



HAL
open science

Les prénoms et les noms de la Bible. Étude quantitative à travers quatre siècles de littérature

Étienne Brunet

► **To cite this version:**

Étienne Brunet. Les prénoms et les noms de la Bible. Étude quantitative à travers quatre siècles de littérature. Onomastique biblique, Yannick Le Boulicaut, Nov 1988, Angers, France. pp.119-147. hal-01544825

HAL Id: hal-01544825

<https://hal.science/hal-01544825>

Submitted on 22 Jun 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.

Les prénoms et les noms de la Bible.

Étude quantitative

à travers quatre siècles de littérature.

Etienne BRUNET

Université de Nice

SUMMARY. The 177 million lexical units provided by the National Institute of French Language allow the lexicographer to trace the evolutions and changes concerning the use of first names in French literature since the XVIIth century. Thanks to linguistic statistics - called "lexicometry" - it is possible to spot quantifiable factors, to follow with acute precision how novelists and poets, playwrights and polemicists have used the 300 names of saints and martyrs scattered through our various calendars. Since the computer arbitrarily notes the number of times a name is used per century or per novel, careful distributional criteria have to be applied in the research, so that the main character of a particularly long novel does not falsify the results. Thus the 1,000 occurrences of Marius in Victor HUGO's *Les Misérables* count only as one. In addition, this statistical approach must be conducted cautiously, as the computer cannot easily distinguish the name of a saint when used as a Christian name from the same one used as a common noun or as a name of a place (i.e. a church, a school, a river or a town...). This study offers a detailed description of people's tastes in Christian names as reflected in literature.

J'ai de la joie à retrouver un cadre où j'ai connu mes premiers émois d'étudiant et mes premiers effrois d'enseignant. Mais cela se passait en des temps très anciens et j'aurai de la peine à ménager la transition avec le dernier cours que j'ai assuré en ces lieux il y a vingt-cinq ans. D'autant qu'en m'éloignant dans l'espace et dans le temps, je me suis quelque peu égaré dans une discipline bizarre qui mêle les chiffres aux lettres et qu'on appelle la "lexicométrie" ou

"statistique linguistique". Cette science s'ingénie à repérer dans les textes des éléments quantifiables. Mais ce qu'on compte, est-ce bien cela qui compte ? La réponse est douteuse. Elle l'est plus encore si l'on passe du profane au sacré et si la machine à lire et à compter s'exerce sur la Bible. On risque alors d'attirer sur soi l'ironie dont ZOLA, pourtant piètre défenseur de la religion et grand amateur de méthodes scientifiques, accable la manie comptable d'un de ses personnages « Vous n'imaginerez jamais à quoi Mouret passe son temps (...) Eh bien! il compte les s qui se trouvent dans la Bible. Il a craint de s'être trompé, et il a recommencé trois fois son calcul (...) Ma foi ! vous aviez raison, il est fêlé du haut en bas, ce farceur-là »⁽¹⁾.

Mais les machines - et les "farceurs" - sont dans le temple même. Les abbayes les accueillent dans des laboratoires, comme à Maredsous. au Centre « Informatique et Bible », dont un représentant se retrouve parmi nous. Et à ces violeurs de secrets la Bible n'oppose qu'un sourire énigmatique, étant moins vulnérable à l'appareillage scientifique que ne fut la Joconde aux rayons infrarouges ou le Saint Suaire de Turin au carbone 14. Mais par prudence notre enquête ne portera pas sur la Bible, mais seulement sur les traces que le texte sacré a laissées dans les lettres françaises. Encore, pour se conformer à l'objet onomastique de ce colloque, allons-nous nous contenter des noms et des prénoms dont l'origine est religieuse, sinon biblique. Et en cela nous poursuivons une enquête, que nous avons conduite parmi les mots communs du vocabulaire religieux et dont les résultats ont été communiqués à Jérusalem en juin 1988⁽²⁾.

I - Les données

A vrai dire, nous ne savons trop si les noms propres que nous allons passer en revue sont bien propres à la Bible. Il s'agit d'un héritage ancien, transmis de l'Ancien Testament, et plus souvent du Nouveau Testament et du martyrologe romain. Il est possible qu'aux époques où la croyance était vive le prénom donné à l'enfant et inscrit sur les registres de la paroisse ait pu faire référence explicite et très consciente à un saint du calendrier. Mais aujourd'hui il ne s'agit guère que d'une disposition laïque de l'état civil, au reste soumise aux variations de la mode. Et l'on pourrait justifier plus facilement une enquête sociologique menée dans les mairies ou dans le vaste champ de l'annuaire du téléphone⁽³⁾. Avec les prénoms laïcisés, souvent étrangers,

régionaux, exotiques, qui ont la faveur des parents modernes, on semble assister à la dissipation de l'héritage. Et même lorsque le choix se porte sur un nom du martyrologe, le saint invoqué n'est guère évoqué, il ne sert que de prête-nom, au point que la fête du saint patron devient la fête « des » *Pierre* et « des » *Paul*, dont un seul - le patronyme - est excepté.

Ainsi notre objet d'étude manque de pureté, puisqu'il confond le profane et le sacré, les premiers siècles de l'Eglise et le monde contemporain. Il confond aussi les êtres réels et ceux de la fiction, les personnages de roman et les figures historiques, le vrai saint Antoine (ou du moins celui que FLAUBERT imagine comme tel) et les multiples héros de l'histoire romaine ou du théâtre moderne que désignent les mêmes syllabes, Il mêle en outre les noms et les prénoms, inextricablement mêlés : comment savoir, hors contexte, si **David** est ou non un patronyme ? Il mêle aussi l'Ancien et le Nouveau Testaments : les premiers martyrs ne disposaient pas d'un martyrologe et leurs noms sont empruntés à toutes les traditions, y compris celle des Juifs. Et le livre liturgique, qui officialise la lise des saints, suit les aléas de l'histoire, acceptant **David** mais refusant **Suzanne**. On doit regretter enfin le mélange des personnes et des lieux. L'expression **Saint Louis** (ou **Sainte Marie** ou **Saint Laurent**) peut désigner une ville, un fleuve, une montagne, un quartier, une église, un cimetière ou une institution. Quand **Laurent** est isolé, comment décider qu'il s'agit d'un fleuve canadien plutôt que d'un prince florentin ? Et, bien sûr, en étendant l'enquête à quatre siècles de littérature, on brouille sinon les époques, du moins les auteurs et les genres. Quelque lueur peut-elle sortir d'un puits si obscur et si trouble ?

Ne désespérons pas trop vite : la statistique intervient généralement dans les situations de ce genre, où le brouillard enveloppe les faits, alors que, par temps clair, le déterminisme suffit pour démêler les causes et les effets. La méthode statistique est un instrument de première urgence qui aide à mettre de l'ordre quand règne le chaos, et à forger quelques présomptions quand manquent les preuves. Mais voyons-la à l'oeuvre.

Il lui faut d'abord des données, en grand nombre. A défaut de qualité, les méthodes dites quantitatives spéculent sur la quantité. Celles que l'Institut National de la langue française a amassées forment une montagne de 177 millions d'unités linguistiques⁽⁴⁾, la plus grosse qu'on ait élevée au monde dans le domaine des lettres. Encore faut-il pouvoir

exploiter cette masse de 3000 textes. Le gisement longtemps enfoui est depuis quelques années à l'air libre, offert au monde scientifique. Constitué en base de données⁽⁵⁾, ce trésor peut être exploré à distance, par la voie télématique, pour n'importe quel mot, n'importe quelle expression, dans n'importe quel texte (ou auteur ou genre ou siècle). La réponse, qui peut être un contexte, un index, une cooccurrence, est donnée immédiatement.

Comme nous voulions une vue panoramique sur l'onomastique religieuse, nous avons choisi la totalité des 3000 textes qui s'échelonnent de 1600 à 1964 et que nous avons répartis en sept tranches, correspondant grossièrement aux demi-siècles⁽⁶⁾. Le corpus étant ainsi fixé, sans esprit parcimonieux, restait à établir la liste des noms, sans exclusive. Dans un premier temps nous avons interrogé le calendrier, pour y relever les noms des saints. Et pour éviter que les fêtes variables ne fassent tort à quelques-uns, nous avons utilisé deux calendriers successifs, ce qui par exemple a sauvé Saint Joseph, que trop de modestie avait fait s'effacer en mars 1988, le jour des Rameaux⁽⁷⁾. C'est donc 330 saints ou martyrs qui ont d'abord été livrés en pâture au monstre électronique. Puis, dans une deuxième approche, conscient que le catalogue officiel de Grégoire XIII ne laissait aucune place à la plupart des noms de l'Ancien Testament, une seconde liste, biblique, a été établie, qui va de **Abdias** à **Zacharie**, et que nous examinerons plus loin.

Voir page suivante :

Tableau 1, Liste alphabétique des noms du calendrier

II - Les occurrences

Le tableau 1 délivre un extrait de la liste fournie par l'ordinateur. Les sept colonnes correspondent aux effectifs relevés dans les sept tranches qu'on a distinguées dans le temps : 1600 à 1700, 1701 à 1749, 1750 à 1788, 1789 à 1848, 1849 à 1892, 1893 à 1932, 1933 à 1964. Ces divisions reprennent certains jalons qu'avait adoptés *le Trésor de la langue française*⁽⁸⁾, mais ne s'y confondent pas absolument : la charnière entre XIXe et XXe siècles a été fixée en 1892 et non en 1880, le souci premier étant d'équilibrer les tranches⁽⁹⁾. Les noms du calendrier

Tableau 1. Liste alphabétique des noms du calendrier (extrait)

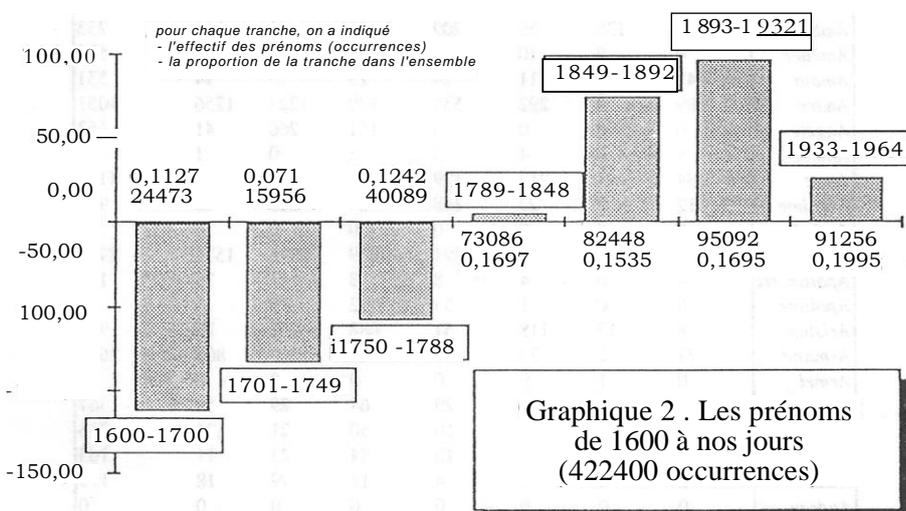
	1600	1700	1750	1789	1849	1893	1933	total
<i>Abel</i>	26	34	24	147	57	113	150	551
<i>Abraham</i>	211	324	273	97	66	142	125	1238
<i>Achille</i>	260	384	226	224	321	136	135	1686
<i>Adèle</i>	0	0	657	1368	435	130	19	2609
<i>Adeline</i>	2	0	12	196	19	2	9	240
<i>Adelphe</i>	0	0	0	0	0	0	0	0
<i>Agathe</i>	10	532	249	197	19	142	134	1283
<i>Agnès</i>	36	83	193	234	132	112	688	1478
<i>Aimé</i>	0	0	3	88	11	452	15	569
<i>Aimée</i>	6	0	0	15	36	193	4	254
<i>Alain</i>	33	145	15	43	16	73	670	995
<i>Alban</i>	7	1	199	95	1	806	27	1136
<i>Albert</i>	234	196	79	1015	359	1116	647	3646
<i>Alexandre</i>	589	372	523	852	411	376	410	3533
<i>Alexis</i>	1010	15	379	44	128	119	646	2341
<i>Alice</i>	0	0	1	6	150	81	92	330
<i>Alida</i>	0	0	0	0	0	0	3	3
<i>Alix</i>	15	0	10	50	30	30	108	243
<i>Alphonse</i>	640	249	54	260	358	161	188	1910
<i>Amandine</i>	0	0	0	0	0	0	2	2
<i>Ambroise</i>	138	135	59	209	44	42	106	733
<i>Amédée</i>	0	9	10	68	30	235	242	594
<i>Amour</i>	442	2	11	24	23	15	14	531
<i>André</i>	189	79	292	533	479	1223	1256	4051
<i>Angèle</i>	0	0	0	5	151	266	41	463
<i>Anicet</i>	1	0	4	3	5	0	1	14
<i>Anne</i>	134	94	217	179	257	2109	1154	4144
<i>Anselme</i>	512	5	27	148	24	202	28	946
<i>Anthelme</i>	0	0	0	0	0	0	33	33
<i>Antoine</i>	666	353	397	991	1909	2846	1574	8736
<i>Apollinaire</i>	4	0	4	8	3	11	75	105
<i>Apolline</i>	0	0	1	53	2	13	1	70
<i>Aristide</i>	8	12	118	51	688	57	25	959
<i>Armand</i>	77	2	22	490	762	505	805	2663
<i>Arnel</i>	0	1	0	0	0	0	0	1
<i>Arnaud</i>	29	13	85	29	64	89	58	367
<i>Arsène</i>	0	1	9	10	50	21	124	215
<i>Aubin</i>	60	2	13	10	44	23	11	163
<i>Aude</i>	2	0	4	4	12	79	18	119
<i>Audrey</i>	0	0	0	0	0	0	0	0
<i>Augustin</i>	965	243	131	356	281	698	2034	4708
<i>Aymard</i>	2	1	0	1	4	1	3	12
<i>Baptiste</i>	68	59	37	127	173	106	94	664
<i>Barbara</i>	0	0	0	8	1	4	18	31
<i>Barnabé</i>	5	0	5	4	2	16	16	48
<i>Barnard</i>	0	0	0	0	0	1	8	9
<i>Barthélémy</i>	0	0	0	1	3	3	32	39
<i>Basile</i>	95	29	23	41	25	63	29	305
<i>Baudouin</i>	33	8	34	16	120	5	346	562
<i>Béatrice</i>	0	0	0	18	12	28	13	71
<i>Bénédicte</i>	4	0	7	0	0	1	0	12
<i>Benjamin</i>	11	6	52	162	59	115	123	528

constituent une masse de près d'un demi-million d'occurrences, dont voici la distribution :

tranche	1	2	3	4	5	6	7	total
année de départ	1600	1701	1750	1789	1849	1893	1933	
effectif	24473	15956	40089	73086	82448	95092	91256	422400

Cela représente le cinquième de l'effectif total des noms propres⁽¹⁰⁾ ou encore une unité linguistique⁽¹¹⁾ sur 400, c'est-à-dire une apparition par page en moyenne.

Quand la machine a digéré dans l'ordre chronologique ces 400.000 pages, on peut lui demander si elle a noté une progression ou une régression dans l'usage littéraire des noms du calendrier. La réponse est dans le graphique 2 :



Il s'agit d'un histogramme dont les éléments sont proportionnels aux effectifs observés, compte tenu de l'étendue de chaque sous-ensemble⁽¹²⁾. La ligne médiane, qui correspond à une distribution régulière ou homogène, fait le partage entre les excédents (dans la zone positive et supérieure) et les déficits dans la zone négative, au bas du graphique. La tendance est nette : les excédents sont relevés dans les

deux derniers siècles, les déficits dans les deux premiers. A quoi tient cette progression ? La machine nous abandonne lâchement dès qu'il faut interpréter. On peut songer d'abord à quelque artefact : comme l'orthographe reste flottante jusqu'au XIXe siècle, certains noms qui comportent des lettres accentuées sont peut-être sous-représentés, parce qu'on n'a pas tenu compte des variantes de l'écriture. En réalité cette distorsion est de faible ampleur et l'évolution se maintient quand on limite l'enquête aux noms qui ne sont pas susceptibles de variations orthographiques. Si l'on admet que les prénoms servent surtout à désigner les personnages de roman, on aurait là peut-être la conséquence des progrès du roman dans notre littérature⁽¹³⁾. Mais il peut y avoir des raisons indépendantes des genres littéraires : en libérant l'individu, et en donnant valeur à la couleur locale, à l'unicité et à la spécificité des personnes et des lieux, le XIXe siècle s'est éloigné de la généralité chère aux siècles classiques et a favorisé l'expansion des noms propres⁽¹⁴⁾.

On vient d'ajouter les lignes les unes aux autres pour considérer le total marginal de chaque colonne, c'est-à-dire de chaque période.

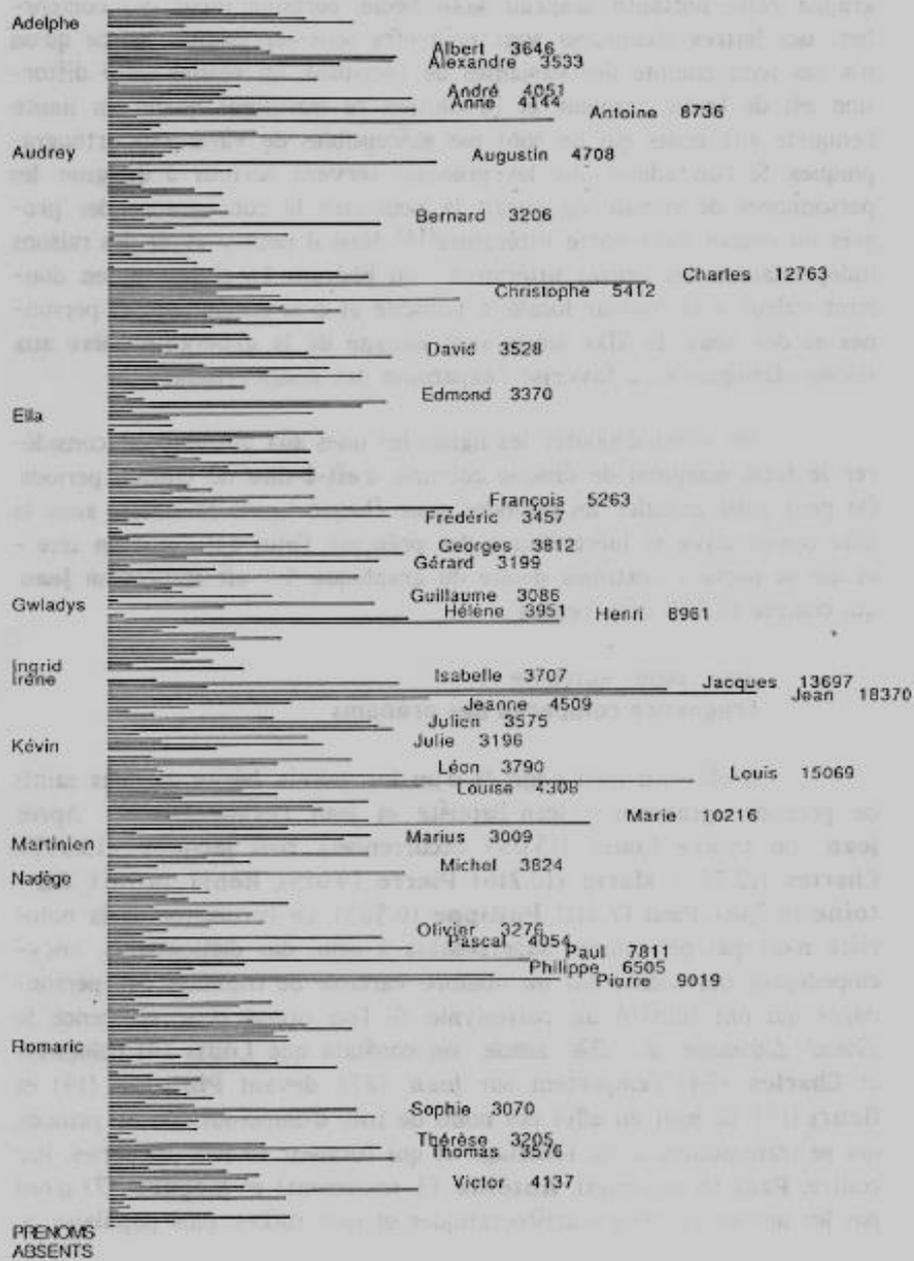
On peut aussi cumuler les périodes pour chaque ligne, et établir ainsi la liste comparative et hiérarchique des prénoms. Celui qui vient en tête - et qui se porte à l'extrême droite du graphique 3 - est le prénom *Jean* qui compte 18.370 occurrences.

Voir page suivante :

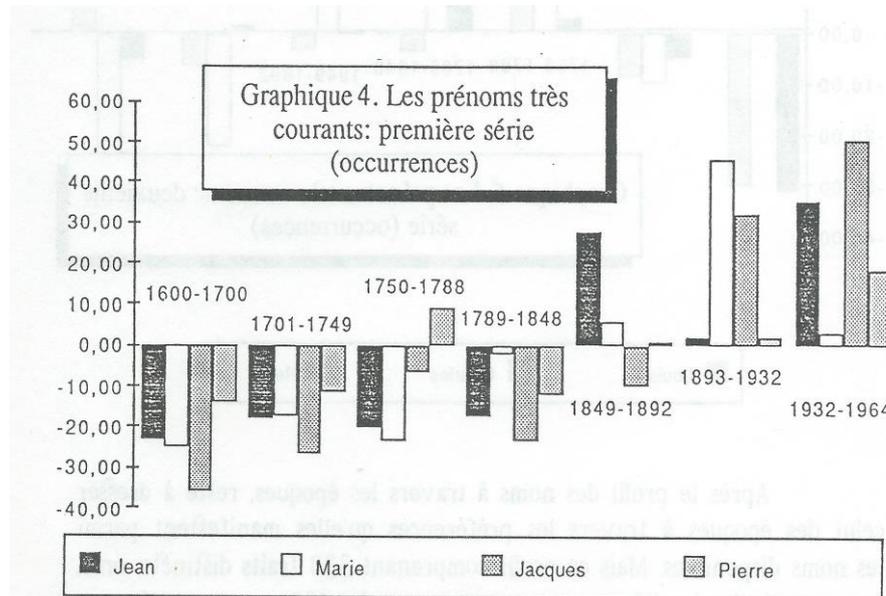
Fréquence comparée des prénoms

Il est vrai qu'il s'agit là d'un lot indivis, hérité de deux saints de première grandeur : Jean-Baptiste et Jean l'évangéliste⁽¹⁵⁾. Après *Jean*, on trouve *Louis* (15.069 occurrences), puis *Jacques* (13.697), *Charles* (12.763), *Marie* (10.216), *Pierre* (9.019), *Henri* (8.961), *Antoine* (8.736), *Paul* (7.811), *Philippe* (6.505). Ce baromètre de la notoriété n'est pas pleinement superposable à celui des dictionnaires encyclopédiques, qui consacrent un nombre variable de colonnes aux personnages qui ont illustré un patronyme. Si l'on prend pour référence le *grand Larousse du 20e siècle*, on constate que *Louis* (41 colonnes) et *Charles* (34) l'emportent sur *Jean* (27), devant *Philippe* (19) et *Henri* (19). Ce sont en effet des noms de rois, d'empereurs ou de princes, qui se transmettent avec l'héritage et qui forment de longues séries. Par contre, *Paul* (6 colonnes), *Antoine* (3 seulement) et *Jacques* (7) n'ont pas les mêmes privilèges aristocratiques et leur succès, plus populaire, a

Graphique 3. Fréquence comparée des prénoms



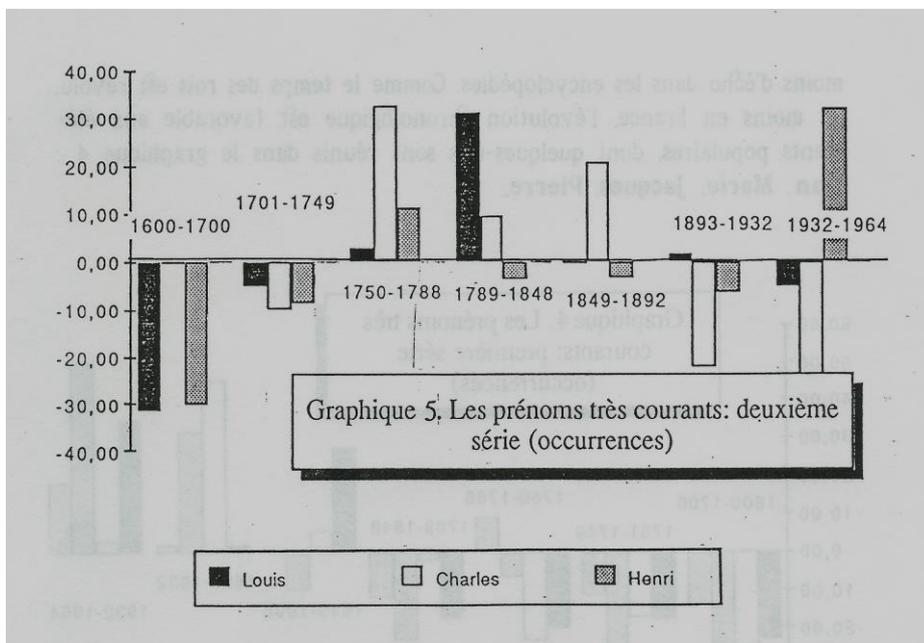
moins d'écho dans les encyclopédies. Comme le temps des rois est révolu, au moins en France, l'évolution chronologique est favorable aux éléments populaires, dont quelques-uns sont réunis dans le graphique 4 *Jean, Marie, Jacques, Pierre,*



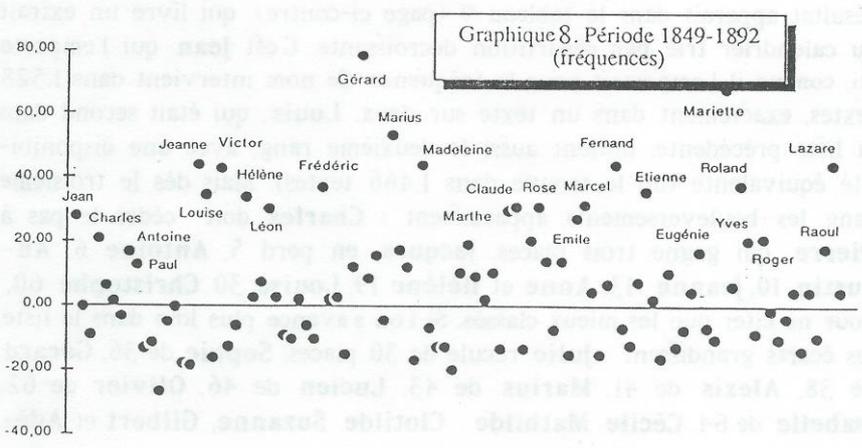
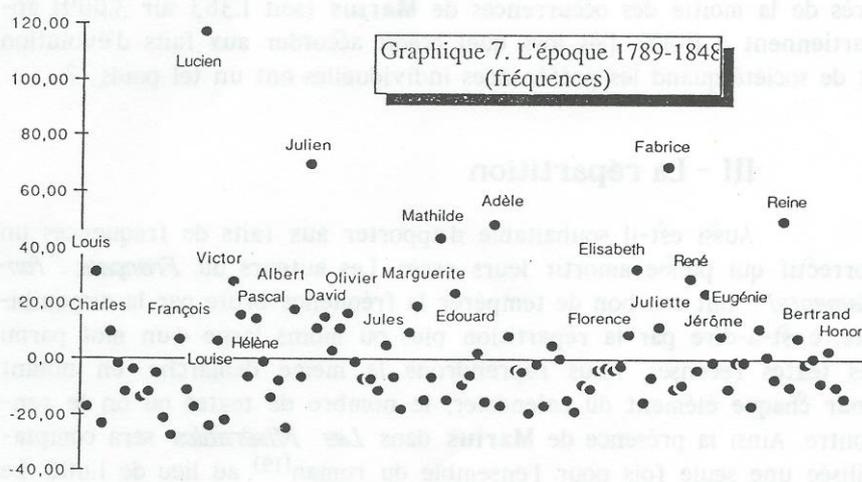
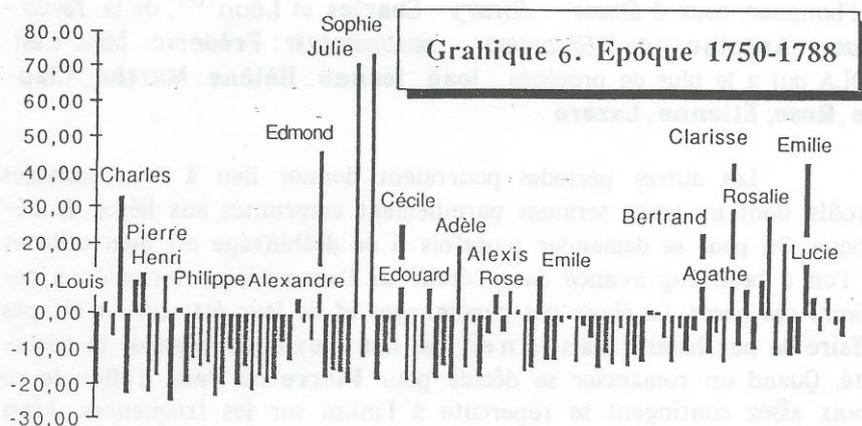
plutôt qu'à ceux qui ont été associés aux dynasties éteintes : Louis, Charles , Henri⁽¹⁶⁾.

Voir graphique 5 page suivante :
Les prénoms très courants : 2e série (occurrences)

On s'est intéressé aux dix éléments les mieux représentés, qui représentent le quart du total des occurrences. On pourrait aussi relever les derniers éléments, notamment ceux qu'aucun auteur n'a daigné citer et qui ne comptent aucun emploi. Ils sont dix également : **Adelphe Audrey, Ella, Gwladys, Ingrid, Irène, Kévin, Martinien , Nadège et Romaric**. Mais ici les commentaires doivent rester prudents, une part devant être laissée au hasard de l'échantillonnage des textes, si étendu soit-il. En choisissant d'autres textes ou d'autres limites chronologiques, les lacunes seraient probablement apparues à d'autres endroits du calendrier.



Après le profil des noms à travers les époques, reste à dresser celui des époques à travers les préférences qu'elles manifestent parmi les noms disponibles. Mais ce profil comprenant 330 traits distincts, on a cru pouvoir le simplifier en ne retenant que les 100 premiers éléments dans l'ordre de fréquence. Les graphiques 6, 7 et 8 en proposent trois qui correspondent aux périodes centrales de notre corpus. L'époque 1750-1788 qui est celle de ROUSSEAU et de DIDEROT met en tête le prénom de *Sophie* qui est commun à ces deux auteurs, Sophie est non seulement la compagne fictive de Emile mais aussi la correspondante réelle de DIDEROT : Sophie Volland. C'est encore la protection de ROUSSEAU qui vaut à *Julie* sa promotion (dans la *Nouvelle Héloïse*), comme aussi à Emile, dans l'ouvrage du même nom. On reconnaît aussi la *Cécile* des *Liaisons dangereuses* mais à qui faut-il attribuer le succès de *Clarisse*, sinon, indirectement, au succès de l'héroïne de RICHARDSON ? Le graphique 6 (page ci-contre qui va de la Révolution de 1789 à celle de 1848, donne le prix à *Lucien*, avec les suffrages conjugués de BALZAC (Lucien de Rubempré) et de STENDHAL (Lucien Leuwen). STENDHAL est aussi le protecteur de *Julien* et de *Mathilde*, dans *Le rouge et le noir*; et de *Fabrice* (*La Chartreuse de Parme*). HUGO se devine derrière *Victor* et *Adèle*, et CHATEAUBRIAND derrière *René*. Hugo est encore reconnaissable dans l'époque qui suit et dont rend compte le graphique 8 : *Victor* reste au premier plan en compagnie de *Marius* ⁽¹⁷⁾, tandis que l'infidèle *Adèle* s'éloigne. Les personnages de FLAUBERT sont à l'honneur, ceux d'*Emma*



Bovary Charles et Léon⁽¹⁸⁾; de la *Tentation* : Antoine ; de l'*Education sentimentale* : Frédéric. Mais c'est ZOLA qui a le plus de protégés Jean , Jeanne , Hélène , Marthe , Claude , Rose, Etienne , Lazare .

Les autres périodes pourraient donner lieu à de semblables profils, dont les traits seraient pareillement empruntés aux héros de l'époque. On peut se demander toutefois si ce déchiffrement est bien utile et si l'on a beaucoup avancé dans l'étude de l'onomastique lorsqu'on a reconnu ces héros. Le choix des personnages et de leur état civil n'est pas affaire de pur hasard, mais ce n'est pas non plus le domaine de la nécessité. Quand un romancier se décide pour Pierre ou Paul, l'effet de ce choix assez contingent se répercute à l'infini sur les fréquences. Ainsi près de la moitié des occurrences de *Marius* (soit 1.363 sur 3.009) appartiennent à HUGO, Dès lors quel crédit accorder aux faits d'évolution et de société quand les préférences individuelles ont un tel poids ?

III - La répartition

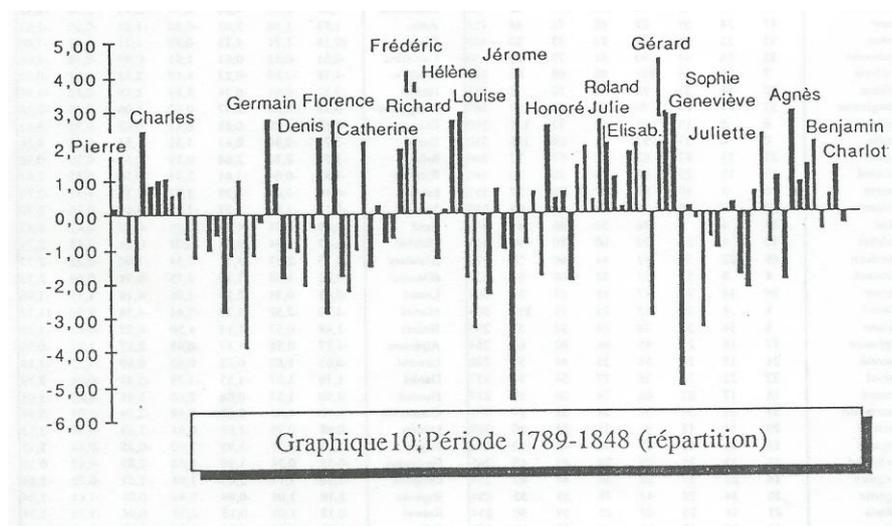
Aussi est-il souhaitable d'apporter aux faits de fréquences un correctif qui puisse amortir leurs excès. Les auteurs du *Français fondamental* ont cru bon de tempérer la fréquence brute par la disponibilité, c'est-à-dire par la répartition plus ou moins large d'un mot parmi les textes recensés. Nous reprendrons la même démarche, en notant pour chaque élément du calendrier le nombre de textes où on le rencontre. Ainsi la présence de *Marius* dans *Les Misérables* sera comptabilisée une seule fois pour l'ensemble du roman⁽¹⁹⁾ au lieu de 1.000. Le résultat apparaît dans le tableau 9 (page ci-contre), qui livre un extrait du calendrier trié par répartition décroissante. C'est *Jean* qui l'emporte ici, comme il l'emportait pour la fréquence. Ce nom intervient dans 1.528 textes, exactement dans un texte sur deux. *Louis*, qui était second dans la liste précédente, obtient aussi le deuxième rang, avec une disponibilité équivalente (on le trouve dans 1.466 textes). Mais dès le troisième rang les bouleversements apparaissent : *Charles* doit céder le pas à *Pierre*, qui gagne trois places. *Jacques* en perd 5, *Antoine* 6, *Augustin* 10, *Jeanne* 13, *Anne* et *Hélène* 19. *Louise* 30, *Christophe* 60, pour ne citer que les mieux classés. Si l'on s'avance plus loin dans la liste, les écarts grandissent : *Julie* recule de 30 places. *Sophie* de 36, *Gérard* de 38. *Alexis* de 41. *Marius* de 43, *Lucien* de 46, *Olivier* de 62. *Isabelle* de 64. *Cécile*, *Mathilde*, *Clotilde*, *Suzanne*, *Gilbert* et

Tableau 9. La répartition des prénoms. Extrait.
(par ordre décroissant)

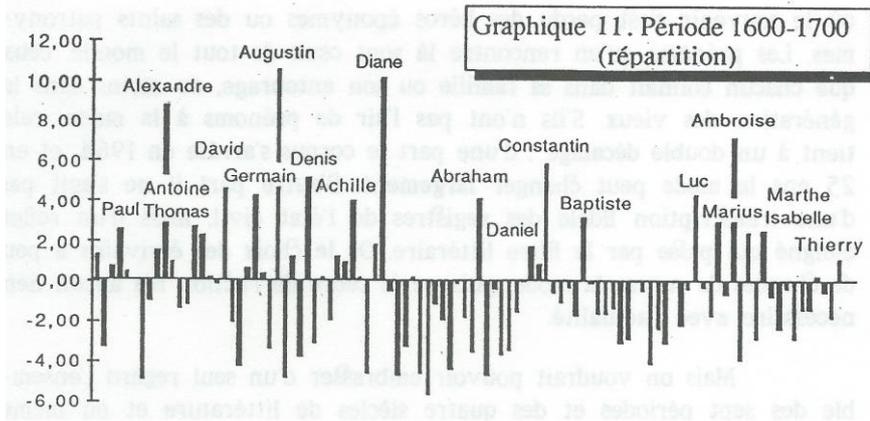
Nombre de textes où figure le prénom								Transformation en écart réduit								
	1600	1701	1750	1789	1849	1893	1933 total		1600	1701	1750	1789	1849	1893	1931	
Jean	1600	1701	1750	1789	1849	1893	1933 total	Jean	1,35	-0,59	0,28	0,22	-1,70	-0,09	0,61	
Louis	146	60	150	276	244	287	365	1528	Louis	-3,35	1,03	3,30	1,73	-0,30	-0,58	-1,34
Pièrre	116	58	140	214	206	217	308	1259	Pièrre	0,79	0,66	1,82	-0,78	-1,18	-1,48	1,03
Paul	127	54	89	191	205	229	293	1188	Paul	2,59	0,54	-2,47	-1,59	-0,33	0,36	1,17
Charles	103	51	143	236	185	182	242	1142	Charles	0,52	0,39	3,35	2,49	-1,26	-2,54	-1,63
Marie	96	38	88	208	204	239	231	1104	Mono	0,13	-1,30	-1,84	0,87	0,74	2,36	-1,82
Jlenri	43	37	131	193	178	179	251	1012	Henri	-4,93	-0,91	3,61	1,02	-0,03	-0,96	1,18
Jacques	77	44	101	191	164	170	251	998	Jacques	-0,98	0,28	0,55	1,07	-0,99	-1,48	1,43
François	91	44	96	169	142	167	199	908	François	1,54	0,92	0,99	0,61	-1,57	-0,37	-0,94
Alexandre	149	63	128	165	128	121	127	881	Alexandre	8,82	4,30	4,96	0,69	-2,41	-3,90	-6,20
Philippe	78	54	108	138	138	146	156	818	Philippe	0,97	3,36	3,49	-0,73	-0,57	-0,75	-2,82
Michel	55	22	70	124	147	159	192	769	Michel	-1,42	-1,89	-0,47	-1,24	108	1,28	1,14
Joseph	55	31	61	134	129	167	170	747	Joseph	-1,20	-0,11	-1,33	0,07	-0,26	2,43	-0,31
Antoine	80	53	82	114	134	108	128	699	Antoine	2,70	4,40	1,91	-1,06	1,07	-2,31	-3,08
Thomas	71	27	72	108	86	99	174	637	Thomas	2,30	0,01	145	-0,59	-2,73	-2,15	2,44
André	54	21	40	82	73	135	206	611	André	0,22	-0,98	-2,57	-2,86	-3,68	2,04	6,13
Robert	51	15	48	97	96	129	172	608	Robert	-0,18	-2,16	-1,43	-1,22	-1,19	1,48	2,95
David	84	35	66	118	79	113	110	605	David	4,65	1,90	1,09	1,07	-2,95	-0,12	-2,94
Jules	36	34	42	67	108	118	173	578	Jules	-2,02	1,97	-1,91	-3,93	0,67	0,95	3,81
Georges	20	13	20	90	99	126	198	566	Georges	-4,29	-2,29	-4,90	-1,21	-0,08	2,06	6,62
Bernard	53	22	54	98	69	119	146	561	Iko-nard	0,73	-0,37	0,02	-0,23	-3,31	1,42	1,57
Germain	76	30	57	124	113	64	90	554	Germain	4,31	1,38	0,55	2,79	1,71	-4,40	-3,89
Guillaume	50	30	54	107	82	110	120	553	Guillaume	0,38	1,39	0,13	0,92	-1,73	0,61	-0,85
Victor	25	8	11	81	130	150	143	548	Victor	-3,36	-3,22	-6,04	-1,87	3,75	5,09	1,59
Augustin	116	35	57	88	73	76	98	543	Augustin	10,62	2,56	0,71	-1,00	-2,56	-2,90	-2,86
1 Mon	15	20	26	87	88	133	173	542	Léon	-4,84	-0,63	-3,80	-1,09	-0,85	3,37	4,79
Marc	55	28	89	78	85	85	120	540	Marc	1,32	1,10	5,42	-2,06	-1,15	-1,86	-0,56
Pascal	21	18	44	91	70	127	159	530	Pascal	-3,80	-0,96	-1,02	-0,40	-2,67	2,99	3,69
Denis	78	37	66	113	79	61	89	523	Denis	5,16	3,23	2,34	2,25	-1,51	-4,21	-3,37
Jeanne	24	10	25	66	114	144	129	512	Jeanne	-3,15	-2,56	-3,63	-2,93	2,76	5,35	1,05
Florence	43	21	56	109	84	83	85	481	Florence	0,28	0,15	1,52	2,76	-0,09	-0,91	-2,89
Albert	28	14	29	68	72	97	157	465	Alben	-1,98	-1,31	-2,46	-1,81	-1,21	1,10	5,38
Simon	47	24	36	63	88	90	107	455	Simon	1,33	1,10	-1,23	-2,23	0,96	0,49	0,14
Nicolas	45	19	40	72	81	90	105	452	Nicolas	1,04	-0,03	-0,54	-1,06	0,16	0,56	0,00
Achille	62	28	80	96	76	48	52	442	Achille	4,08	2,20	6,06	2,13	-0,24	-4,31	-5,71
Claude	39	15	34	66	67	73	146	440	Claude	0,21	-0,86	-1,34	-1,56	-1,32	-1,22	4,94
Madeleine	10	7	13	80	128	104	93	435	Madeleine	-4,68	-2,72	-4,68	0,30	6,46	2,68	-0,92
Anne	47	24	59	69	65	76	84	424	Artnc	1,83	146	3,01	-0,84	-1,24	-0,50	-1,67
Diane	95	25	66	70	84	53	30	423	Diane	10,18	1,71	4,19	-0,70	1,21	-3,33	-7,86
Catherine	33	16	44	90	81	73	83	420	Catherine	-0,53	-0,43	0,61	1,92	0,89	-0,78	-1,68
Frédéric	7	5	36	101	86	69	85	389	Frédéric	-4,78	-2,89	-0,23	4,18	2,32	-0,57	-0,65
Hélène	15	14	33	98	80	70	78	388	Helene	-3,32	-0,61	-0,74	3,81	1,55	-0,42	-1,46
Marguerite	33	17	23	70	87	70	67	367	Marguerite	0,28	0,38	-2,17	0,62	3,06	0,10	-2,26
icone	6	8	30	65	68	72	111	360	Etienne	-4,69	-1,89	-0,82	0,11	0,63	0,55	3,41
	0	6	37	53	71	68	115	350	Emile	-5,73	-2,34	0,61	-1,32	1,31	0,27	4,26
Julien	23	23	47	62	41	67	77	340	Julien	-1,20	2,32	2,64	0,19	-2,69	0,39	-0,26
Riclmdr	19	12	25	80	56	58	90	340	Richard	-1,97	-0,64	-1,41	2,74	-0,56	-0,85	1,41
Louise	6	3	20	81	97	73	57	337	Louise	-4,46	-3,05	-2,29	2,97	5,38	1,31	-2,75
Vincent	28	7	20	47	69	66	99	336	Vincent	-0,17	-1,96	-2,27	-1,85	1,40	0,36	2,70
René	19	4	8	36	54	66	145	332	Rond	-1,86	-2,74	-4,45	-3,33	-0,65	0,47	8,82
Ibèresse	10	4	21	59	60	80	98	332	Thérèse	-3,63	-2,74	-2,03	-0,03	0,21	2,43	2,71
Abraham	49	23	50	42	44	60	58	326	Abraham	4,15	2,53	3,51	-2,34	-1,96	-0,22	-2,33
é Douard	4	8	38	63	52	66	93	324	éDouard	-4,72	-1,58	1,30	0,75	-0,74	0,69	2,33
Laure	24	14	16	47	81	65	56	303	Lazare	-0,41	0,34	-2,55	-1,06	4,16	1,15	-1,96
Marcel	8	4	20	18	23	76	154	303	Marcel	-3,70	-2,52	-1,77	-5,41	-4,58	2,76	11,37
Jérôme	8	14	33	78	44	51	57	285	Jérôme	-3,48	0,57	1,13	4,20	-0,97	-0,42	-1,29
Alphonse	17	14	20	45	64	62	62	284	Alphonse	-1,57	0,58	-1,47	-0,88	2,17	1,27	-0,56
Lament	24	18	28	54	55	44	57	280	Laurent	-0,01	1,82	0,22	0,63	0,89	-1,35	-1,14
Daniel	'	22	20	38	27	54	84	277	Daniel	1,76	3,07	-1,35	-1,79	-3,44	0,26	2,79
ihonoré	28	17	27	66	74	26	39	277	honoré	0,90	1,57	0,08	2,60	3,97	-4,03	-3,61
Constant in	51	27	54	52	31	32	27	274	Constantin	5,92	4,62	5,68	0,49	-2,74	-3,04	-5,24
Lucien	20	14	18	52	31	54	80	269	Lucion	-0,68	0,79	-1,62	0,64	-2,63	0,50	2,53
Roger	18	11	16	35	41	45	100	266	Roger	-1,06	-0,08	-1,99	-2,00	-0,95	-0,81	5,55
Ferdinand	21	12	35	56	29	47	65	265	Ferdinand	-0,39	0,24	1,99	1,40	-2,85	-0,47	0,50
Grégoire	14	21	37	58	36	44	47	257	Grégoire	-1,80	3,14	2,61	1,98	-1,52	0,72	-1,88
Baptiste	36	14	20	48	45	39	52	254	Baptiste	3,18	1,01	-0,94	0,44	0,04	-7,43	-1,04
Roland	21	14	25	62	45	37	50	254	Roland	-0,18	1,01	0,13	2,73	0,04	-1,75	-1,34
Julie	14	14	43	60	68	32	19	250	Julie	-1,69	1,07	4,08	2,54	3,97	-2,45	-5,85
Clément	14	12	47	51	32	46	46	248	Clément	-1,65	0,47	5,00	1,12	-1,95	-0,13	-1,75
Bertrand	15	8	28	45	39	43	67	245	Bertrand	-1,38	-0,75	0,97	0,21	-0,70	-0,53	1,52
Christophe	7	13	30	54	40	46	52	242	Cleisticpic	-3,16	0,88	147	1,82	-0,45	0,05	-0,64

Adèle de plus de 60. Ces déclassements frappent les héros de roman que la faveur particulière et arbitraire d'un écrivain avait portés au devant de la scène. Inversement le critère de la répartition rend justice à **Paul** qui progresse de 6 rangs, à **Michel** qui gagne 12 places, à **Thomas (+14)**, **Simon (+18)**, **Daniel (+20)**, **Catherine (+20)**, **Claude (+22)**, **Jules (+24)**, **Roger (+27)**, **Jérôme (+28)**, **Armand (+29)**, **René (+30)**, **Étienne (+31)**, **Honoré et Denis (+34)**, **Grégoire (+39)**, **Germain (+40)**, **Gaston et Vincent (+42)**, **Marc (+43)**. On assiste même à des promotions éclatantes et méritées : de **Nicolas, Dominique, Raymond, Clément, Baptiste, Abel, Benjamin**, l'écart dépassant alors 50 places. Au total, sur les 100 premiers de la liste des occurrences, 82 seulement figurent parmi les 100 premiers de la nouvelle liste, à des places d'ailleurs fort éloignées.

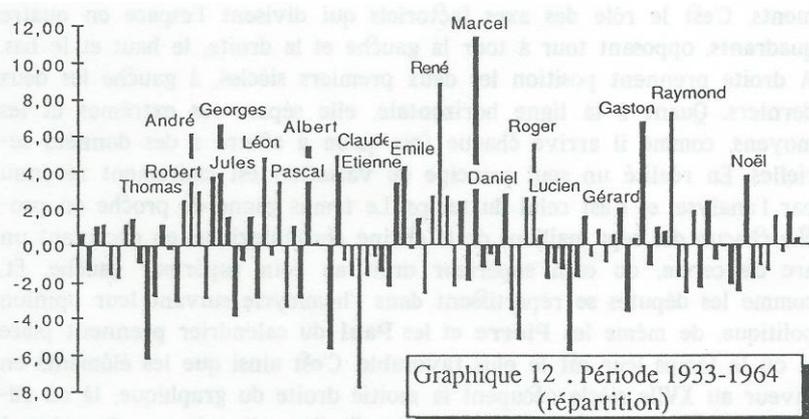
Ce changement de perspective est très sensible dans la figure 10, qui est consacrée à la période romantique et qu'il faut rapprocher de la figure 7, où la même époque est considérée. Les héros de roman **Julien, Lucien et Fabrice** ont disparu, et sur les 22 noms qui ont des excédents significatifs, 8 seulement se retrouvent dans le graphique 7 - **Charles, Juliette, Louise, Jérôme, Honoré, Hélène, Florence** ⁽²⁰⁾ et **Elisabeth**.



Pour saisir l'évolution générale, il peut être révélateur d'opposer les deux tranches extrêmes.



Le XVII^e siècle, représenté dans ce graphique 11, cite volontiers les représentants de la culture gréco-latine (Alexandre, Achille, Diane, Constantin, Marius), ou de la tradition biblique (David, Abraham, Daniel). Le siècle classique puisant souvent ses sujets dans l'antiquité, il s'agit moins ici de prénoms et de saints que de héros historiques ou légendaires.



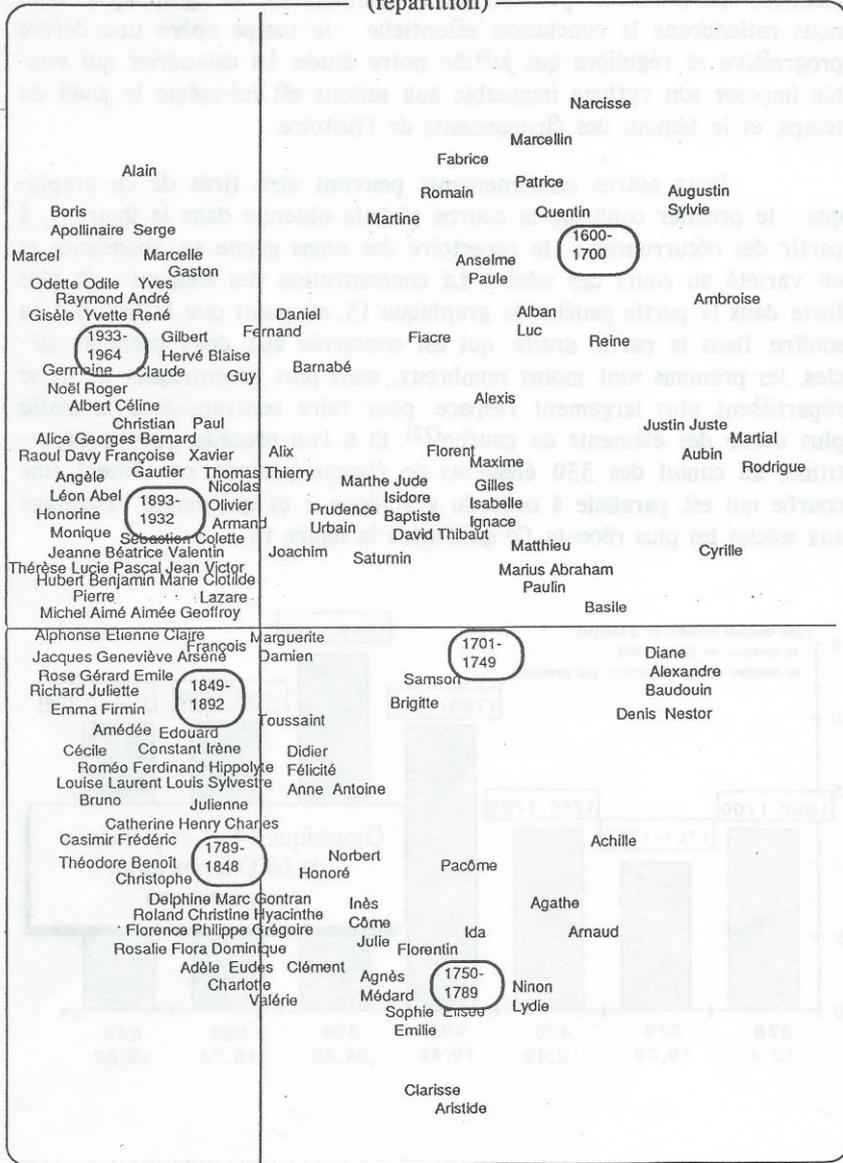
En revanche, la liste qui rend compte de la période la plus récente (graphique 12) est celle du calendrier banalisé, actualisé, laïcisé,

où le souvenir s'est perdu des héros éponymes ou des saints patronymes. Les prénoms qu'on rencontre là sont ceux de tout le monde, ceux que chacun connaît dans sa famille ou son entourage, au moins dans la génération des vieux. S'ils n'ont pas l'air de prénoms à la mode, cela tient à un double décalage : d'une part le corpus s'arrête en 1964, et en 25 ans la mode peut changer largement. D'autre part il ne s'agit pas d'une transcription fidèle des registres de l'état civil, mais d'un reflet éloigné qui passe par le filtre littéraire. Or le choix des écrivains a peu de chances de suivre la mode, puisque le cadre de l'action n'a aucun lien nécessaire avec l'actualité.

Mais on voudrait pouvoir embrasser d'un seul regard l'ensemble des sept périodes et des quatre siècles de littérature et en même temps l'ensemble des 330 éléments du calendrier, au lieu de focaliser l'attention sur une seule époque et sur les premiers de la liste. Il existe une méthode mathématique qui permet le large survol des données, quand celles-ci sont trop abondantes pour être saisies d'un seul coup par la conscience du chercheur. Cette vaste procédure de synthèse, qui, porte curieusement le nom d' "analyse factorielle", produit le résultat de la figure 13.

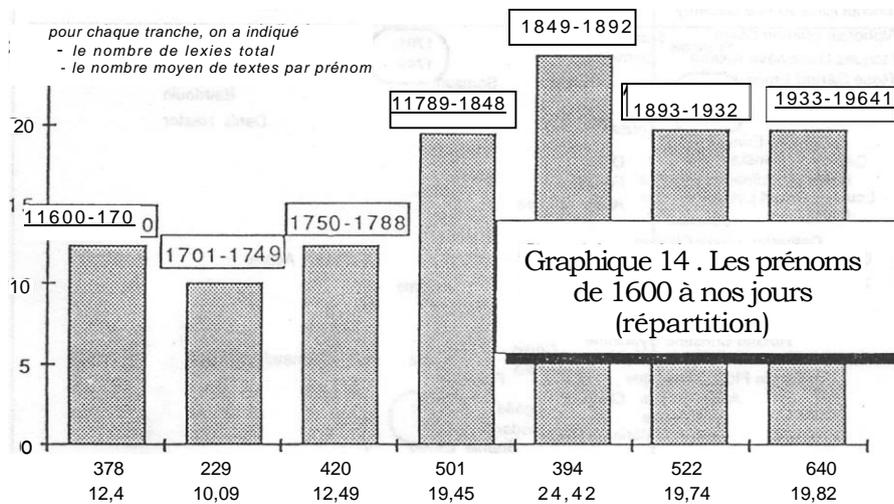
A vrai dire, il y a tant de gens dans cette enceinte que le regard a besoin d'être guidé pour comprendre le principe des regroupements. C'est le rôle des axes factoriels qui divisent l'espace en quatre quadrants, opposant tour à tour la gauche et la droite, le haut et le bas. A droite prennent position les deux premiers siècles, à gauche les deux derniers. Quant à la ligne horizontale, elle sépare les extrêmes et les moyens, comme il arrive chaque fois qu'on a affaire à des données sérielles. En réalité un seul principe de variation est clairement reconnu par l'analyse, et c'est celui du temps. Le temps gagne de proche en proche chacun des sept maillons de la chaîne chronologique, en décrivant un arc de cercle, du coin supérieur droit au coin supérieur gauche. Et, comme les députés se répartissent dans l'hémicycle suivant leur opinion politique, de même les **Pierre** et les **Paul** du calendrier prennent place là où le temps leur est le plus favorable. C'est ainsi que les éléments en faveur au XVIIe siècle occupent la moitié droite du graphique, là où l'analyse a placé les premières périodes. Et l'on n'aurait pas de peine à retrouver dans ces parages les éléments en excédent que le graphique 11 avait mis en relief, par exemple **Augustin** , **Constantin**⁽²¹⁾, **Ambroise** , **Alexis** , **Diane**, **Alexandre** , **Achille**, etc. De même le quadrant supérieur droit, qui est réservé au XXe siècle contient la plupart des noms de la figure 12

Figure 13. Analyse factorielle de 330 éléments du calendrier (répartition)



et pareillement les éléments du graphique 10 se répartissent en bas, à gauche, là où l'époque romantique prend position. Bien des indications particulières peuvent être extraites de ce graphique, dont nous retiendrons la conclusion essentielle : le temps opère une dérive progressive et régulière qui justifie notre étude. Le calendrier qui semble imposer son rythme immuable aux saisons est lui-même le jouet du temps, et le témoin des changements de l'histoire.

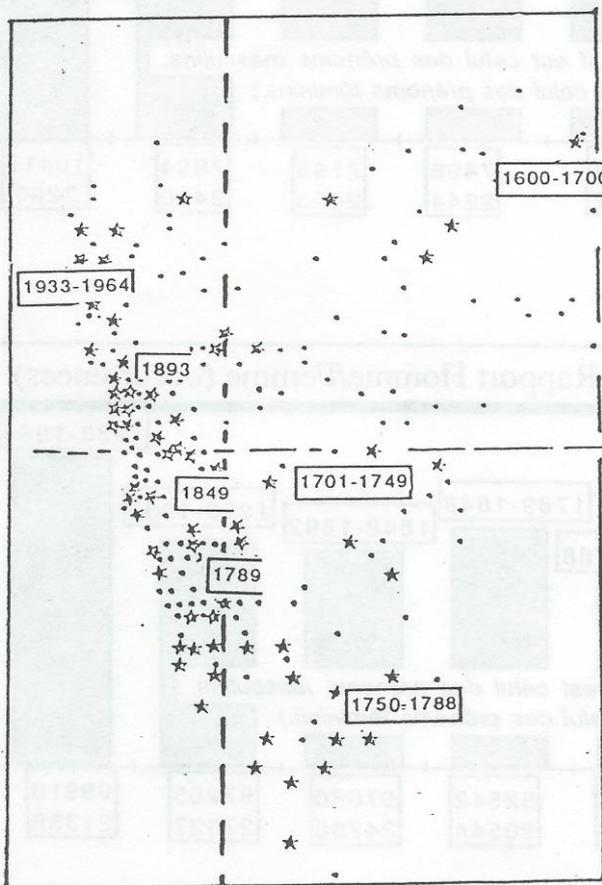
Deux autres enseignements peuvent être tirés de ce graphique : le premier confirme la courbe globale obtenue dans la figure 2, à partir des **occurrences** : le répertoire des noms gagne en abondance et en variété au cours des siècles. La concentration des éléments est plus forte dans la partie gauche du graphique 13, au point que la lisibilité en souffre. Dans la partie droite, qui est consacrée aux deux premiers siècles, les prénoms sont moins nombreux, mais plus excentriques, et ils se répartissent plus largement l'espace, pour faire contrepoids à la masse plus dense des éléments de gauche⁽²²⁾. Et si l'on procède, pour la répartition, au cumul des 330 éléments de chaque période, on obtient une courbe qui est parallèle à celle du graphique 2 et qui donne l'avantage aux siècles les plus récents. Ce qu'illustre la figure 14.

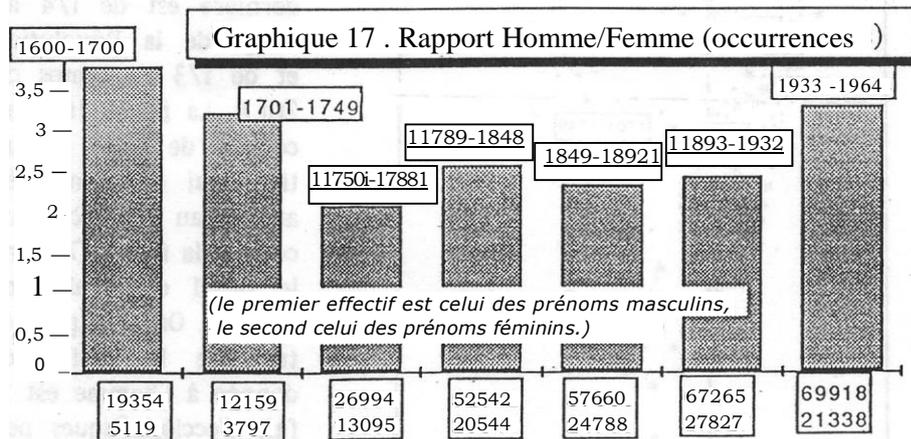
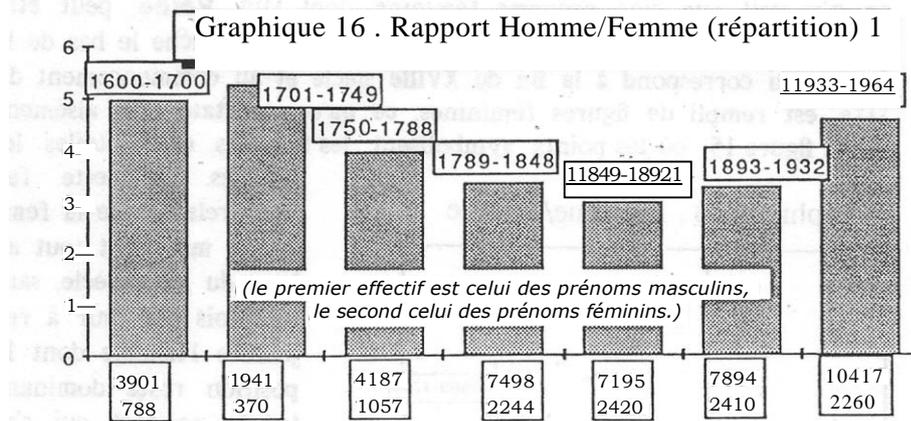


Le second enseignement est moins perceptible : observons pourtant le quadrant supérieur droit. Sur une cinquantaine d'éléments,

on n'y voit que cinq prénoms féminins, dont l'un, **Reine**, peut être récusé comme étant un titre et non un nom. En revanche le bas de la figure, qui correspond à la fin du XVIIIe siècle et au commencement du XIXe, est rempli de figures féminines, ce qu'on constate plus aisément sur la figure 15 où les points symbolisent les hommes et les étoiles les femmes. Et cette faveur relative de la femme se maintient tout au long du XIXe siècle, sans toutefois parvenir à rejoindre l'homme dont la position reste dominante. Le rapport, qui s'abaisse à 1/5 dans les premières périodes et la dernière, est de 1/4 au temps de la Révolution et de 1/3 au temps de ZOLA. La figure 16 rend compte de cette évolution, qui est confirmée au niveau des occurrences par la figure 17, dont le profil est également concave. On dira peut-être que la préférence donnée à l'homme est le fait d'ecclésiastiques peu portés au féminisme. On a tout lieu de penser que le nombre des martyrs des deux sexes s'équilibre : le bourreau ne fait pas de distinction coupable. Mais les autorités religieuses qui ont officialisé le martyrologe n'ont pas trouvé les mérites égaux et les saints authentifiés sont deux fois plus nombreux que les saintes (214 contre 116, quotient : 1,845). Il se trouve que les écrivains renchérissent sur les privilèges masculins, tant au niveau des occurrences (rapport variable entre 2,06 et 3,78), que de la répartition moyenne par texte (rapport entre 2,973 et 5,245) voir graphiques 16 et 17.

Graphique 15 . Homme/Femme





IV - Les noms bibliques

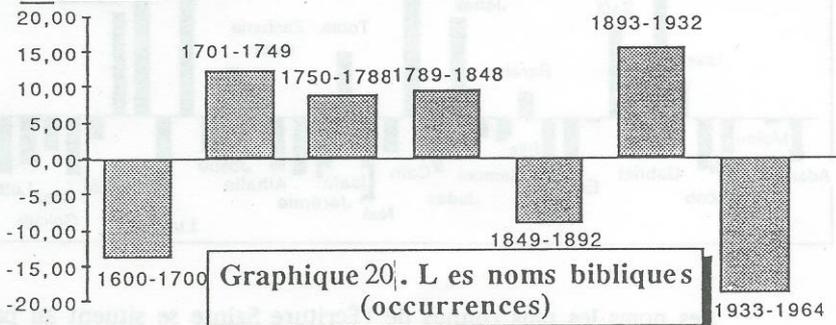
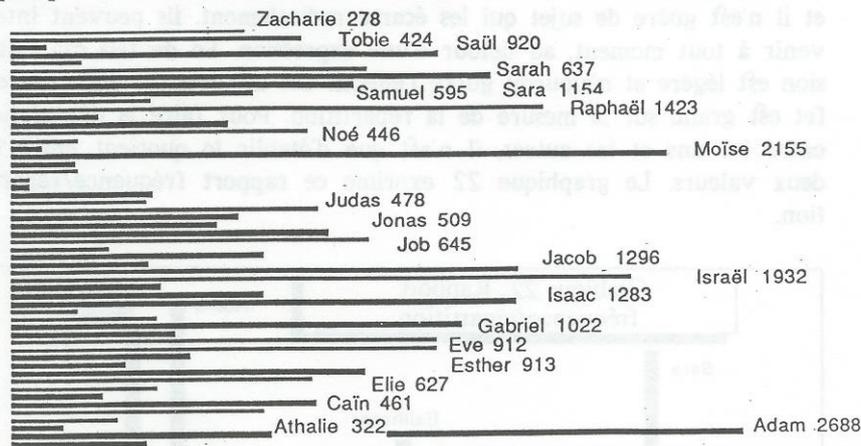
Le martyrologe de Grégoire XIII avait frappé d'ostracisme les noms de la Bible, même ceux des grandes figures et des grands prophètes, dès lors qu'aucun saint ou martyr ne pouvait être cité qui portât le même nom. Cela n'a pas empêché les écrivains de puiser à la source biblique, puisqu'il leur est loisible de faire concurrence à l'état civil. On trouve chez GIRAUDOUX un personnage qui s'appelle **Moïse** et chez ZOLA

un autre qui répond au surnom de **Goliath**. Et dans la société même, on rencontre sans peine des **Eve**, des **Adam**, des **Sarah**. Il y a donc quelque arbitraire à ne considérer que le calendrier liturgique, puisque ni la société civile, ni le royaume littéraire ne sont tenus d'obéir aux dispositions de l'autorité religieuse. Nous avons donc élargi le cercle de notre enquête aux noms les plus connus de l'Ancien Testament, en nous arrêtant au seuil de la divinité : ni Dieu, ni Jéhovah, ni Yahvé, ni le Messie, ni le Christ n'ont été pris en compte, non plus que Satan. Car la plupart de ces noms ont une fréquence très élevée qui eût écrasé le reste de la série⁽¹³⁾. Au reste ils ont déjà été traités dans notre étude du vocabulaire religieux auquel nous renvoyons le lecteur.

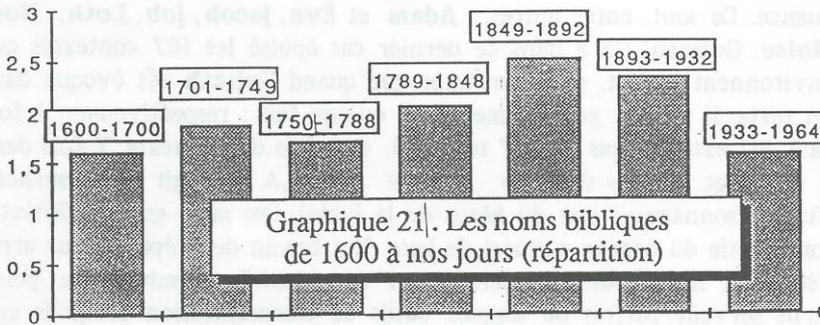
Les données sont restituées dans le tableau 18 (page suivante), tant en ce qui concerne la fréquence que la répartition. C'est le premier homme, **Adam**, qui marche en tête, avec 2.688 occurrences réparties dans 655 textes. Il précède **Moïse** (2.155 occ.) et **Israël** (1.932 occ.). En queue de liste se serrent les petits prophètes qui ont des effectifs étiques : **Osée** 39, **Habacuc** 23, **Michée** 21, **Sophonie** 7, **Abdias** 6 et **Nahum** 5. Voir graphique 19.

En regroupant cette cinquantaine d'éléments on obtient les deux histogrammes 20 et 21, lesquels ne semblent pas en parfaite harmonie, au moins pour la période centrale. Si l'on se fonde sur les occurrences, les éléments de la Bible paraissent en faveur au XVIIIe siècle (essentiellement à cause de VOLTAIRE), à l'époque romantique (CHATEAUBRIAND y est pour quelque chose) et lors du renouveau spiritualiste qui se développe à la charnière du XIXe et du XXe siècles. Ils restent discrets au siècle classique (sauf lorsqu'on met en scène un sujet biblique)⁽²⁴⁾, ils s'effacent aussi à l'ère du positivisme, et se font plus rares encore à notre époque. Tout cela est fort plausible et s'accorde avec l'évolution du vocabulaire religieux, tel qu'on l'observe parmi les noms communs. Pourquoi donc les faits de répartition ne donnent-ils pas exactement la même image ? En réalité l'évolution est bien la même, sauf sur un point : la seconde moitié du XIXe siècle produit un sommet au lieu d'un affaissement. La culture biblique reste donc vivante au temps de RENAN, puisqu'on en trouve l'empreinte dans de nombreux textes. Mais elle ne donne pas lieu alors à des ouvrages techniques dont l'influence est plus **sensible sur les occurrences**.

Graphique 19 . La fréquence comparée des noms bibliques



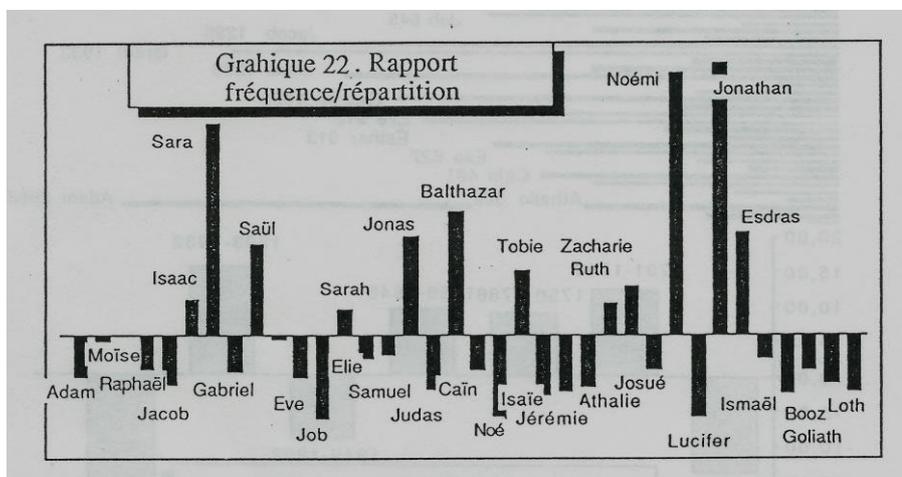
Graphique 20. Les noms bibliques (occurrences)



Graphique 21. Les noms bibliques de 1600 à nos jours (répartition)

On a le soupçon que les éléments du tableau n'ont pas les mêmes propriétés relativement à la fréquence et à la répartition. Certains ont une distribution concentrée. Ils groupent leurs occurrences en certains textes, où les attire quelque raison particulière, qui ne se repro-

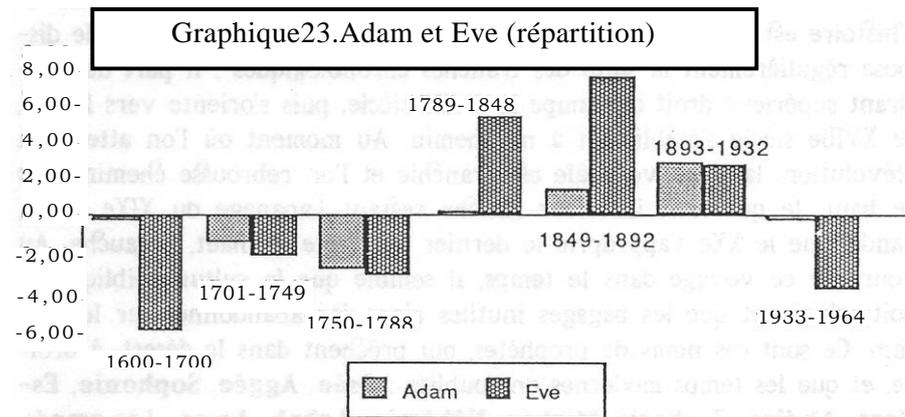
duit pas ailleurs. D'autres appartiennent au fonds commun de la culture et il n'est guère de sujet qui les écarte radicalement. Ils peuvent intervenir à tout moment, au détour d'une expression. En de tels cas l'allusion est légère et n'alourdit guère l'effectif des occurrences. Mais son effet est grand sur la mesure de la répartition. Pour faire la décantation entre les uns et les autres, il n'est que d'établir le quotient entre les deux valeurs. Le graphique 22 exprime ce rapport fréquence/répartition.



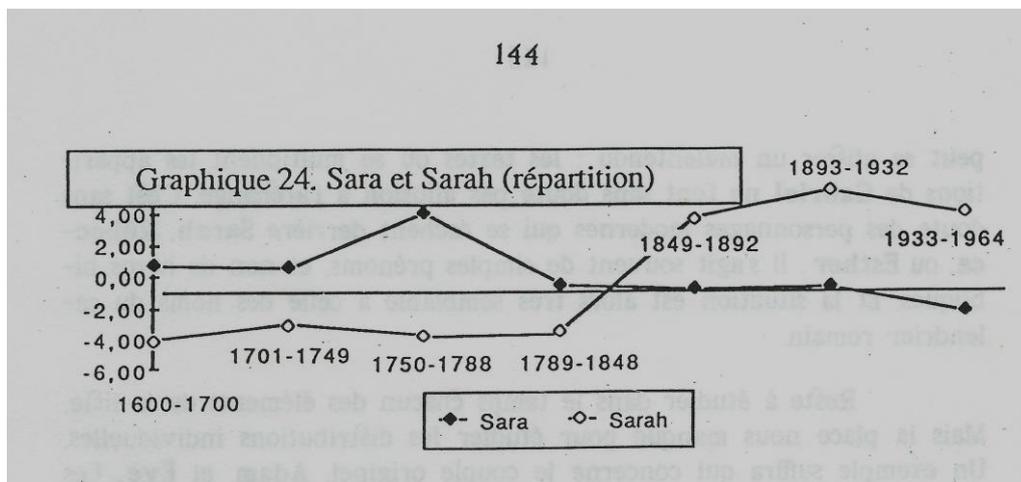
Les noms les plus connus de l'Écriture Sainte se situent au bas du graphique, là où la répartition l'emporte relativement sur la fréquence. Ce sont, entre autres, **Adam** et **Eve**, **Jacob**, **Job**, **Loth**, **Noé**, **Moïse**, **Goliath**. On a dans ce dernier cas épuisé les 107 contextes qui environnent le mot, pour constater que quand **Goliath** est évoqué dans un texte, il n'y est généralement cité qu'une fois : respectivement 1 fois dans 31 textes, 2 fois dans 7 textes, 3, 4, 6 fois dans 1 texte, 7 fois dans 2 textes et 35 fois dans *La Débâcle* de ZOLA (il s'agit là du surnom d'un personnage et non du héros de la Bible). Des mots comme **Goliath** font partie du bagage culturel de base. Nul besoin de préparer leur arrivée dans le discours. L'interlocuteur les identifie d'emblée, au point qu'ils servent parfois de simples outils de communication et qu'ils entrent dans des expressions ou aphorismes : « pauvre comme Job, la tenue d'Eve, l'arche de Noé ».

L'exemple du **Goliath** de *La Débâcle* nous montre aussi qu'il peut se glisser un malentendu : les textes où se multiplient les apparitions de **Gabriel** ne font sans doute pas allusion à l'archange. C'est sans doute des personnages modernes qui se cachent derrière **Sarah, Rébecca, ou Esther**. Il s'agit souvent de simples prénoms, et non de héros bibliques. Et la situation est alors très semblable à celle des noms du calendrier romain.

Reste à étudier dans le temps chacun des éléments de la liste. Mais la place nous manque pour étudier les distributions individuelles. Un exemple suffira qui concerne le couple originel, **Adam** et **Eve**. Ces deux noms restent associés dans le sentiment littéraire et le couple reste uni dans la courbe 23 qui les enlace. A aucun moment **Adam** ne se trouve séparé **d'Eve** par la ligne qui sépare les déficits des excédents, et quand ils franchissent cette ligne, cet ensemble, la main dans la main.

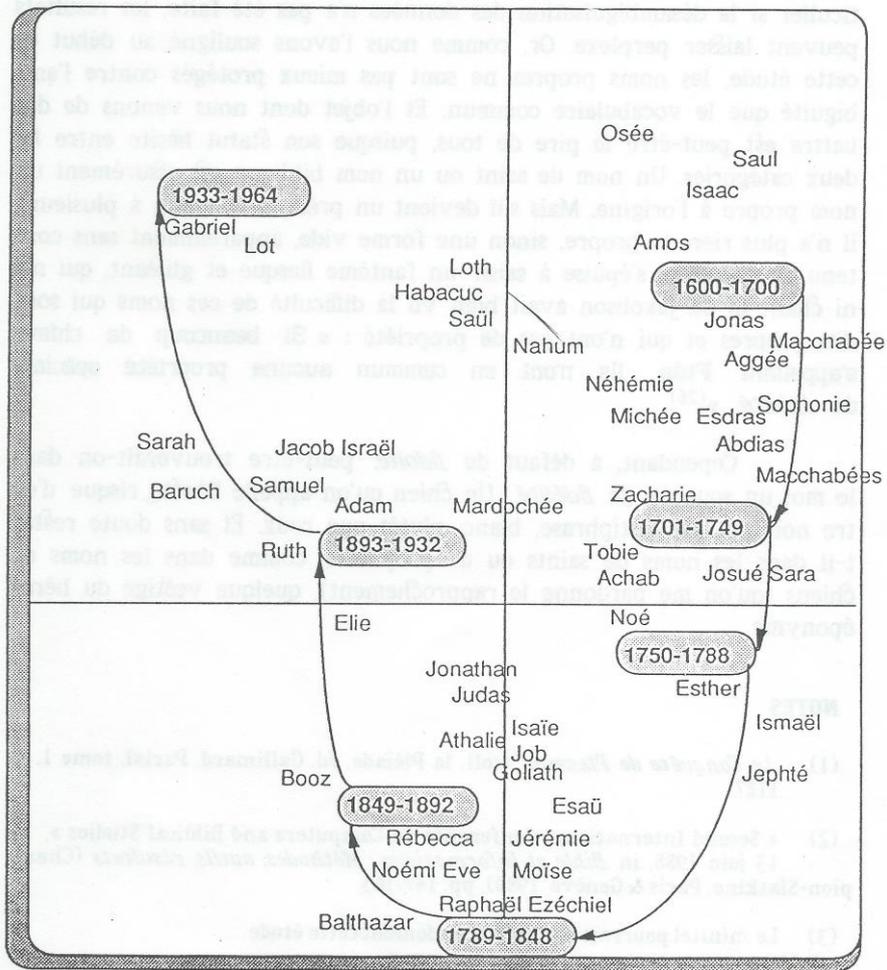


C'est le contraire qui se produit pour ce faux couple que **Sarah** forme avec **Sara**. Il s'agit en réalité de la même personne sous deux graphies voisines. Dès lors c'est la concurrence qui règne, et le partage haineux du même territoire. **Sara**, qui dominait jusqu'à la Révolution, doit désormais céder le pas à sa rivale⁽²⁵⁾.



L'analyse factorielle, reproduite dans le graphique 25, rend compte de ces menues observations. **Eve** et **Adam** sont situés sur la partie gauche, à une distance qui n'a rien de scandaleux chez un vieux couple. Par contre **Sarah** (à l'extrême gauche) s'oppose violemment à **Sara** (sur l'autre marge). Comme dans la figure 13, le mouvement de l'histoire est clairement reconnu par la machine. Et un arc de cercle dispose régulièrement la suite des tranches chronologiques : il part du quadrant supérieur droit où campe le XVII^e siècle, puis s'oriente vers le bas, le XVIII^e siècle s'établissant à mi-chemin. Au moment où l'on atteint la Révolution, la ligne verticale est franchie et l'on rebrousse chemin vers le haut, le quadrant inférieur gauche restant l'apanage du XIX^e siècle, tandis que le XX^e s'approprie le dernier qui prend place en haut, à gauche. Au cours de ce voyage dans le temps, il semble que la culture biblique se soit allégée et que les bagages inutiles aient été abandonnés sur le chemin. Ce sont ces noms de prophètes, qui prêchent dans le désert, à droite, et que les temps modernes ont oubliés : **Osée, Aggée, Sophonie, Esdras, Abdias, Zacharie, Michée, Néhémie, Achab, Amos**. Les grands prophètes résistent mieux et leur voix se fait entendre avant et après la Révolution, au bas et à gauche du graphique où se groupent **Isaïe, Elie, Jérémie, Ezéchiel**, en compagnie des grandes figures de la Bible comme **Job** et **Moïse**. De façon générale le premier axe qui sépare les deux derniers siècles des deux premiers ne conserve pour l'époque moderne que les noms illustres de l'histoire sainte, qui sont les plus fréquents, et tous les premiers de la lise se portent à gauche, dans la zone contemporaine : **Adam, Eve, Jacob, Israël, Samuel**. En revanche les curiosités et les raretés ont cessé de faire partie de l'héritage culturel. Il semble même que la tranche 1933-1964, qui occupe la zone nord-ouest, se désintéresse des choses de la Bible et que cette désaffection soit sensible sur le graphique, où cette zone ne

Graphique 25 . Analyse factorielle des noms bibliques



retient que de rares éléments, dont le statut est d'ailleurs douteux. **Gabriel** risque d'être un simple prénom, et non l'archange préposé aux messages célestes. **Lot** n'est pas nécessairement la seconde écriture de **Loth**, ce peut être une rivière ou un département.

On voit par ce dernier exemple assez grotesque que la statistique opère en aveugle, et qu'elle ne procède qu'à des manipulations formelles qui ne tiennent aucun compte de l'identité et du contenu des éléments traités. Si le statut de ces éléments est mal défini, et en particulier si la désambiguïsation de données n'a pas été faite, les résultats peuvent laisser perplexe. Or, comme nous l'avons souligné au début de cette étude, les noms propres ne sont pas mieux protégés contre l'ambiguïté que le vocabulaire commun. Et l'objet dont nous venons de débattre est peut-être le pire de tous, puisque son statut hésite entre les deux catégories. Un nom de saint ou un nom biblique est assurément un nom propre à l'origine. Mais s'il devient un prénom commun à plusieurs, il n'a plus rien de propre, sinon une forme vide, apparemment sans contenu. Et l'analyse s'épuise à saisir un fantôme flasque et glissant, qui n'a ni chair, ni os. Jakobson avait bien vu la difficulté de ces noms qui sont dits propres et qui n'ont pas de propriété : « Si beaucoup de chiens s'appellent **Fido**, ils n'ont en commun aucune propriété spéciale de *fidoïté* »⁽²⁶⁾.

Cependant, à défaut de *fidoïté*, peut-être trouverait-on dans le mot un soupçon de *fidélité*. Un chien qu'on appelle *Black* risque d'être noir ou, par antiphrase, blanc, plutôt que roux. Et sans doute reste-t-il dans les noms de saints ou de prophètes, comme dans les noms de chiens (qu'on me pardonne le rapprochement) quelque vestige du héros éponyme.

NOTES

- (1) *La Conquête de Plassans* (coll. la Pléiade, éd. Gallimard, Paris), tome 1, p. 1127.
- (2) « Second International Conference on Computers and Biblical Studies », 913 juin 1988, in *Bible et Informatique Méthodes, outils, résultats* (Champion-Slatkine, Paris & Genève, 1989), pp. 147-165.
- (3) Le minitel pourrait faciliter grandement cette étude.
- (4) Nous hésitons à parler ici de "mots", car les signes de ponctuation sont compris dans ce chiffre.
- (5) Cette base porte le nom de FRANTEXT. Son créateur est Jacques Dendien.

- (6) Le dix-septième siècle, cependant, n'a pas été morcelé, parce que son étendue était moindre.
- (7) On n'a malheureusement pas pris garde que certaines fêtes civiles, qui tombent à date fixe, ont aussi fait des victimes parmi les saints. Ainsi l'armistice du 11 novembre a été fatal à *Martin* qui fut pourtant le saint le plus fêté dans les anciennes paroisses françaises. Nous ne regrettons pourtant qu'à moitié cet oubli, car *Martin* est aussi un nom de famille extrêmement répandu, et le départ entre le prénom et le nom eût été fort malaisé à réaliser.
- (8) Voir le *Dictionnaire des fréquences*, Didier, Paris. Voir aussi notre *Vocabulaire français de 1789 à nos jours* (Slatkine, Genève).
- (9) Cet équilibre n'est qu'approximatif, la dernière tranche dépassant la somme des deux premières. Voici l'étendue relative de chacune : 11%, 7%, 13%, 17%, 15%, 17%, 21%.
- (10) On en a relevé 2.234.448 aux XIXe et XX siècles, c'est-à-dire dans les quatre dernières tranches du présent corpus.
- (11) Nous entendons par là non seulement les mots, mais aussi les signes inclus dans le discours et notamment les signes de ponctuation.
- (12) La pondération n'est pas un simple pourcentage, mais une mesure probabiliste connue sous le nom d'écart réduit. Bornons-nous à en indiquer la formule :

$$z = \frac{o - np}{\sqrt{npq}}$$

pour o = effectif de la catégorie (ou du mot) observé dans le sous-ensemble

n = effectif de la catégorie (ou du mot) dans l'ensemble

p = proportion du sous-ensemble dans l'ensemble

q = probabilité complémentaire ($q = 1 - p$)

(np représente l'effectif théorique de la catégorie ou du mot dans le sous-ensemble)

- (13) On a déjà remarqué que les noms propres ont partie liée avec le roman (et le théâtre) et qu'ils se raréfient au contraire dans les recueils de poésie et les ouvrages techniques. Voir notre *Vocabulaire français* tome 1, p. 280.
- (14) Victor Hugo et Chateaubriand sont les témoins privilégiés de ce changement de goût : voir notre *Vocabulaire de Hugo* (Slatkine, Genève) tome 1, p. 309.
- (15) La faveur dont jouit ce prénom tient peut-être aussi au recouvrement du calendrier solaire et du calendrier romain : le solstice d'été coïncide avec la fête de Saint Jean-Baptiste tandis que la fête de l'évangéliste se situe aux environs du solstice d'hiver, au temps de Noël,

- (16) Il y aurait beaucoup à dire sur les rapports complexes de la notoriété et de la popularité. D'une part un prénom peut devenir populaire lorsqu'il s'incarne dans un héros. D'autre part, les rois n'ont pas cherché l'originalité et l'exclusivité, mais bien au contraire, comme le peuple, le patronage des grands saints.
- (17) C'est aux *Misérables* qu'il faut imputer aussi la présence insolite de **Madeleine**, l'homonymie ayant confondu ce nom d'emprunt de Valjean avec un prénom féminin.
- (18) **Emma** est absente du graphique parce que seuls les cent prénoms les plus fréquents ont été pris en compte.
- (19) En réalité ce roman est si long qu'il a été découpé en 10 unités, ou 10 "textes-machines", une contrainte technique obligeant, à l'époque du dépeuillement, à ne pas dépasser la limite de 100.000 mots. L'effet de cette contrainte est plutôt heureux dans le cas présent, puisque cela contribue à équilibrer les "textes" et à donner à la notion de répartition une justification meilleure, La mesure de la répartition suppose en effet que les textes sont de taille égale ou voisine.
- (20) Il vaudrait mieux ne rien dire de **Florence**, qui désigne aussi la ville de Toscane, très présente dans *Lorenzaccio* L'Italie est alors à la mode et MUS- SET n'a pas d'exclusivité.
- (21) Beaucoup de noms en **-in** se trouvent ici, comme si ce suffixe latin était plus prisé à cette époque : en dehors de **Constantin** et **Augustin** , on note la présence de **Quentin, Romain, Paulin, Aubin, Urbain, Saturnin** .
- (22) Par un principe analogue à celui du levier en physique.
- (23) Le mot **Dieu**, qui a 79.312 occurrences, pèse plus que tous les autres éléments réunis.
- (24) Or les pièces bibliques de Racine, *Esther* et *Athalie*, sont absentes du corpus.
- (25) On observe une distorsion du même genre avec les deux orthographes de **Saül**.
- (26) *Essais de linguistique générale* (Paris, 1968), p. 178.