HAL
open science

Peut-on régler son compte à la “ raison ” ?

Étienne Brunet

▶ **To cite this version:**

Étienne Brunet. Peut-on régler son compte à la “ raison ” ?. Vuillemin, Lenoble. Littérature, Informatique, Lecture, PULIM, pp.135-145, 1994, 2-84287-108-1. hal-01477568

HAL Id: hal-01477568

https://hal.science/hal-01477568v1

Submitted on 4 Mar 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.

Peut-on régler son compte à la « raison » ?

Etienne Brunet

Est-il bien raisonnable d'aborder la *raison* par les chiffres ? De même qu'il est difficile de circonscrire dans le cerveau le territoire occupé par la raison, et de délimiter les lobes et les cellules dévolus à cette faculté, de même il semble peu plausible qu'on puisse rendre compte avec des calculs et des décimales de la place que tient la raison dans le discours. Pourtant la raison a partie liée avec les comptes. Un livre de raison est un livre de compte. Et le sens mathématique du mot remonte à la racine étymologique. Le verbe latin *"reor"* signifie d'abord calculer. Et curieusement c'est aussi le cas de *"putare",* ce qui justifie le nom que les anglophones donnent à l'ordinateur. Oserons-nous suggérer, en prenant appui sur les mots, que le calcul est peut-être le premier état de la pensée. Si l'on entreprenait d'apprendre à raisonner à l'animal, il serait probablement judicieux de commencer par les nombres. En tout cas c'est par les chiffres que s'est fait le dressage initial de cet animal domestique qu'est l'ordinateur.

Et sans plus attendre lançons le limier électronique sur la trace de la "raison". Le terrain de chasse est très vaste, ce qui convient à l'ardeur dévorante de la machine. Il s'étend sur quatre siècles de littérature, de 1600 à nos jours, et comprend les 3 222 textes répertoriés dans *Frantext.* D'un coup c'est 155 millions de mots qui se trouvent disponibles et accessibles par la voie télématique. Encore faut-il désigner la cible à l'ordinateur, dont la naïveté risquerait d'être abusée par les déguisements pourtant transparents de la dérivation. On a donc commencé par dresser la liste des formes qui s'appuient sur le radical "raison". Et pour faire bonne mesure on a recensé aussi celles qui sont bâties sur le radical latin *"ratio".* Par contre, dans cette première approche, on a négligé les préfixes et les formes composées (du type "déraison"), et l'on s'est contenté de l'infinitif du verbe "raisonner".

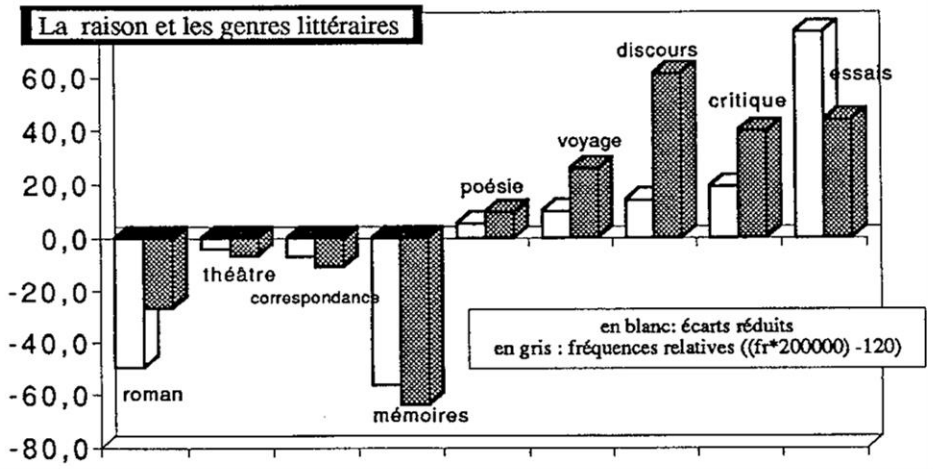
Quelques secondes ont suffi à la machine pour débusquer le gibier. Aucun des 92 921 repaires n'a pu échapper à son oeil de faucon, comme on le voit dans le tableau de chasse qu'est le tableau 1. Les espèces y ont été distinguées, mais aussi les territoires y sont étiquetés, ce qui aboutit à un tableau à deux dimensions où les lignes représentent les formes et les colonnes les genres littéraires. Précisons que nous avons adopté la catégorisation des textes établie à Nancy, sans trop approfondir, cas

par cas, le bien-fondé des étiquettes. Car on pourrait parfois hésiter avant de décider qu'un texte est un récit de voyage ou une oeuvre de mémorialiste, ou avant de ranger dans la correspondance un roman par lettres (sans parler des mélanges dont le nom même est un aveu). Mais la statistique est habituée à ces situations impures, où les données semblent moins données que fabriquées, avec un dosage suspect d'imprécision, d'incomplétude, et d'arbitraire.

Tableau 1. Les dérivés de la "raison". Effectifs par genre littéraire.

	roman	*théâtre*	*corresp*	*mémoire*	*poésie*	*voyage*	*oratoire*	*mélange*	*essai*	*TOUT*
raison	15450	3631	2919	2919	2919	2068	728	2505	24227	57366
raisonnablement	184	20	34	34	34	4	10	25	361	706
raisonnable	2403	388	373	373	373	87	108	345	2532	6982
raisonnant	62	7	4	4	4	10	8	16	129	244
raisonné	182	16	35	35	35	6	17	71	506	903
raisonnement	1136	107	159	159	159	47	61	319	2933	5080
raisonner	1162	175	149	149	149	91	63	299	1768	4005
raisonneur	117	14	7	7	7	21	0	25	114	312
raisonneuse	27	7	5	5	5	3	1	1	12	66
raisons	4221	692	888	888	888	211	151	582	5124	13645
ratio	6	0	1	1	1	1	1	4	28	43
ratiocination	3	0	2	2	2	1	0	0	8	18
ratiociner	10	1	1	1	1	1	1	1	5	22
ratiocineur	1	0	0	0	0	0	0	0	2	3
rational	4	1	0	0	0	1	0	0	11	17
rationalisation	3	0	0	0	0	0	0	6	128	137
rationaliser	0	0	0	0	0	0	0	2	23	25
rationalisme	31	0	5	5	5	0	4	4	366	420
rationaliste	51	0	2	2	2	0	0	3	238	298
rationalité	0	0	0	0	0	0	0	0	178	178
rationnel	80	3	27	27	27	3	4	21	2052	2244
rationnellement	7	0	1	1	1	0	0	2	195	207
total	25140	5062	4612	4612	4612	2555	1157	4231	40940	92921
étendue	54163610	8963391	8502767	16581985	7136490	3516123	1281813	5316524	50208835	155671538
p	0,3479	0,0576	0,0546	0,1065	0,0458	0,0226	0,0082	0,0342	0,3225	1
q	0,6521	0,9424	0,9454	0,8935	0,9542	0,9774	0,9918	0,9658	0,6775	0

Figure 2. La "raison". dans les genres. Représentation graphique



Le mot "raison" domine largement la famille des mots dérivés (57 366 occurrences sur 92 921, soit plus de la moitié.) De même les genres ne sont guère équilibrés, la production romanesque représentant un tiers du corpus, comme les essais, tandis que les autres genres se partagent le troisième tiers (dans lequel le

partage est également léonin, les mémoires s'arrogeant la meilleure part). Il importe donc de pondérer, ce qu'on peut faire de deux façons. On peut d'abord se contenter de simples pourcentages et *Frantext* s'engage dans cette voie qui désoriente moins les esprits littéraires. Il suffit d'établir le rapport entre la fréquence du mot considéré et l'étendue du texte ou du genre littéraire qu'on a isolé. Cela s'appelle fréquence relative. *Frantext* propose un histogramme, établi à partir des valeurs de cette fréquence relative. Nous l'avons reproduit en gris dans le graphique 2. Mais il nous a semblé que les écarts aléatoires pouvaient créer de dangereuses illusions d'optique lorsque les effectifs sont trop inégaux. C'est pourquoi nous avons préféré une mesure plus sûre : celle de l'écart réduit. Comme nous envisageons ici l'intégralité du corpus, soit N= 155 millions de mots (ponctuations comprises), la probabilité p , ou rapport de chaque genre à la totalité, soit Ni/N, est appliquée à la fréquence totale du mot ou de la série de mots, que *Frantext* peut fournir aussi (ici 92921). Ainsi, prenant l'exemple du roman, et sachant que ce genre occupe 35% du corpus (p 54163610/155671538 =0,3479, d'où $q = 1-p$ = 0,6521) et que le champ de la "raison" a 25140 occurrences romanesques et 92921 en tout, on en déduit :

*effectif théorique = 92921 * 0,3479 = 32330*
écart absolu = 25140 - 32330 = -7190
*écart réduit = -7190 / (32330 * 0,6521) = -49,52*

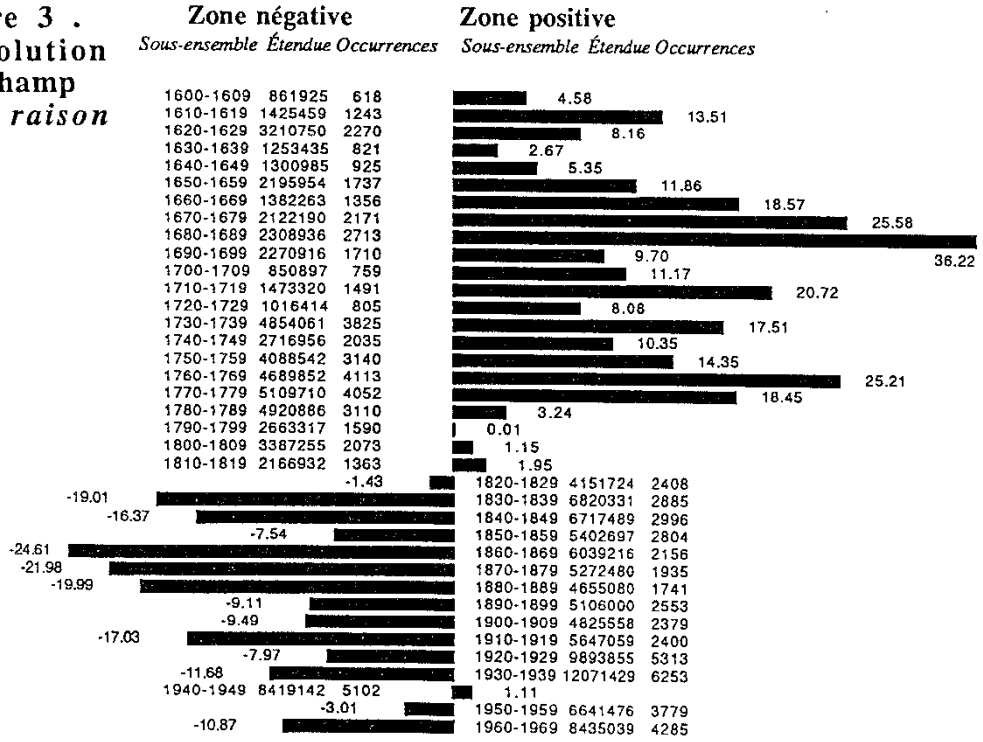
Le graphique 2 établit le calcul pour les genres comparés (histogrammes en blanc) et confirme les indications fournies par les fréquences relatives : déficit de la "raison" dans les oeuvres de fiction (roman et théâtre) et dans l'écriture personnelle (correspondance et mémoires), excédents au contraire dans les oeuvres techniques qui prennent pour objet le monde réel et le soumettent à la représentation mentale, à la discussion, en fin de compte à la raison : essais, critique, discours. Deux surprises apparaissent cependant : pourquoi les récits de voyage se rangent-ils de ce côté alors que les mémoires choisissent le camp opposé ? Et pourquoi la poésie fait-elle appel à la raison ? En réalité les genres littéraires n'étant pas également représentés à toutes les époques, on voit là l'effet de l'influence du facteur temps. Poésie et récits de voyage sont plus en faveur dans les siècles passés, au moment où la raison est elle-même dominante, comme nous le verrons plus loin. Deux influences ici se croisent et interfèrent et il faut se garder d'attribuer au genre ce qui appartient au temps.

Isolons maintenant la chronologie. Ici les procédures que propose *Frantext* sont plus simples et plus directes. Alors que le tableau précédent avait été obtenu colonne par colonne (chaque genre exigeant une interrogation particulière), ici l'histogramme est reproduit du premier coup, après qu'on a indiqué au programme la

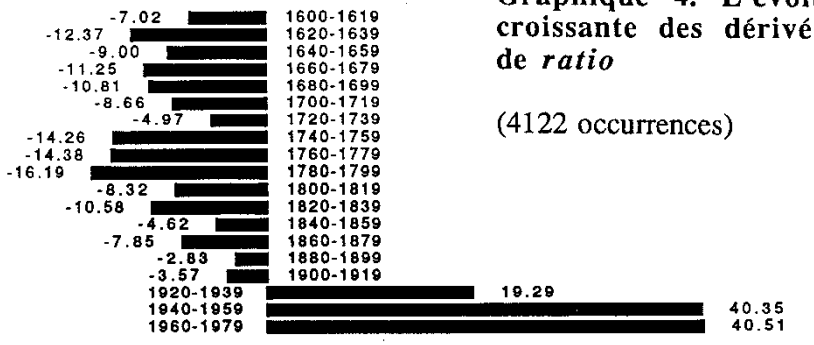
série de mots à analyser et la taille des tranches chronologiques qu'on souhaite appliquer à la segmentation des quatre siècles couverts par les données. Mais comme précédemment nous refuserons la mesure trop naïve de la fréquence relative en faveur de l'écart réduit. Le passage de l'une à l'autre mesure est d'ailleurs aisé car on peut retrouver la donnée manquante, *Ni* ou étendue de chaque sous-ensemble, à partir des deux données fournies par *Frantext :* fréquence absolue *(Fabs)* et relative *(Frel)* du mot : *Ni = Fabs / Frel.*

Le résultat se lit dans le graphique 3 dont la progression dans le temps se fait au pas de 20 ans. Pour chaque tranche on y lira successivement les dates limites, l'étendue en nombre de mots, l'effectif du champ de la "raison" et enfin l'écart réduit, lequel sert d'abscisse à la représentation graphique. Point n'est besoin d'être clerc ou mathématicien pour saisir l'orientation et la signification de la courbe. Toutes les tranches sont excédentaires de 1 600 à 1 800 et toutes sont déficitaires au-delà. On savait bien que les siècles classiques étaient dévoués à la raison et cette dévotion avait été clairement affichée, de Boileau à Robespierre. On voit grâce aux chiffres que la pratique de ce culte ne se démentit pas jusques et y compris la Révolution (qui va jusqu'à élever un temple à la déesse Raison) et qu'elle ne faiblit qu'à l'avènement du romantisme. Quand un mot est lourd de sens, comme le mot "raison", et qu'il porte une part des valeurs d'une civilisation, son destin reproduit l'évolution des mentalités et c'est ce qui rend si clair le graphique 3 où se reflète l'histoire des idées.

Figure 3 . L'évolution du champ de la *raison*



Un léger correctif doit cependant être apporté à la courbe précédente qui regroupe et confond les familles de "raison" et de *"ratio".* On peut vouloir isoler cette dernière famille, issue directement de la souche latine. Comme le poids de cette branche reste faible (1/20 de l'ensemble), la courbe générale en subit peu l'influence. Mais l'évolution y est inverse, et croissante. Toutes les ressources de la dérivation sont mises à profit pour enrichir le paradigme de formes savantes en *"iste",* en *"isme",* en *"lion",* en *"té",* en *"el",* en *"ali"* et en *"iser".* La variété des formes l'emporte sur celle de la raison-mère. Et parallèlement les occurrences de cette branche intellectuelle se développent, surtout au XXe siècle, en participant au progrès vers l'abstraction qu'on a observé à partir de l'ensemble des suffixes[1].



Graphique 4. L'évolution croissante des dérivés de *ratio*

(4122 occurrences)

Avouons cependant que ces graphiques revêtent un caractère abstrait dans la mesure où rien ne subsiste de l'auteur et du texte. Celui qu'on obtient ci-dessous - là encore directement à partir de *Frantext* - apparaîtra un peu moins désincarné car il reproduit le classement des écrivains selon la plus ou moins grande faveur qu'ils accordent à la "raison". Précisons que *Frantext* délivre une longue liste de plusieurs centaines de noms et qu'on a cru bon, pour plus de lisibilité, de ne retenir que les auteurs les mieux représentés dans le corpus (ayant plus de 300 000 occurrences). Ajoutons que si le classement du graphique 5 est bien celui *deFrantext,* établi sur les fréquences relatives, les valeurs qui fondent l'histogramme sont celles de l'écart réduit. Le premier de la liste est commun aux deux classements : c'est d'Holbach. Il est accompagné de beaucoup des philosophes du XVIIIe siècle : Condillac, Rousseau, Destutt de Tracy, mais aussi de penseurs moins anciens comme Proudhon, Renan, Bergson ou Alain. A l'inverse de l'observation précédente, c'est ici le genre littéraire qui perturbe la pureté du mouvement chronologique. L'évolution ne se vérifie correctement qu'à genre constant. Il suffirait d'écarter de la liste les philosophes de profession, pour opérer une décantation qui porterait à gauche parmi les partisans de la "raison" la

1. Voir notre *Vocabulaire Français de 1789 à nos jours,* éd. Slatkine, Genève, 1981, tome I, p. 415-602.

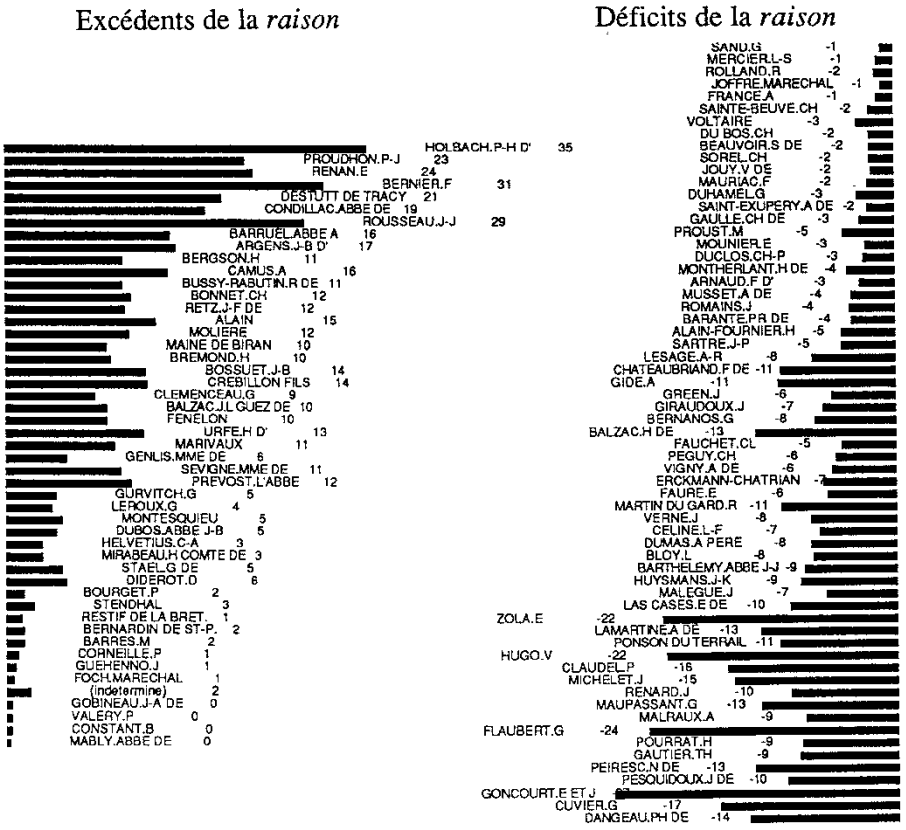
plupart des auteurs classiques ou post-classiques (Bussy-Rabutin, Retz, Molière, Bossuet, Fénelon, Guez de Balzac, Crébillon, Mme de Genlis, H. d'Urfé, Marivaux Mme de Sévigné, abbé Prévost, Montesquieu, Mme de Staël, Diderot, Corneille, etc.). Quant à la liste de droite, qui est par définition défavorable aux philosophes, elle exclut aussi systématiquement tous les auteurs antérieurs à la Révolution. La seule exception digne d'être notée concerne Voltaire. Quoique philosophe et fort attaché à la raison, Voltaire utilise relativement peu le mot "raison" (*z* -3). Nous avons montré ailleurs[2] que le mot "religion" est au contraire surabondant sous sa plume. C'est ici que le postulat statistique, qui infère de la fréquence à l'importance, trouve ses limites. Il peut arriver qu'on parle peu de ce qui est implicite et va de soi (comme la raison pour Voltaire) et qu'on s'acharne sur ce qui fait problème (la religion, par exemple, chez ce même auteur). La fréquence est le signe de l'intérêt, de la passion, voire de la hantise. Mais l'intérêt ou la passion peuvent être faits d'hostilité autant que d'adhésion.

Mais de quelle raison parle-t-on ? S'agit-il de la faculté pensante ? ou d'une proportion ? ou d'une cause ? La polysémie obère ici gravement les calculs et les conclusions. *Frantext* n'est pas doté de mécanismes qui puissent dissoudre les ambiguïtés - encore que de telles procédures sont à l'étude, mais leur application à un si gros corpus reste problématique. Cependant *Frantext* peut permettre à l'utilisateur de résoudre les cas d'homographie ou de polysémie en lui fournissant des contextes sélectifs. Rien n'empêche qu'on reçoive la totalité des contextes, et la machine ne répugnera nullement à délivrer les 57 366 contextes de la seule forme "raison". Mais il faudra de l'espace (plus de 20 millions de caractères), du temps (plusieurs heures de communication) et de l'argent (nous n'avons pas fait l'évaluation). Il a paru plus raisonnable de n'envisager que certaines expressions convenues qui relèvent de l'une ou l'autre des acceptions du terme et qu'on peut cataloguer immédiatement, sans qu'un examen cas par cas soit nécessaire. Afin de neutraliser notre propre subjectivité, nous avons puisé cette liste d'expressions dans *lePetit Robert.* La voici, avec le nombre d'occurrences pour chaque élément :

" comme de raison"	131	*" à tort ou à raison"*	163
" (ni) rime ni raison "	30	*" (plus) que de raison"*	43
" perdre (la) raison "	245	*" donner raison"*	400
" raison suffisante"	211	*" en raison de "*	429
" en raison d' "	72	*" (par ou pour) la raison" (que ou qu)*	387
" à raison de"	633	*"raison directe"*	78
" à raison d"	93	*" raison sociale "*	45
" raison inverse"	190		

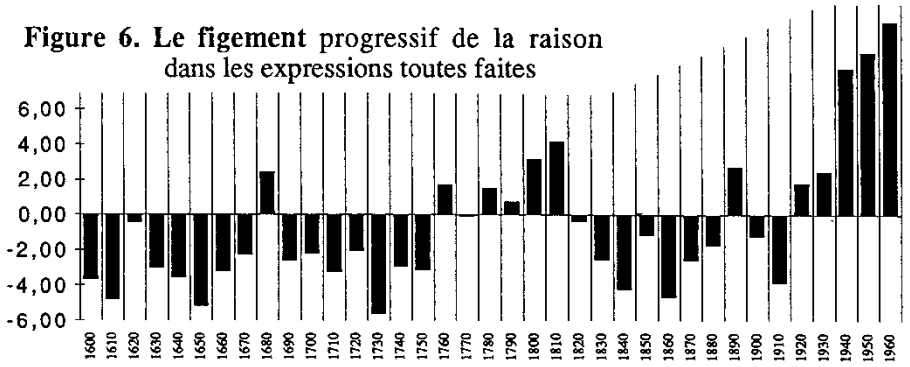
2. « Le vocabulaire religieux dans trois siècles de littérature française », *Actes du Colloque de Jérusalem, Bible et informatique,* Juin 1988, Slatkine, p.147-165.

Graphique 5. Les écrivains face à la raison



La phraséologie ainsi constituée du mot raison n'est pas complète. Mais on a reculé devant la fréquence débordante de "avoir raison", "sans raison", "avec raison". Les 3150 exemples retenus constituent toutefois une base suffisante pour observer une tendance : le figement progressif du mot dans les expressions toutes faites, qui apparaît, non sans à-coups, dans la courbe ci-dessous. La "raison" semble perdre avec le temps son sens plein de faculté pensante, du point de vue sémantique, et, dans l'ordre syntaxique, la pleine liberté d'association du substantif. Il serait intéressant d'observer si ce phénomène de vieillissement ou d'engourdissement du langage - ou de rhumatisme articulaire qui limite la variété des gestes et des combinaisons - s'étend aussi à d'autres mots.

Figure 6. Le figement progressif de la raison dans les expressions toutes faites



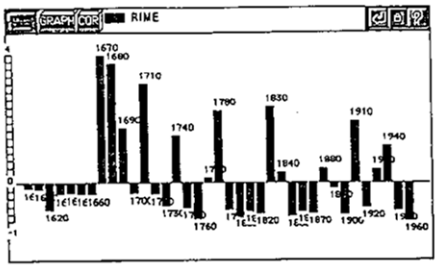
Cependant la phraséologie qui se forme autour du mot "raison" est elle-même un nuage flottant qui se modifie selon les vents de l'histoire. Certaines agglomérations se forment et puis se dissolvent tandis que d'autres naissent ailleurs. Huit des expressions étudiées ont trouvé place dans le graphique 7. Les histogrammes, fondés sur l'écart réduit, y reproduisent le cours du temps, de 1600 à nos jours. Or l'évolution est très différente si l'on considère l'expression "ni rime ni raison" - qui est de moins en moins employée (figure 7a) - et l'expression "à tort ou à raison" - qui l'est de plus en plus (figure 7b). De même on "perd" de moins en moins "la raison" - sans que le monde en devienne plus sage (figure 7c) - et il arrive plus souvent que les événements ou les autres vous "donnent raison" - à moins que ce soit l'inverse (figure 7d).

Notons ici que l'article, présent la plupart du temps dans la première expression, disparaît dans la seconde. C'est en effet la faculté pensante qu'on entend au siècle classique, et l'article souligne l'autonomie du substantif qui est ici une substance, alors que le terme s'est dégradé dans les siècles suivants pour ne plus désigner qu'un simple rapport de causalité, qui fait corps avec le verbe et exclut tout déterminant. Ce n'est pas que l'acception causale soit absente des siècles classiques (et l'expression "par" ou "pour la raison que", graphique 7e, tend à le montrer[3]), mais sa prépondérance y est moins affirmée. Quant à l'acception plus rare de "proportion", le corpus n'en montre guère d'exemple avant le milieu du XVIIIe siècle, qu'il s'agisse de "raison directe" (graphique 7h) ou de "raison inverse" (7f).

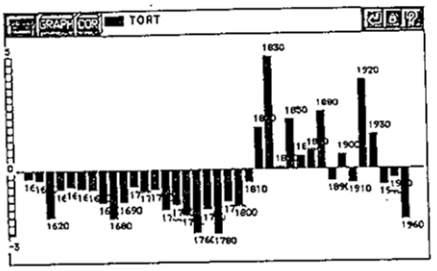
3. Cette expression a un peu vieilli (comme "pour ce que") au bénéfice de "parce que". En revanche l'expression "c'est la raison pour laquelle" est devenue une *scie* du discours politique, au détriment de "c'est pourquoi". C'est que l'orateur a besoin de gagner un peu de temps et la formule développée lui sert de cheville rhétorique.

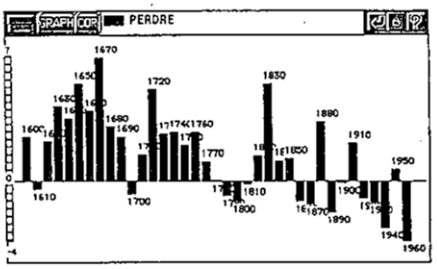
Figure 7. Quelques courbes de la phraséologie de la *raison*

sans rime ni raison

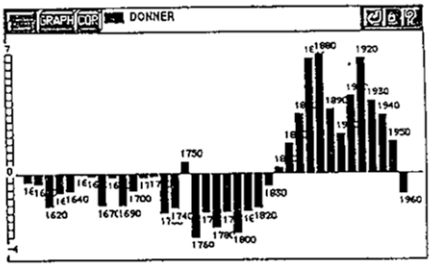


à tort ou à raison



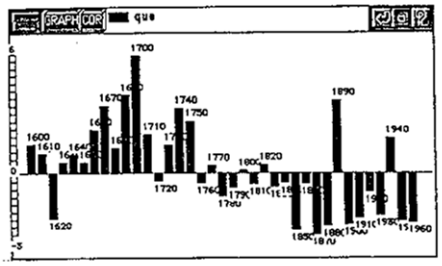


perdre la raison

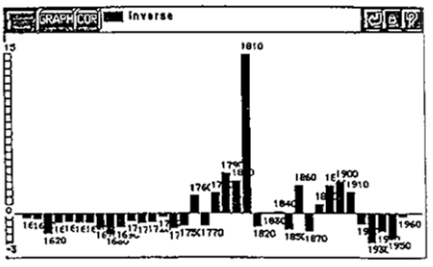


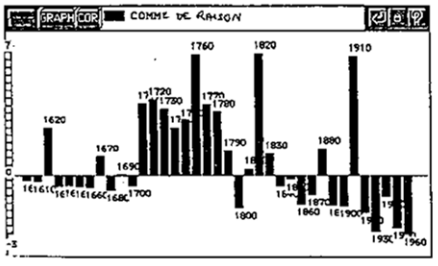
donner raison

pour ou par la raison que ou qu'

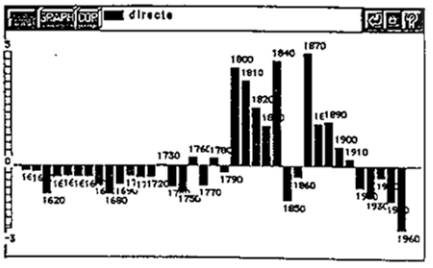


raison inverse





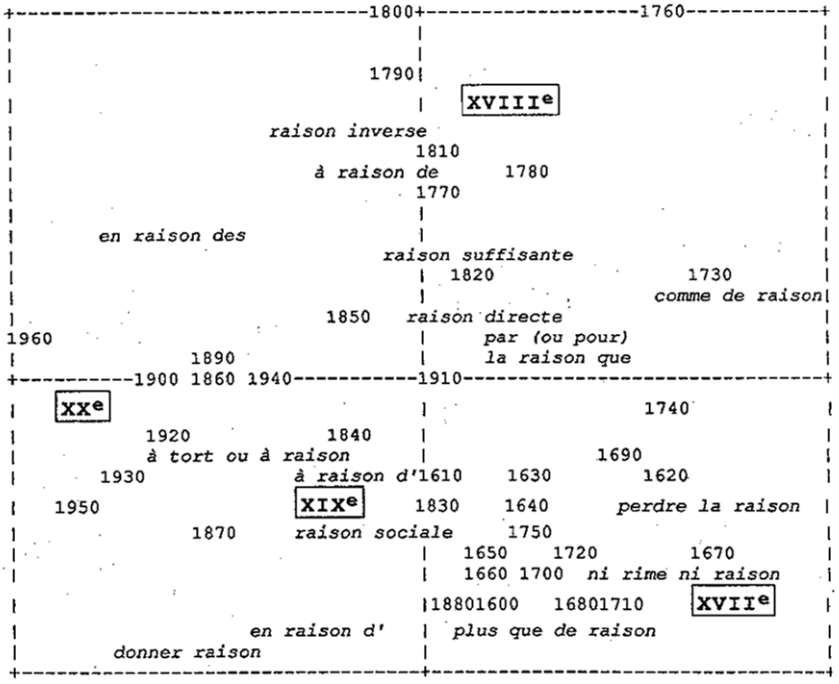
comme de raison



raison directe

Mais pour avoir une idée synthétique des mouvements chronologiques qui affectent la phraséologie de la "raison", il ne suffit pas de multiplier les courbes. On pourrait ajouter aux histogrammes de la figure 7 les 7 autres séries, qui sont pareillement intéressantes. Mais la superposition des points de vue particuliers risque de brouiller le dessin général et il vaut mieux faire appel aux techniques éprouvées de l'analyse factorielle. Cette procédure mathématique produit la figure ci-dessous (figure 8) qui confirme les conclusions fragmentaires inspirées par les courbes. Un mouvement chronologique s'y lit qui s'oriente de la droite vers la gauche. A droite c'est la faculté pensante qui est en faveur, parmi les tranches du XVIIe siècle : 1600, 1610, 1620, 1630, 1640, 1650, 1670, 1680, 1700, 1710, 1720, et qui sous-tend les expressions "perdre la raison", "plus que de raison", "comme de raison", "ni rime ni raison". Les tranches suivantes (de 1760 à 1820) se portent au haut du graphique, à proximité de l'axe des x. C'est l'époque du mouvement encyclopédique, prolongé par la Révolution. La raison y est prise souvent dans son sens mathématique de proportion ("raison inverse", "raison directe", "à raison de"). Enfin sur la gauche du graphique se concentrent les tranches du XIXe siècle, près de l'axe central, et celles du XXe, à l'extrême gauche. L'acception causale de la "raison" y devient majoritaire ("en raison de", "en raison d", "en raison des").

Figure 8. Analyse factorielle de la phraséologie de la *raison*



Nous arrêterons là notre enquête sur le mot "raison", en réservant à un autre article une analyse plus approfondie[4] de la question. Notre but n'était ici que de montrer la puissance de *Frantext* dans la nouvelle version offerte au public en 1991. Nous n'avons guère exploré ici que les fonctions statistiques, qui, il est vrai, ont été considérablement développées dans le nouveau logiciel d'interrogation. Les fonctions documentaires y ont été aussi grandement améliorées et la recherche des contextes d'un mot, d'un thème, d'une figure, y est réalisée à une vitesse surprenante pour un coût fort modéré[5]. En même temps l'amélioration des performances des liaisons télématiques permet de transférer de grosses masses de texte, qu'on peut traiter commodément ensuite avec les ressources et les méthodes dont on dispose. Ainsi avons-nous recueilli plus de 6000 contextes appartenant à des oeuvres expressément philosophiques et contenant le mot "raison". Comme les logiciels de communication permettent généralement la capture des résultats, on peut constituer ainsi un nouveau corpus, très sélectif et très concentré, où par construction chaque phrase contient au moins une fois le mot (ou les mots) cherché. On dispose alors d'un nouveau "texte", formé de contextes mis bout à bout[6]. Un simple logiciel de traitement de texte pourrait à la rigueur permettre certaines recherches à l'intérieur du corpus ainsi réalisé. Mais on obtient des conclusions plus rapides, plus riches et plus sûres, si on utilise un logiciel spécialisé. Ce n'est pas le lieu de développer les fonctionnalités de celui que nous avons réalisé : HYPERBASE. Il nous suffit d'indiquer que *Frantext* non seulement délivre à foison des informations aussi riches, aussi précises et aussi rapides qu'on peut le souhaiter mais qu'il s'articule sans difficulté aux habitudes de travail du chercheur et à sa méthodologie propre. Immense réservoir d'exemples littéraires et linguistiques, il peut jouer, selon le désir de la clientèle, le rôle de la boutique du coin de la rue, ou bien celui de la grande surface, ou même celui de grossiste pour produits demi-finis.

C'est, à notre connaissance, le réseau de distribution le plus souple et le plus puissant jamais vu sur le marché linguistique.

4. A paraître dans les Actes du *Colloque raison/ratio* qui s'est tenu à Rome en janvier 1992, au *Lessico Intellettuale Europeo.*

5. Le coût de la consultation s'abaisse encore lorsqu'on utilise le CD-ROM *Discotext,* qui contient plus de 500 textes principalement du XIXe et pour lequel la liaison télématique n'est plus nécessaire.

6. En réalité, le filtrage retient non seulement le contexte proprement dit (en général une phrase, sauf si la longueur en est démesurée), mais aussi les références (auteur, texte, date, page) qu'un traitement approprié prend en compte.