

# JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

MARIE-LOUISE DUFRÉNOY

## **La statistique et les tendances littéraires**

*Journal de la société statistique de Paris*, tome 95 (1954), p. 196-204

[http://www.numdam.org/item?id=JSFS\\_1954\\_\\_95\\_\\_196\\_0](http://www.numdam.org/item?id=JSFS_1954__95__196_0)

© Société de statistique de Paris, 1954, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

### La statistique et les tendances littéraires

La Statistique et la Littérature... Le rapprochement de ces deux mots évoque, dans la plupart des esprits, une irréductible antinomie, un antagonisme irréconciliable.

En l'espace de vingt minutes, nous espérons faire ressortir l'utilité des méthodes statistiques dans les recherches d'histoire littéraire. Nous nous efforcerons de prouver que ces méthodes permettent d'organiser logiquement une vaste matière bibliographique en révélant l'amplitude des mouvements et le dessin des tendances qu'une compilation arithmétique avait simplement indiqués.

Un mouvement donné étant connu dans sa ligne continue et les points d'inflexion de sa courbe étant déterminés, il devient possible d'étudier de façon précise les influences qui ont agi sur son destin et les causes de son apparition, de son développement et de son déclin.

En vue d'expliquer pourquoi, quand et comment, les méthodes statistiques peuvent être employées avec fruit dans les travaux d'histoire littéraire, j'aurai recours à l'expérience que j'ai acquise lorsque j'ai entrepris la tâche de retracer l'histoire du Roman Oriental en France au XVIII<sup>e</sup> siècle.

Comme M. André Maurois l'a très justement fait remarquer dans l'analyse qu'il a bien voulu consacrer à mon *Orient romanesque en France* : « Que l'Orient ait inspiré quelques chefs-d'œuvre du XVIII<sup>e</sup> siècle, comme *Les Lettres Persanes* ou *Zadig*, tout étudiant des lettres françaises le savait. Ce que l'on ignorait, c'était l'extraordinaire amplitude de la vogue de l'Orient littéraire au XVIII<sup>e</sup> siècle et son étonnante diversité. »

Je ne le soupçonnais pas, moi non plus, au début de mes recherches, sans quoi le courage m'eût sans doute manqué pour les entreprendre.

J'ai donc débuté par une patiente enquête bibliographique qui devait m'amener, en fin de compte, à réunir quelque 700 titres de romans et contes orientaux en une bibliographie de 509 pages.

M'étant mise en devoir d'organiser cette abondante matière, j'ai distingué des catégories : celle des ouvrages purement imaginatifs et fantastiques dans la veine des *Mille et une Nuits*, les contes galants, puis érotiques ; la satire qui revêt plusieurs formes : voyages imaginaires, lettres ou écrits allégoriques ; enfin, la littérature narrative idéologique, représentée par un petit nombre d'ouvrages particulièrement significatifs.

Ayant établi une bibliographie chronologique générale, je l'ai décomposée en bibliographies partielles, groupant les œuvres selon leur caractère dominant.

Ainsi ont commencé à se dessiner plusieurs vagues qui, comme celles de l'Océan, se recouvraient partiellement et dont le mouvement demeurerait assez mollement indiqué.

C'est ici que les méthodes statistiques et graphiques sont venues m'offrir un secours important, et les résultats obtenus apportent la réponse aux trois questions que nous nous sommes posées. Pourquoi recourir aux mathématiques en histoire littéraire ? Pour interpréter un grand nombre de documents bibliographiques précédemment dénombrés.

A quel moment doit-on faire appel à la statistique ? Lorsque les éléments à étudier sont groupés en catégories homogènes. Comment doit-on procéder ?

Voici la méthode que j'ai adoptée.

Si nous inscrivons, sur une ligne horizontale, les périodes quinquennales de la fin de 1694 à la fin de 1804, et que, de chaque point de cette échelle des temps, correspondant aux années 1695, 1700, ... 1750, ... 1800, 1805, nous élevions une droite verticale, d'une hauteur proportionnelle au nombre des romans d'inspiration orientale, publiés en France durant la période quinquennale précédente, nous obtenons une série de barres dont le sommet dessine le contour d'une « courbe en cloche ».

Au lieu de rendre chaque barre proportionnelle au nombre des romans publiés au cours de chaque période quinquennale successive, élevons chaque barre à une hauteur proportionnelle au nombre (c'est-à-dire à la fréquence cumulative) de tous les romans publiés depuis 1695, nous obtenons un second graphique par barres : le sommet des barres dessine le contour d'une courbe en S, ou « courbe sigmoïde » représentative de l'allure de la croissance du Roman Oriental en France, au cours du XVIII<sup>e</sup> siècle.

La croissance débute lentement, s'accélère à partir de 1715, atteint son maximum de vitesse vers 1750, au milieu du siècle, alors que sont déjà publiés 50 % des Romans Orientaux, dont le total atteindra près de 700 au début du XIX<sup>e</sup> siècle (Graphique I, *Journal Soc. Statistique*, 96<sup>e</sup> année, 1945, p. 263).

Il est facile de déterminer l'équation qui peut le plus fidèlement rendre compte de l'allure de la courbe, c'est-à-dire interpréter mathématiquement la vogue du Roman Oriental en France au XVIII<sup>e</sup> siècle.

Il suffit, en effet, d'utiliser l'un des papiers conçus en vue de l'établissement de graphiques transformant en une ligne droite des courbes en S.

La courbe en S, telle que nous venons de la définir, représente la courbe d'intégration de la courbe en cloche; elle doit donc pouvoir se convertir en droite, par transformation des « fréquences cumulatives » en « pourcentages du total » (c'est-à-dire dans le cas du Roman Oriental, en pourcentage de 700) et par l'utilisation, pour la représentation graphique de ces pourcentages, d'une échelle de « probabilité normale ».

Par exemple, si, dans le graphique par barres ayant servi à établir la courbe, les fréquences sont représentées par ( $y$ ), les pourcentages correspondants, sur échelle de probabilité normale, seront

$$p = \frac{100}{\sqrt{2} \pi} \int_{-\infty}^x e^{-\frac{x^2}{2}} dx.$$

Cette fonction permet de transformer la courbe sigmoïde en une droite; il suffit de calculer chaque fréquence cumulative ( $y$ ) sous forme de pourcentage du total (ici 700), puis de reporter ces pourcentages sur l'échelle de probabilité normale, chacun à la date qui lui correspond. On obtient une série de points qui s'alignent très sensiblement sur une droite (Graphique Ia). On peut donc conclure que la courbe en S représentant la tendance du Roman Oriental au xviii<sup>e</sup> siècle est la courbe d'intégration de la courbe en cloche.

Dans la mesure où la « croissance » du Roman Oriental peut se comparer à celle d'une population d'individus, elle doit pouvoir se représenter par la courbe logistique; cette courbe peut être convertie en une droite lorsque la transformation des fréquences ( $y$ ) en pourcentages ( $p$ ) met en œuvre la fonction

$$p = \frac{100}{1 + e^{-l}}.$$

Les pourcentages étant alors portés sur l'échelle logistique définie par cette fonction, on obtient une série de points; l'alignement linéaire des points est satisfaisant pour la période 1715-1795 (Graphique Ib).

Pour cette période au moins, la courbe de croissance du Roman Oriental peut être considérée comme correspondant à une courbe logistique.

La fonction inverse définit le « logit »

$$l = \ln \frac{p}{100 - p}.$$

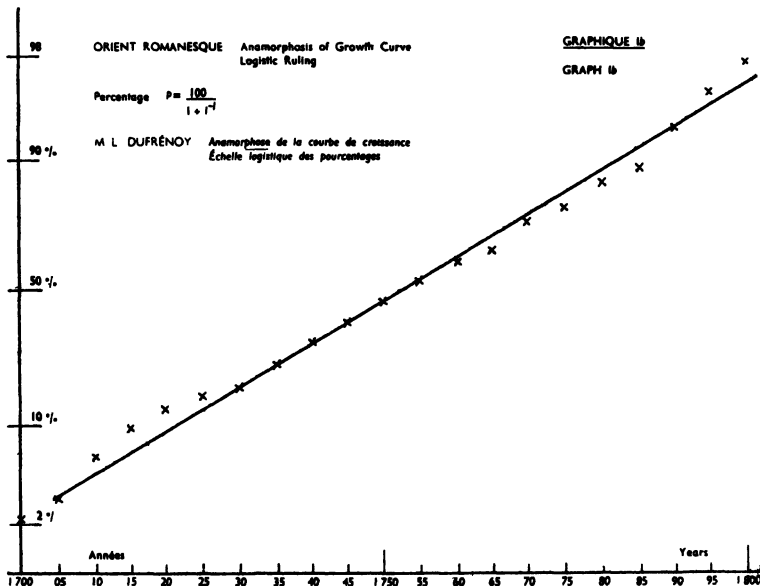
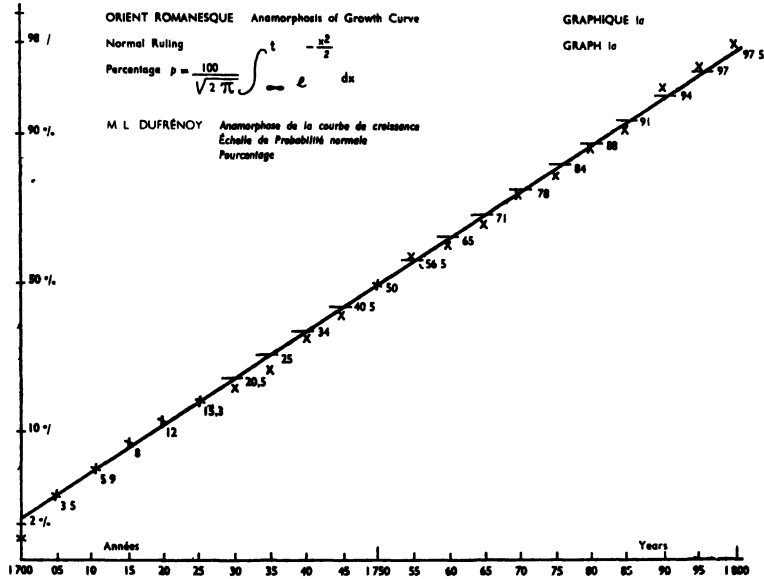
L'introduction du « logit » met en évidence le fait que la vitesse de la croissance est, à chaque moment, proportionnelle au pourcentage de la croissance déjà effectué ( $p$ ) et inversement proportionnelle au pourcentage de la croissance restant à effectuer ( $100-p$ ). Cette relation définit la croissance « autocatalytique », si l'on compare la croissance organique à la marche d'une réaction chimique.

Une relation du même type pourrait, d'ailleurs, être calculée pour rendre compte de la vogue du Roman Oriental considérée comme la manifestation d'un phénomène contagieux.

Les transformations généralement proposées en vue d'obtenir l'anamorphose en droite d'une courbe de croissance portent sur les fréquences ( $y$ ), qu'il s'agit

de transformer en probabilités ( $p$ ), puis de porter sur une échelle convenablement choisie.

Cette anamorphose peut aussi s'obtenir, en conservant les fréquences ( $y$ ),



mais en utilisant pour la représentation des temps l'échelle « arc-tangente » (Graphique 2).

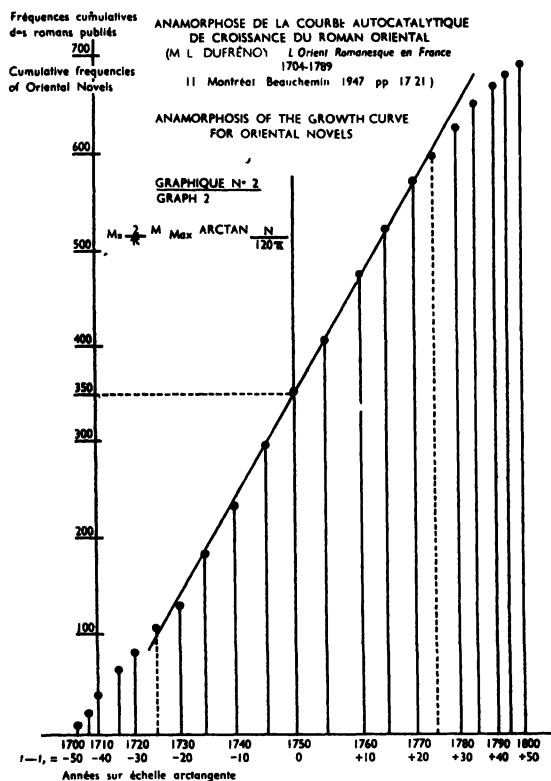
Au lieu d'employer une échelle de probabilité normale après transformation des fréquences ( $y$ ) en probabilité ( $p$ ) selon la relation

$$p = \left( \frac{100}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx \right)$$

on emploie l'échelle arc-tangente après transformation des temps (N) selon la relation

$$M = \frac{2}{\pi} M_{\max} \arctan \frac{N}{120\pi}$$

où M représente les distances linéaires mesurées à partir de zéro, sur l'échelle,



et  $M_{\max}$  la longueur totale de l'échelle employée, comptée à partir de l'origine zéro.

Les fréquences cumulatives de la colonne (1) ont été calculées en considérant la courbe de croissance du Roman Oriental comme une courbe autocatalytique, en utilisant l'équation  $\log \frac{x}{700-x} = 0.155 (t-t_1)$  et en employant les *Tables for the computation of curves of autocatalysis* de T. B. Robertson (*U. C. Publ. in Physiology*, 4 : 221-8, 1915).

Les fréquences de la colonne (2) ont été déterminées d'après la transformation en droite de la courbe de croissance observée, par utilisation de l'échelle arc-tangente comme échelle des temps.

Les pourcentages (col. 3) résultant de la transformation en « pour cent de 700 » des fréquences de la colonne (1) étant reportés sur échelle « normale » ou sur échelle « logistique » donnent des séries de points d'après lesquels on peut estimer les chiffres des colonnes 4 et 5.

TABLEAU I

*Fréquences de publications d'inspiration orientale*

DATES	PÉRIODES quinquen- nales comptées à partir de 1750	FRÉQUENCES CUMULATIVES			POURCENTAGES		
		Observées (0)	Calculées		Observés (3)	Théoriques	
			(1)	(2)		Normaux (4)	Logistiques (5)
1700. . . . .	— 10	11	19		1,6	2,5	2,5
1705. . . . .	— 9	24	27	10	3,4	3,5	3,5
1710. . . . .	— 8	40	39	20	5,9	5,9	5
1715. . . . .	— 7	64	54	40	9,0	8,5	6,5
1720. . . . .	— 6	84	80	80	12,0	11,9	9,5
1725. . . . .	— 5	107	111	120	15,3	15,6	14
1730. . . . .	— 4	130	137	155	18,8	21,5	19
1735. . . . .	— 3	169	180	200	23,0	27,5	26
1740. . . . .	— 2	227	230	245	32,4	34	33
1745. . . . .	— 1	278	288	295	39,9	42	41
1750. . . . .	0	349	350	350	50,0	50	50
1755. . . . .	1	403	412	405	57,5	58	59
1760. . . . .	2	438	470	455	62,3	66	67
1765. . . . .	3	478	520	500	68,3	72,5	74
1770. . . . .	4	542	563	545	77,4	78,5	81
1775. . . . .	5	577	599	580	82,4	84,4	86
1780. . . . .	6	609	630	620	87,0	88,1	89,5
1785. . . . .	7	631	656	660	90,1	91,5	92,5
1790. . . . .	8	666	675	680	95,1	94,1	95
1795. . . . .	9	677	683	690	96,9	96,5	96,5
1800. . . . .	10	687	691		98,1	97,5	97,5

La vogue de l'Orient Romanesque, qui, dans l'ensemble, peut être comparée à une vague, se compose de plusieurs ondulations secondaires, pouvant chacune se représenter par une courbe de croissance autocatalytique. Deux des manifestations les plus importantes de l'Orient Romanesque sont le Conte galant oriental et la Satire orientale (Graphique IV, *ibid.*, p. 267).

Établies par les méthodes que nous venons de décrire, les courbes de ces deux mouvements ont fait apparaître avec précision le point d'inflexion représentant les apogées successives des deux vagues sensiblement concomitantes.

Les sommets des courbes nous ont permis de circonscrire dans le temps les minutieuses investigations qui devaient nous amener à découvrir les causes déterminantes de l'impulsion donnée à un genre, ainsi que la genèse des phénomènes d'inhibition destinés à provoquer le ralentissement du mouvement et son arrêt.

Nous avons montré, dans *l'Orient Romanesque en France*, que le traducteur des *Mille et une Nuits*, Antoine Galland, ayant ouvert la voie aux Orientalistes littéraires, Montesquieu a fait école en réalisant la fusion harmonieuse de la matière orientale et de la satire, moins de vingt ans avant que Crébillon fils n'imprimât à l'Orient galant une déviation fatale, en suscitant par le succès du *Sopha*, nombre de pastiches inférieurs au modèle.

Le principal responsable du renversement des tendances est cependant l'auteur des *Mémoires d'un homme de qualité* et de *l'Histoire de Manon Lescaut*, qui fut aussi le traducteur de la *Pamela* de Richardson.

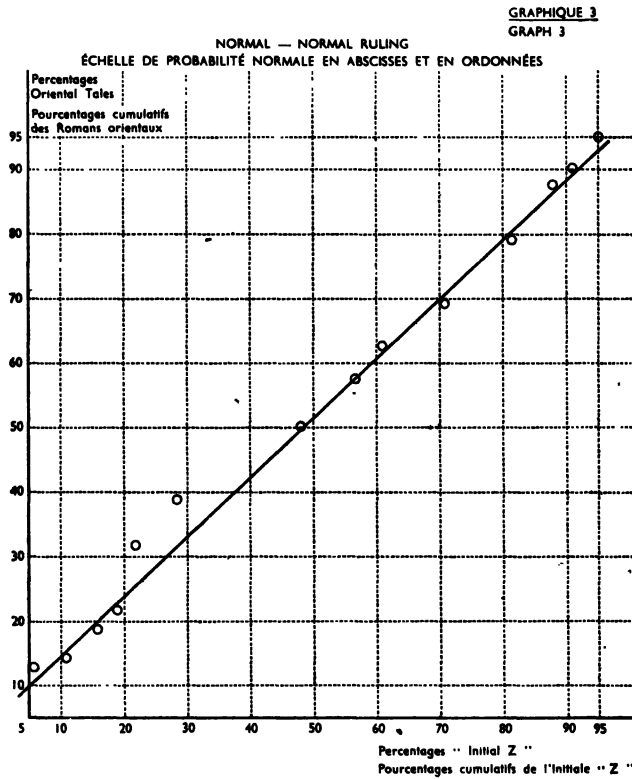
L'abbé Prévost, adepte d'un exotisme éclectique, a précipité le roman orien-

tal vers sa perte en se faisant l'artisan d'une nouvelle vogue : celle du roman anglais.

Nous pouvons d'ailleurs faire apparaître le cycle de l'Orient romanesque, non plus à la faveur de l'analyse statistique des fréquences de publications, mais par l'étude de quelque autre manifestation d'inspiration orientale dans la littérature d'imagination : l'emploi de la lettre Z comme initiale du nom des personnages de roman, par exemple.

La lettre Z a été employée comme initiale de 225 noms de héros, dans les romans publiés entre 1699 et 1804 (*Journal Soc. Statist.* 86 : 265).

Les fréquences cumulatives par périodes quinquennales se distribuent selon



une courbe sigmoïde : la somme des fréquences manifestées durant la première moitié du cycle, de janvier 1699 à juin 1752, devrait donc être théoriquement de  $225/2$  soit 112.5.

Le graphique III (*Journal Soc. Statist.* 86 : 266) illustre l'accord satisfaisant entre la courbe de croissance autocatalytique et la distribution des points représentant les fréquences observées.

Le graphique VI (*ibid.*, p. 268) fait apparaître le synchronisme relatif des manifestations de la vogue de l'Orient : fréquences des romans orientaux (x), des romans galants ( $\Delta$ ), de la lettre Z comme initiale (o), exprimées, en ordonnées sur l'échelle des « probits » d'après la Table de Fisher et Yates.

C'est encore à la statistique que nous avons eu recours pour mettre en évidence l'amplitude du mouvement romanesque oriental au XVIII<sup>e</sup> siècle, en France et en Angleterre.



Un seul graphique, sur papier logarithmique (VII, *ibid.*, p. 269), permet la comparaison directe des fréquences cumulatives des romans en général, des romans orientaux publiés en France, des « Oriental Tales » et des traductions anglaises de romans français publiées en Angleterre.

Les points représentant les fréquences s'alignent sur des segments de droites, avec changements de pente concomitants des changements de tendances. L'époque de la plus grande vogue du roman en France (1709-1749) a été aussi celle de la plus grande vogue du roman oriental; la vogue en Angleterre s'est efforcée de suivre, avec dix ans de retard. Mais, par un phénomène de réaction, l'influence anglaise devait, dans la seconde moitié du siècle, susciter en France des tendances qui supplantèrent progressivement celles auxquelles le roman oriental devait un si grand essor.

Aux érudits qui se déclarent horrifiés des incursions de la statistique dans la littérature, nous répondrons que nous n'avons pas la responsabilité de l'invention. L'idée, en effet, n'est pas neuve.

Pascal estimait que les lois élaborées par les mathématiciens français du xvii<sup>e</sup> siècle, pour prévoir les résultats des jeux de hasard, pouvaient s'appliquer aux résultats de l'activité intellectuelle. « La fièvre a ses frissons et ses ardeurs », dit-il, « (le flux et le reflux)... Les inventions des hommes de siècle en siècle vont de même ».

Au siècle suivant, Montesquieu devait avoir la prescience de l'« histoire naturelle des lois ». « Les phénomènes de l'Histoire », aux yeux de Montesquieu, « ne se distingueraient qu'en apparence de ceux de la nature, mais en réalité seraient soumis comme eux à des lois invariables ». « La détermination de ces lois », tel serait, pour F. Brunetière (*Questions de critique*, Paris, 1889, p. 102), « l'objet propre de l'Esprit des Lois ».

Dans le chapitre consacré à l'« Histoire littéraire » de *La méthode dans les sciences* (Paris, Alcan, 1911), G. Lanson a nettement pris position contre Brunetière en déclarant que le « parti pris... d'employer les formules des sciences physiques et naturelles » condamnait ce dernier « à déformer et à mutiler l'histoire littéraire ». Ainsi s'exprime G. Lanson :

« Désabusons-nous des courbes, dont nous faisons le symbole du développement des idées littéraires. Mais elles y supposent ou elles y introduisent : 1<sup>o</sup> l'unité; 2<sup>o</sup> la continuité. Or, il y a des mouvements qui éclatent comme des épidémies en plusieurs lieux à la fois, et des genres qui naissent deux ou trois fois avant de vivre. »

Lanson n'avait pas envisagé les conséquences qui découlent naturellement du caractère épidémique attribué à une vogue littéraire. Il est, d'ailleurs, excusable, car, de son temps, les phénomènes de contagion n'avaient pas encore fait l'objet d'une interprétation mathématique.

G. Lanson reconnaît, cependant, que l'« influence de la littérature sur la Révolution ne sera tant bien que mal perceptible que lorsqu'on aura observé patiemment de 1715 et même de 1680 à 1789 les échanges multiples qui se sont faits sans interruption entre la littérature et la vie ».

Dès 1797, Chateaubriand s'était efforcé d'établir une semblable corrélation, dans son *Essai historique, politique et moral sur les Révolutions*. Quarante ans plus tard, Chateaubriand avait révisé beaucoup de ses jugements, mais, dans

son *Essai sur la littérature anglaise et Considérations sur les génies des hommes, des temps et des révolutions*, il a mis de nouveau en évidence les rapports étroits qui lient une révolution littéraire à une révolution politique. Il revendiquait la responsabilité d'une volte-face radicale dans le domaine des lettres et fit même allusion au caractère contagieux des mouvements littéraires : « Novateur né, j'aurai peut-être communiqué aux générations nouvelles la maladie dont j'étais atteint » (*Mémoires d'Outre-Tombe*, livre IV, chap. 12). Chateaubriand montre comment la désintégration du roman du xviii<sup>e</sup> siècle a fourni les éléments essentiels à la naissance du roman « romantique ».

Un siècle plus tard, Silas Paul Jones a soigneusement compilé une liste chronologique des romans publiés en France de 1700 à 1750 et a essayé de montrer, au moyen de graphiques, l'accroissement numérique progressif des œuvres publiées pendant la première moitié du xviii<sup>e</sup> siècle.

Le roman en général comprend un trop grand nombre de types pour se prêter à une étude numérique des tendances. Un tel travail devient significatif lorsqu'il est limité à un genre plus nettement défini.

Longtemps avant que je n'eusse retracé le flux et le reflux du roman oriental au xviii<sup>e</sup> siècle, d'éminents lettrés avaient mené à bien une tâche similaire et leurs résultats avaient déjà confirmé la théorie émise par Pascal dans ses *Pensées*. Les genres qui naissent deux ou trois fois avant de vivre trahissent le rythme intérieur de l'évolution de la pensée que la statistique essaie d'appréhender et de traduire.

Bien que, comme la bibliographie que j'ai pu établir le fait ressortir, la statistique ait été surtout employée pour l'étude du vocabulaire, quelques articles récemment parus sont consacrés à la théorie de la contagion appliquée à la diffusion de différentes formes d'activité intellectuelle et à la détection des tendances.

Marie-Louise DUFRÉNOY.

---