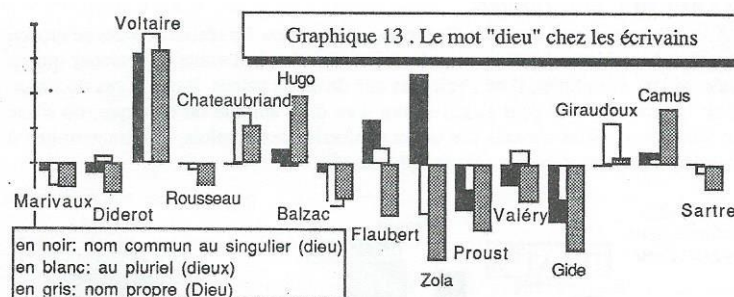


Les trois formes s'accordent à peu près dans le choix des genres. La divinité refuse d'apparaître dans le roman et ne se plaît guère non plus dans la correspondance. C'est en terre poétique qu'elle établit sa demeure privilégiée. Mais la trinité se divise lorsqu'il faut choisir entre le théâtre et les essais. La minuscule est admise sur la scène, parce que certains dramaturges modernes, comme Giraudoux, Claudel, Sartre ou Anouilh, exploitent les mythes grecs et représentent des Electre ou des Antigone assaillies par les dieux. La majuscule qui sied à la majesté divine préfère les ouvrages sérieux, les essais ou traités théologiques.

Et de la même façon c'est *Dieu* en personne qui est en faveur dans les époques de croyance religieuse, à l'époque romantique et lors du renouveau spiritualiste de 1900 (les *dieux* restant dans la zone négative). Inversement le XVIII^e siècle a quelque pudeur à s'adresser directement à un *Dieu* personnel, et c'est un *dieu* moins exigeant et plus conventionnel qui apparaît, comme dans la tradition classique (graphique 12).



La répartition selon les écrivains est riche d'enseignement. Si Voltaire admet toutes les variantes, Marivaux, Balzac, Proust et Gide les écarte toutes. Les écrivains portés vers Dieu, comme Chateaubriand¹³ et Hugo, utilisent la majuscule. Par contre les réalistes et les naturalistes se contentent de la minuscule, ce qui est bien suffisant pour un juron (graphique 13).

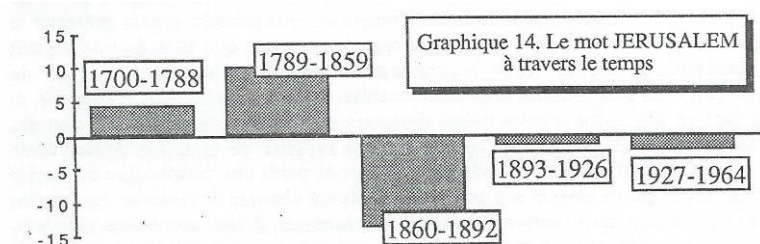


La distribution de chacun des mots de notre champ sémantique pourrait faire l'objet d'une étude semblable. Nous n'en donnerons qu'une autre illustration, relative à la ville sainte qui nous accueille présentement : **Jérusalem**. Le graphique 14 ci-dessous rend compte de la fortune du mot à travers le temps. Il montre que le mythe de Jérusalem est plus fort dans les premières tranches, non seulement parce qu'il est le pôle terrestre où convergent plusieurs grandes religions, mais parce que la génération romantique, que l'expédition d'Egypte de Bonaparte a fait rêver et qui se souvient des croisades, aspire à la découverte du monde et se lance dans des expéditions lointaines, dont Jérusalem est souvent le but

13. L'intrigue particulière des *Martyrs*, dont le cadre est grec et antique, explique l'abondance des *dieux* chez Chateaubriand.

et le couronnement. Chateaubriand avec son *Itinéraire de Paris à Jérusalem* et Lamartine avec son *Voyage en Orient* en portent témoignage¹⁴. Jérusalem a donc alors un double visage : l'un diaphane, l'autre profane. L'un est l'image de la ville sainte, l'autre reflète les sortilèges de l'orient. Ces deux valeurs semblent s'être un peu estompées de nos jours, comme l'indique le graphique. Mais un autre visage est apparu dans les dernières décennies, qui est historique et politique. Jérusalem est devenue la capitale d'un Etat. Mais notre corpus s'arrête trop tôt pour que la trace de cette transformation y soit visible.

On peut encore creuser une galerie plus profonde et obtenir pour un mot particulier la liste de tous les contextes qui l'environnent. C'est ce que nous avons fait pour le mot *Jérusalem*, au moins dans les textes poétiques. Comme les quatre syllabes du mot sont plus faciles à placer dans un verset que dans un vers, la plupart des 68 occurrences sont relevées chez Péguy (16 emplois) et Claudel (22). Mais on en trouve 11 dans les vers de Hugo, surtout dans la *Légende des siècles*. Les autres poètes sont plus discrets : 3 exemples chez Aragon, 2 chez Voltaire, un seul chez Lamartine. Partout le sens religieux du mot est apparent, soit qu'on déplore les malheurs de la ville sainte, soit qu'on évoque les épisodes de la vie de Jésus, soit qu'on fasse de ce haut lieu le symbole de la terre promise et du pacte avec Dieu.



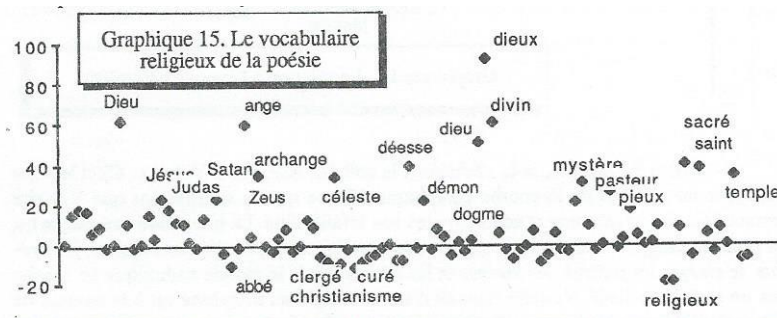
Mais on ne saurait s'appesantir sur chacun des emplois de chacun des mots de notre corpus. La vision globale d'un corpus de cette dimension impose qu'on prenne du recul, même si à distance les lignes finissent par se brouiller. Mais avant de prendre du champ, observons quelques colonnes.

14. Notons que si le récit de Lamartine est incorporé à la base de données, celui de Chateaubriand n'en fait pas partie.

2 – Le profil religieux d'un genre, d'une période, d'un écrivain

Les colonnes sont plus riches d'information que les lignes, car elles sont plus longues. Chacune contient plus de cent données, ce qui permet de dessiner des profils de plus grande précision.

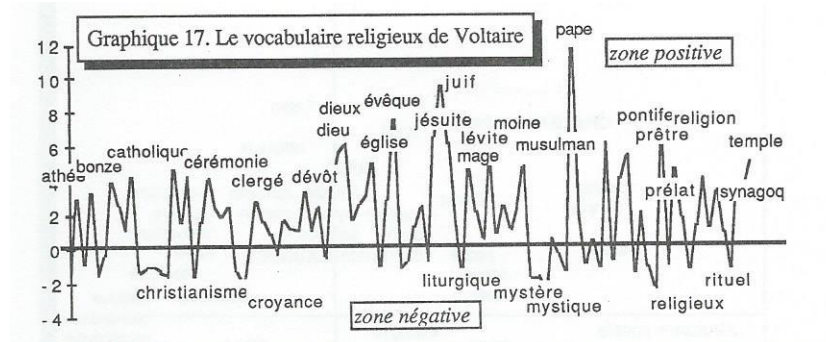
a – Le graphique 15, représenté ci-dessous, indique le choix de la **poésie** dans la boutique lexicale où la religion offre ses produits.



Les poètes n'ont que faire du clergé et de l'administration ecclésiastique, non plus que des dogmes et des débats théologiques qui s'expriment avec des mots en *isme* (comme par exemple le *christianisme*). Les unités abstraites du vocabulaire religieux sont déficitaires en poésie, à commencer par le mot *religion* et l'adjectif *religieux*. Les unités trop concrètes et trop épaisses comme un *abbé* ou un *curé* n'ont pas non plus sa faveur : ce sont plutôt des personnages de roman. Ce qui attire les poètes dans la religion, c'est ce qui est sensible et vague à la fois, et beaucoup d'adjectifs ont cette double qualité : *sacré*, *saint*, *pieux*, *divin*, *céleste*. La divinité elle-même a ces deux attributs, sous quelque forme, antique ou moderne, bienveillante ou maléfique, qu'elle apparaisse : *Dieu*, *dieu*, *dieux*, *déesse*, *Zeus*, *Jésus*, *Satan*, *ange*, *archange*, *démon*. Le mot *mystère* résume ce que le poète demande à la religion.

b – On pourrait opposer à la courbe de la poésie celle du roman, ou celle des essais, la première débordant du côté des réalités tangibles, la seconde du côté des notions abstraites. Mais, comme la place manque, nous changeons de point de vue, pour considérer la perspective chronologique. Prenons pour exemple la tranche 4, qui va de 1860 à 1892. L'histogramme 16, qui rend compte de cette période, semble présenter le négatif du portrait précédent. Les personnages ensoutanés que la poésie dédaignait prennent ici leur revanche : l'*abbé*, le *curé*, le

laisse Voltaire indifférent, si ce n'est la surface des pratiques religieuses (*liturgique, rituel* ont des écarts négatifs), et les profondeurs de la foi (déficit de *croissance, mystère, mystique*).



3 – Synthèse

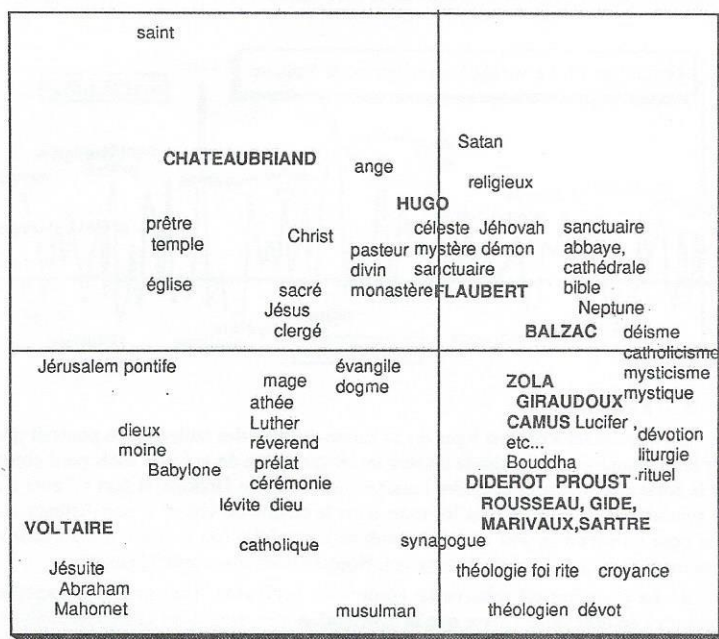
En s'attachant à chaque ligne ou à chaque colonne des tableaux, on pourrait multiplier les instantanés et prendre la mesure de chaque arbre de la forêt. Mais pour contempler la forêt elle-même et entendre l'oracle des chênes de Dodone, il faut s'élever à une vue synthétique, qui garde tous les mots dans le champ de vision, à une distance suffisante pour qu'aucun ne soit caché. L'analyse factorielle (ou multidimensionnelle) est l'outil mathématique qui permet ce survol. Nous en donnerons trois illustrations.

a – La première est relative au corpus des **écrivains**. Une première expérience s'étant révélée décevante (parce que la distribution trop irrégulière et trop spécialisée de quelques mots avait trop d'influence), l'analyse a été refaite en excluant 8 éléments perturbateurs. Consignée dans le graphique 18, elle accorde un quadrant à chacun des trois écrivains que le graphique 10 avait reconnus comme seuls excédentaires. Voltaire a établi son tribunal sur le quadrant gauche inférieur. C'est là qu'il juge et compare les religions, opposant *Mahomet* à *Abraham*, les *catholiques* aux *musulmans*, *Jérusalem* à *Babyone*, les *moines* aux *bonzes*, les *jésuites* aux *lévites* et les *mages* aux *athées*. Chateaubriand officie dans le quadrant supérieur gauche dans une atmosphère *sacrée, divine, sainte* et *céleste*. Pour célébrer dignement le *culte*¹⁶, il a convoqué à l'*église* (ou au *temple*) le *clergé*, les *pasteurs*, les *prêtres*, les *diacres* et jusqu'aux *anges*. C'est la procession de la fête-Dieu. Le troisième

16. On notera que ce culte est une relation personnelle et qu'il s'adresse à *Jésus-Christ*.

quadrant est moins nettement l'apanage d'un homme. Hugo le partage avec les écrivains de son temps : Flaubert et Balzac. Ici la religion se fait plus trouble, le principe du mal combattant celui du bien. On voit *Satan* s'opposer à *Jéhovah*, les *démons* aux *archanges*, *Babel* aux *cathédrales*. Le quatrième quadrant enfin (en bas à droite) est attribué indistinctement à tous les autres auteurs, de Marivaux à Sartre, et il n'y a pas grand chose à dire de cet héritage indivis¹⁷.

Graphique 18. Analyse factorielle du vocabulaire religieux des écrivains (facteurs 1 et 2)

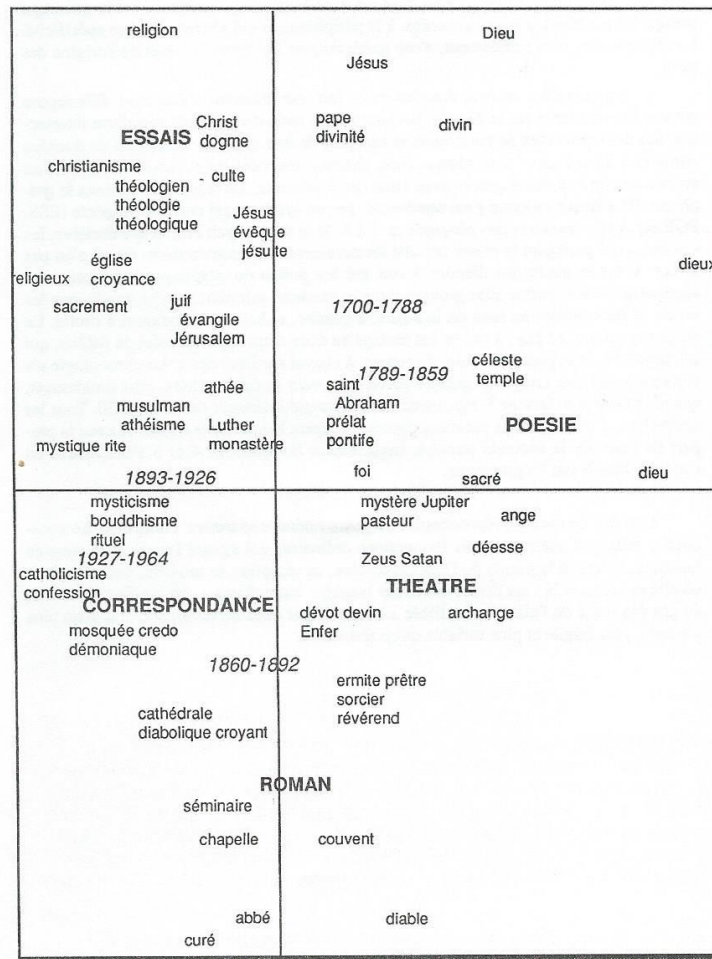


b – La seconde analyse s'appuie à la fois sur les tableaux des genres et des époques. Chaque mot est doté de dix informations, dont cinq sont relatives au **genre** et cinq à la **chronologie**. Les dix variables se livrent dans le graphique 19 à un jeu complexe d'alliances et d'oppositions.

17. Une autre analyse a été réalisée avec les mêmes données qui ont produit le graphique 18, en choisissant un principe différent de pondération. Si Voltaire, Chateaubriand et Hugo y gardent leurs distances respectives, Zola se détache dans le quatrième quadrant, où il attire des *prêtres*, des *curés* et des *abbés*, pour leur donner un rôle dans les *Rougon-Macquart*.

Graphique 19. Analyse factorielle du vocabulaire religieux
Les genres et les périodes (facteurs 1 et 2)

Graphique 19. Analyse factorielle du vocabulaire religieux
des genres et des périodes (facteurs 1 et 2)



On distingue d'abord sur la droite l'isolement de la poésie, qui ne veut pas se mêler aux genres prosaïques, surtout à la correspondance et aux essais. La poésie n'accepte autour d'elle que les noms nobles et classiques comme *dieu*, *dieux*, *déesse*, *ange*, *temple* et les adjectifs en grande livrée : *divin*, *céleste*, *sacré*. C'est de ce côté que s'orientent les deux premières tranches, parce qu'elles sont les moins éloignées de l'âge

classique (époque 1700-1788) ou les plus favorables à la poésie (période 1789-1859). La seconde ligne de partage oppose le haut et le bas, c'est-à-dire les essais et les romans. Les premiers mettent en oeuvre le vocabulaire technique du théologien : *théologie, théologien, théologique, catholique, religion, religieux, christianisme, croyance, dogme, évangile, divinité, athée, athéisme, culte, rite, sacrement, clergé, pape, église*. Les seconds, au bas du graphique, mettent en action des personnages et des lieux, qui ont un rapport avec le monde ecclésiastique : *curé, abbé, révérend, ermite, prêtre, sorcier, couvent, chapelle, cathédrale, séminaire*. Ce sont précisément les éléments que le graphique 16, consacrée au naturalisme, avait déjà révélés.

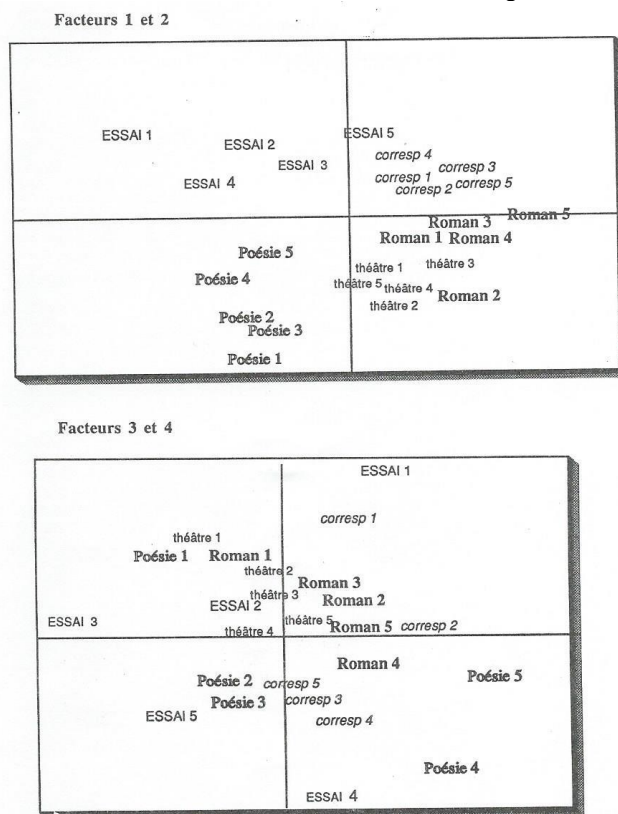
La chronologie détermine son choix une fois que le genre a dessiné la carte et imposé sa loi. On a déjà noté que l'époque romantique s'orientait vers la poésie. On constate que la seconde moitié du XIX^e siècle se range du côté du roman et que les deux tranches les plus récentes vont à gauche, là où sont les essais et la correspondance. Il n'y a donc pas pleine indépendance des genres et de la chronologie. Reste à savoir lequel des deux facteurs domine l'autre. La réponse est dans le graphique même : c'est le genre qui occupe les positions les plus avancées, à la périphérie, et qui affirme ainsi sa spécificité. La chronologie, plus timidement, n'ose guère risquer ses éléments loin de l'origine des axes.

c – Une dernière analyse donnera de ce fait une illustration éclatante. Elle repose comme la précédente sur le genre et la chronologie, mais il s'agit cette fois d'une intersection des deux principes de répartition, et non plus de leur réunion. Le tableau de données comprend 25 colonnes pour chaque mot, chacune recevant les occurrences du mot au croisement de l'un des 5 genres avec l'une des 5 périodes. Le résultat se lit dans le graphique 20. Chaque colonne y est représentée par un symbole qui combine le genre (ESS, POE, etc.) et la tranche chronologique de 1 à 5. Si le temps avait l'influence décisive, les symboles qui partagent la même période formeraient des agglomérations. Or tel n'est pas le cas. C'est le genre qui dispose à son gré les points du graphique : tous ceux qui appartiennent à la poésie sont groupés dans le quadrant inférieur gauche, tandis que les essais se rassemblent au haut de la figure, à gauche, et la correspondance, à droite.

Le dernier quadrant, en bas, à droite, est partagé en deux clans séparés : celui du théâtre, qui se rapproche de la poésie, et celui du roman, à cheval sur l'axe des x. La chronologie n'a son mot à dire que

lorsque le genre a parlé, et sa voix se fait entendre, plus timidement, quand intervient le facteur 3, représenté dans la moitié inférieure du graphique 20. Tous les symboles qui relèvent de la première tranche occupent le haut de la figure, et aussi la plupart de ceux de la seconde tranche. Inversement les tranches 4 et 5 s'établissent en majorité dans le bas de la figure.

Graphique 20. Analyse factorielle du vocabulaire religieux
Au croisement des genres et des époques



Une fois de plus la prépondérance du genre littéraire se trouve confirmée. Le vocabulaire religieux obéit donc aux fluctuations ordinaires qui agitent l'esprit et le langage humains. Même si la parole de Dieu est révélée, sa réception se modifie, selon l'endroit où elle s'exerce, et la voix divine arrive par bouffées intermittentes aux oreilles humaines, au gré des vents de l'histoire. La Bible a beau se figer dans un testament, rien n'est plus contesté, plus fragile et plus variable qu'un testament.