

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
Publications de l'Institut National de la Langue Française
Directeur : Robert Martin

DICIONNAIRIQUE ET LEXICOGRAPHIE

Collection dirigée par Bernard Quemada

volume 2



FRANTEXT
AUTOUR D'UNE BASE DE
DONNÉES TEXTUELLES

témoignages d'utilisateurs et voies nouvelles

DIDIER ÉRUDITION

UTILISATION DE LA BASE FRANTEXT EN RECHERCHE LITTÉRAIRE

La chose commence à se savoir : la France possède la plus grosse banque de textes (à majorité littéraires) du monde, avec les centaines de millions d'occurrences de FRANTEXT. Et cela suscite immédiatement deux types de réactions contradictoires.

La première, lyrique, non dénuée d'orgueil national, est souvent empreinte de magie : rendez-vous compte, tout ce qu'il sera possible de faire avec un tel trésor ! et de citer les dénombrements multiples, les tris croisés, les recherches de lettres, de formes ou de syntagmes que la machine peut nous procurer en un temps éclair.

La seconde, désenchantée et revenue de tout, déplore que l'informatique, à quelques rares exceptions près, n'ait pas donné les études littéraires nouvelles qu'on attendait d'elle, ni même qu'elle ait suscité une réflexion théorique appropriée¹.

Comme toujours, la bonne démarche se trouve entre ces deux attitudes extrêmes. Elle commence par une analyse du contenu de la base, qui n'est pas seulement, comme dans le projet d'origine, une vaste collecte d'exemples pour les lexicographes. Moyennant quelques très légères modifications (sur lesquelles je reviendrai à la prochaine occasion) tendant à reproduire intégralement le texte saisi, avec ses citations et ses notes, FRANTEXT peut être considérée, en effet, comme la plus importante banque de littérature française. En effet, ladite littérature y est présente dans son déroulement diachronique, avec ses variations génériques, régionales ou nationales (il est de bon ton de parler de francophonie, concept pour le moins ambigu, comme si la littérature française avait jamais représenté la nation française !), ses différents niveaux de diffusion et de réception.

130

Cette masse énorme de textes est accessible directement, en ligne ou, pour une partie, chez soi, sur un CD-ROM. Mais, chaque fois que j'ai assisté à une présentation de FRANTEXT devant des étudiants, et même des enseignants-chercheurs, j'ai été frappé par le fait que ce public, le plus souvent émerveillé, il faut en convenir, ne cherchait qu'à y trouver des attestations de faits de langue et se préoccupait peu de la parole, de la voix spécifique d'un écrivain, de l'originalité même d'un texte, en somme, de ce qui fait la "littérarité", pour reprendre un terme bien français qui nous est revenu de l'Est via Roman Jakobson.

Il est clair que, sous forme de listings ou bien sur écran, on sait tirer parti de la base FRANTEXT dans les domaines de la lexicographie et de la lexicométrie, qui sont toutes deux des disciplines indispensables à l'étude littéraire.

Mais, ne peut-on aller plus loin et se servir de FRANTEXT comme d'un instrument de nature différente, faisant éclater le texte littéraire de base en autant de particules élémentaires que le chercheur peut regrouper à sa guise, en fonction de l'objet de sa recherche ? Depuis le début des années 80, notre URL5* a considéré de son devoir d'initier les jeunes chercheurs à l'utilisation de la Base, en l'état où elle se trouvait alors. Nous entendions leur apprendre à se servir des différents outils informatiques actuellement disponibles pour étayer leurs hypothèses, les vérifier, en formuler de nouvelles ; en somme, introduire de la rigueur là où, le plus souvent, règnent l'à peu près, l'empirisme et la subjectivité. Il me souvient d'avoir écouté avec étonnement le parti que l'on pouvait tirer d'encombrantes "sorties papier" sur les différents aspects de la temporalité dans *Nadja*, la circularité des désignations personnelles dans *Capitale de la douleur*, la couleur dans *Le Feu* d'Henri Barbusse, les marques de l'onirisme dans *Plume* d'Henri Michaux etc. Autant de travaux qui s'appuyaient sur une intuition de départ et qui, à force de va-et-vient du texte aux index et concordances, construisaient une interprétation de l'oeuvre à laquelle il nous était difficile de rien ajouter.

Depuis la mise au point, par Jacques Dendien, du logiciel STELLA, cette époque historique est bien révolue. Elle aura néanmoins appris aux jeunes chercheurs de l'URL à se poser les bonnes questions, à élaborer une stratégie d'investigation adaptée à la fois au texte et aux outils disponibles. C'est pourquoi j'ai souhaité donner la parole à trois d'entre eux. Chacun a choisi une interface différente : la banque FRANTEXT pour Patrick Rebollar, la Minibase de textes "révolutionnaires" pour Liliane Pouzadoux, enfin un produit tout nouveau, qui n'est pas encore sur le marché à l'heure où j'écris, DISCOTEXT 1 pour Michel

* Unité de Recherche en Linguistique

Bernard. Plutôt que de fournir les résultats acquis, comme font, en bonne méthode, les littéraires, ils ont préféré relater les différentes étapes de leur recherche, indiquant par là même leurs difficultés, leurs doutes, leurs précautions, leurs réserves aussi sur certaines procédures, les choix tactiques qu'ils ont dû adopter faute de pouvoir élaborer un programme spécifique. C'est, me semble-t-il, la seule démarche convenable, loin du triomphalisme ou de son contraire, le mutisme ébaubi devant la machine.

Si l'on veut changer quelque peu l'allure des études littéraires, en répudiant d'une part les théories absolutistes plaquées artificiellement sur n'importe quel objet, les remarques impressionnistes à valeur d'impératif catégorique d'autre part, je n'y vois que cette façon, faite de probité, de rigueur et de modestie, avec un esprit critique toujours en éveil.

Aussi ces trois variations sur l'usage de FRANTEXT par des chercheurs en littérature doivent-elles être tenues pour ce qu'elles sont : le journal de bord du laboratoire, avant qu'une version plus conviviale de STELLA ne rende caduques bien des remarques.

Certes, il aurait été possible de faire l'économie de certaines difficultés dont font état nos chercheurs, en attendant la livraison d'un outil parfaitement accessible aux littéraires totalement vierges de toute éducation en informatique. Qui ne voit que ce serait là le meilleur moyen de condamner toute évolution dans ce domaine ? Le mouvement se prouve en marchant : il n'en va pas autrement dans la recherche littéraire assistée par ordinateur.

Henri BÉHAR
Université Paris III
Paris

NOTE

1. Sur ce point, voir les références données et répétées par Alain Vuillemin dans *Informatique et littérature*, Genève, Slatkine, 1990, *passim*.

Chercheurs à l'URL 5 de l'INaLF, nous avons eu le privilège d'utiliser la base FRANTEXT sous plusieurs formes, et dès qu'elles étaient disponibles. Les interfaces dont nous avons pu bénéficier sont à ce jour au nombre de trois : STELLA, la base de textes sur la Révolution et DISCOTEXT 1. On lira ici les relations de trois recherches menées à l'aide de ces trois outils. Utilisateurs privilégiés, nous avons aussi le sentiment d'inaugurer des voies et des procédures qui devraient être celles de toute la recherche littéraire des années qui viennent. En ce sens, nos perplexités et nos tâtonnements sont à la mesure d'un tel changement qui, jetant un pont entre esprit de géométrie et esprit de finesse, fera entrer l'informatique, son herméneutique et ses modes de démonstration dans le champ clos de la recherche littéraire.

I - UN NOUVEL UNIVERS DE RECHERCHES : FRANTEXT

La critique de l'outil

L'interrogation que nous utilisons ici en exemple consiste à rechercher un choix d'adjectifs en "esque" (de l'italien *esco* : à la manière de), moyennement et peu courants contenus dans des textes littéraires, de toutes les époques possibles.

Nous ne traitons donc pas de FRANTEXT dans son ensemble, bien que des parties de cette procédure soient répétées ou similaires quelle que soit l'interrogation.

La connexion sur FRANTEXT est en soi un problème que les néophytes ont du mal à résoudre : être abonné (ou se rendre dans une bibliothèque universitaire), connaître le numéro de code, ne pas confondre le code d'identification et le mot de passe, connaître le nom du service où se trouve FRANTEXT, sont les premières opérations ou connaissances. Il s'ensuit que le service n'est accessible qu'à des chercheurs avertis, ou, pour les autres, par l'utilisation des services bibliothécaires (d'ailleurs fort utiles, puisque ces personnes spécialement formées à FRANTEXT aident aussi à "formuler" une question).

L'interrogation nécessite en effet que l'on connaisse au préalable la structure et la syntaxe de FRANTEXT¹ (ce qui peut être fait en se reportant au manuel). Mais aussi que l'on adapte les recherches que l'on veut faire : il faut pour cela fractionner le problème à résoudre en une série ordonnée de questions simplement formulées (codifiées). Or - et FRANTEXT n'est pas en cause -, il n'existe pas de formation à la *formulation des problèmes* à l'intérieur des études de lettres, bien que l'on sache maintenant combien sont nécessaires les cognitiens, dont c'est le métier, pour le moins, d'adapter questionnement et structures de réponse existantes ou à créer.

En ce qui nous concerne, nous avons utilisé la commande qui permet "l'extraction" du vocabulaire désiré. On indique le nom voulu pour la liste de mots, et les conditions : critères de filtrage (ici, "%esqu(e/es)", où "%" demande une troncature quelconque et "(e/es)" deux finales possibles autour du pivot "esqu"). La touche *entrée* déclenche la recherche, qui donne le nombre de formes sélectionnées ; mais qui ne propose pas de les voir.

Notre liste doit contenir des formes en "esque" à supprimer qui ne sont pas des adjectifs : ainsi du substantif "fresque", du nom propre "Fiesque" (le comte de), ou du latin "compedesque" ; et aussi des adjectifs trop fréquents qui ne nous intéressent pas ici car leur étude demanderait trop de temps ("grotesque", "romanesque", etc.).

Il existe cependant des formes dont on ne peut dire si elles sont adjectivales ou autres tant que l'on ne les lit pas au sein de leur contexte, opération que l'on appelle la désambiguïsation (ou tant que le système ne précisera pas la nature des mots dans les listes sélectionnées - ce qui pose le problème épineux de la reconnaissance grammaticale automatique puisqu'il n'est pas question que des êtres humains encodent un par un les mots des textes !...) : "arabesque" peut être adjectif ou substantif, "magiquesque" pourrait *a priori* être également un adjectif (on y reviendra plus bas), tandis que l'écartement de "compedesque" présuppose une connaissance du latin.

Nous enlevons donc le moins de mots possible de la liste, ce qui entraînera que nous aurons plus de références et de citations à récolter, à lire pour désambiguïser.

Préalablement (bien que l'opération soit tout à fait indépendante), nous avons utilisé la commande qui permet de sélectionner le corpus à l'intérieur duquel doit se faire une recherche : par date, par genre, nom d'auteur ou titre d'oeuvre (et croisement possible de ces conditions)². Dans notre cas, nous réduisons le corpus aux genres dits "littéraires", ceux où *a priori* l'emploi de mots nouveaux ou rares est le plus fréquent.

Enfin la commande de recherche dans les textes "interroge" le corpus sur les formes indiquées comme "composants".

Il faut ensuite choisir la "taille" des citations. À certaines recherches suffisent les concordances d'une ligne ; dans notre cas, le besoin de comprendre tout à fait le sens du mot nous pousse à choisir 2 phrases de chaque côté. Après la validation de ce choix, commence le défilement des citations correspondant à la demande. Ces résultats qui défilent à l'écran, peuvent être demandés à FRANTEXT sous forme imprimée (tout comme pouvait l'être l'ensemble du questionnement, pour le chercheur pressé ou seulement occupé des résultats). Ils

peuvent également être directement enregistrés sur le poste de travail au moyen d'une "capture" que permet le logiciel de connexion de l'utilisateur ; bien que cette fonction ne dépende pas de FRANTEXT, on peut regretter que le système n'invite pas à activer une éventuelle capture avant le défilement des résultats.

L'état des textes, comme on le voit sur l'exemple suivant, est lui aussi problématique puisque sur certains matériels tous les accents sont à côté des lettres, impliquant un travail de remise en état avant une éventuelle citation dans un ouvrage ou dans un cours³. S'y ajoute, dans cet exemple, une faute de frappe, puisque la ligne est : «des vieillards magiques que l'ancien monde tenait».

Notons que la création d'un tel adjectif, selon le mode de formation en "esque", eût été viable ; seul le contexte, où «l'ancien monde [...]» devenait un groupe nominal parataxique incompréhensible, permet de douter et de recourir au texte sur papier...

*C'est que l'abaissement ou la disparition des autres
puissances a fait du prêtre, pourtant si
étroitement mêlé en apparence à la vie sociale, un
être plus particulier, plus inclassable qu'aucun
des vieillards MAGIQUESQUE l'ancien monde tenait
enfermés au fond des temples, ainsi que des animaux
sacrés, dans la seule familiarité des dieux. D'autant
plus particulier, plus inclassable qu'il ne se
reconnaît pas pour tel, presque toujours dupe
d'apparences grossières, l'ironie des uns, la
déférence servile des autres. Mais, à mesure que la
contradiction, d'ailleurs moins religieuse que
politique et dont s'est trop longtemps nourri leur
orgueil, se résout peu à peu en une sorte*
page 1494/K674/BERNANOS. G/MONSIEUR OUINE/1943

L'édition de référence n'est pas précisée lors de l'affichage des résultats. Il faudrait la demander à part ou la chercher dans le catalogue papier. Souvent pourtant, cette édition n'est plus courante, et retrouver la référence prend un temps non-négligeable. Du côté de FRANTEXT le problème est quasiment insoluble car il paraît illusoire d'indiquer une édition qui puisse perdurer, tout comme il semble lourd à gérer de réactualiser sans cesse les textes en fonction des éditions dites "de référence".

En conclusion de notre critique du mode opératoire, disons que l'accès et l'interrogation de FRANTEXT accomplissent parmi les clients potentiels une sélection qui lui est défavorable. De plus, l'état des accents dans les textes, les possibilités de coquilles et les problèmes de référence bibliographique, n'en font pas l'outil fiable et pratique dont les chercheurs ont besoin. Néanmoins, comme nous le montrons ensuite par un résumé de l'analyse à laquelle nous a mené cette interrogation, FRANTEXT est un outil d'une absolue nécessité, utilisable de façon efficace et rapide pour peu que l'on passe ces chicanes. Il est donc souhaitable que le système soit adapté à la convivialité et à la sécurité auxquelles d'autres logiciels nous ont habitués.

Sur la pratique adjectivesque

Le relevé qui suit (p. 144) provient, selon les références que l'on voit apposées aux occurrences, du Robert Électronique (R)⁴, de l'interrogation de FRANTEXT (F), des deux (sans indice), ou de l'oeuvre de Claude Simon, sur laquelle nous travaillons par ailleurs, lorsque la néologie est de lui (S).

Quelques commentaires sur les tableaux (p. 145) s'imposent :

1. FRANTEXT recèle plus d'occurrences que le ROBERT Électronique : c'est que malgré d'importants efforts pour élargir le champ encyclopédique (citations nombreuses, références bibliographiques, hypertextualité), le dictionnaire ne parvient pas à rendre compte des hapax et néologies des écrivains.
2. La portion commune est inférieure à la moitié : constat logique, puisque le ROBERT recense principalement les plus usuels. Mais en intégrant d'autres textes, FRANTEXT devrait faire croître cette proportion jusqu'à un taux satisfaisant (alors croîtra également le nombre des hapax...).
3. Chez Simon : en l'absence de données complètes sur la présence de ces adjectifs chez d'autres auteurs (recherche possible avec FRANTEXT), aucune conclusion n'est à tirer d'un point de vue statistique. Chaque lecteur pourra s'interroger sur différents auteurs en les lisant ; nous pensons que les écrivains utilisant plus de 20 adjectifs en "esque" sont rares. Plus rares encore sont ceux qui en auraient forgé plus de 5. Il semble pourtant que la deuxième moitié du XIX^e siècle soit très productive dans ce domaine (peut-être à cause de la vogue des journaux satiriques) : il faudrait donc effectuer une étude plus précise de cette période littéraire et la mettre en perspective d'une étude historique globale.
4. L'origine de la formation adjectivale est bien loin. En effet moins de 10 % des mots renvoient au nom d'un peintre — c'est-à-dire autant qu'à des personnages fictifs, ce qui est étymologiquement absurde puisque ces derniers n'ont pas existé (mais ils furent si bien décrits !...).

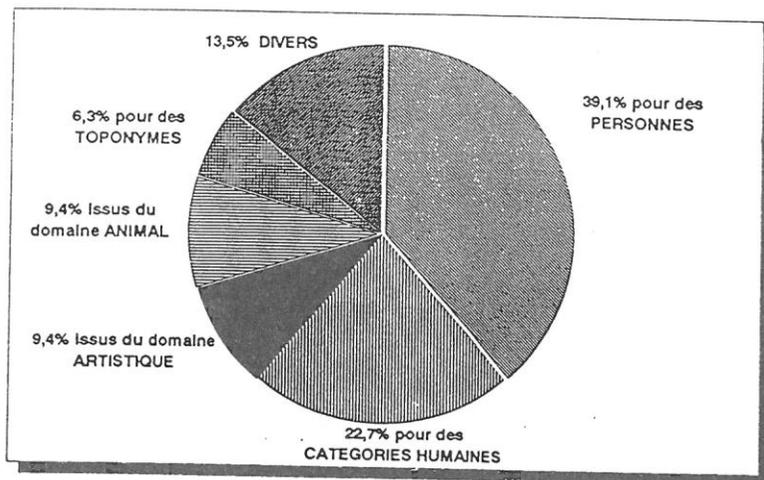
Par contre, le référent humain reste le plus important puisqu'il compte presque pour 72 % de la liste.

- abracadabrantisque (R)
 alhambresque (F)
 animalesque (R)
 antipittoresque (F)
 arabesque
 aristophanesque
 atridesque (S)
 babelesque
 bacchanalesque (F)
 banvillesque (F)
 barbaresque
 bouffresque (F)
 boy-scoutesque (S)
 burlesque
 cambronnesque (F)
 caméléonesque
 cannibalesque
 canularisque (R)
 caravagesque (R)
 carbonaresque (R)
 cardinalesque (F)
 caricaturesque
 carnavalesque
 cassandresque (S)
 cauchemaresque
 chaplinesque (S)
 chardinesque (F)
 charenton(n)esque
 charivaresque (R)
 charlatanesque
 chatesque (F)
 chatnoiresque (F)
 chevaleresque
 churrigueresque (R)
 clownesque
 courtelinesque
 courtisanesque
 crocodilesque (R)
 cyclonesque (S)
 danaesque (S)
 dandyesque (F)
 dantesque
 diafoiresque (F)
 donjuanesque (S)
- don(-)quichottesque
 edmondschresque (F)
 éléphanthesques
 errantesque (F)
 faubourg-St-germanesque (F)
 faunesque
 festivalesque (R)
 feuilletonesque (R)
 figaresque
 fourbesque (R)
 fourmillesque (S)
 funambulesque
 gargantuesque
 gentilhommesque (R)
 gigantesque
 giottesque (R)
 godwinesque (F)
 goyesque (S)
 grandguignolesque (R)
 grenouillesque (F)
 grotesque
 guignolesque
 héliogabalesque (F)
 hippopotamesque
 hoffmannesque
 ingresque
 jargonesque (R)
 jordanesque
 lalesque (F)
 léonardesque
 livresque
 madrigalesque (R)
 marinesque
 mastodontesque (R)
 m[au/o]resque
 méropesque (F)
 michelang(el)esque
 mirliton(n)esque
 molièresque
 ossianesque (R)
 paganinesque (F)
 parmegianesque
- parmesanesque (F)
 paysagesque (F)
 paysan(n)esque (F)
 pédantesque
 pétrarchesque (F)
 pharaon(n)esque
 phidiesque (R)
 picaresque
 pittoresque
 plateresque
 polichinellesque (F)
 porphyresque (S)
 poussinesque
 prud(')hom(m)esque
 putanesque (R)
 rapha[é/è]lesque
 rembranesque
 rocambolesque
 romanesque
 sandesque (F)
 sardanapalesque
 sauvagesque (F)
 schumannesque (F)
 scipionesque (F)
 sheridanesque (F)
 simiesque
 soldatesque
 somnambulesque
 sultanesque (F)
 tabarinesque (F)
 tailoresque (F)
 tauresque (F)
 tintamar((R))esque
 titanesque
 titian(n)esque
 troubadouresque
 truquesque (F)
 tudesque
 turlupinesque (R)
 ubuesque (R)
 vaticanesque (R)
 vaudevillesque

Source	total	%	exclusif	%
liste fusionnée	128	100		
FRANTEXT	95	74,8	37	29,1
ROBERT	81	63,8	23	18,1
communs aux deux	58	45,7		
chez SIMON	34	27	néologismes : 10	7,2

Origine	sous-catégorie	nombre	Pourcentage
PERSONNES		50	39,1
	modernes	19	15
	peintres	12	9,4
	fictifs ¹	12	9,4
CAT. HUMAINES ²		29	22,7
DOM. ARTISTIQUE ³		12	9,4
DOM. ANIMAL		12	9,4
TOPONYMES		8	6,3
DIVERS ⁴		18	13,5

1. personnages de théâtre ou de roman (Diafoirus, Gargantua, etc.)
2. désigne des groupes liés à des qualités ou un état d'existence (sauvage, boy-scout, clown, pédant, etc.)
3. du vocabulaire principalement de la peinture ou du théâtre (vaudeville, marine, etc.)
4. inclassables, dont des néologismes au sens flou (porphyre, cyclone, bouffre, etc.)



5. Certains de ces adjectifs sont aujourd'hui incompréhensibles, y compris en lisant le texte dont ils proviennent : c'est qu'ils sont liés à la connaissance qu'avait un public particulier à un moment donné (il faudrait donc effectuer une enquête littéraire, journalistique, voire historique ou sociale pour comprendre pleinement le sens de "chatnoiresque", du Chat Noir, cabaret de Paris de 1881 à 1897 où se réunissaient poètes, chansonniers, dessinateurs, etc., dont Alphonse Allais ; ou de "charivaresque", du *Charivari*, journal satirique de 1832 à 1860)

Pour beaucoup d'entre eux, une bonne culture générale, picturale, historique et mythologique est malgré tout nécessaire et suffisante.

Par contre l'invention au présent ne demandera quasiment aucune recherche, ce qui est en accord avec le principe de leur formation⁵ : on comprend d'emblée "une ville beyrouthesque", "un but papinesque"... (entendus à la télévision).

La compréhension du mot, si elle ne nécessite pas toujours une connaissance approfondie du champ culturel en question, présuppose, au minimum, le recours à un référent dont le nom est reconnu en ôtant le suffixe "esque" (et la possibilité de consulter un dictionnaire ou une encyclopédie).

L'usage courant de certains de ces mots (*titanesque*, *romanesque*) nous a déshabitués de les comprendre *a priori* comme des comparaisons (en accord avec leur origine : *comme un titan*, *comme dans un roman*) elle fait choisir à chacun, dans le procès comparatif, certains attributs seulement du comparant (grand, puissant, fort, résistant, etc. comme un titan ; fabuleux, exalté, chimérique... comme un personnage de roman) — cette cible n'étant pas le résidu culturel des référents (ce qui serait une contingence), mais bien un choix du locuteur (même s'il lui est imposé par l'usage appris de l'adjectif en question...) : un titan est encore autre chose *aujourd'hui* qu'un géant fort, un roman est encore autre chose *aujourd'hui* qu'une suite de péripéties scabreuses.

Par conséquent, l'usage de ces mots sous-entend un accord tacite de la communauté linguistique, mis en place lors de l'apprentissage scolaire (ou autre), et qui ne laisse subsister que quelques traits distinctifs dans le champ du comparant.

En opérant cette réduction, et du fait de l'absence d'un opérateur comparatif habituel (*comme*, *tel*, *ainsi que*, etc.) ce qui était étymologiquement une comparaison travaille la langue de fait comme une métaphore : *goyesque* m'impose à l'esprit un schème caricatural où la figure est torturée, où les couleurs sont violentes, sombres et profondes à la fois (mais pas, à supposer qu'elle existe, la taxinomie intellectuelle des traits distinctifs de l'oeuvre de

Goya, *in extenso* !) — c'est MA réception du mot et, à ce titre, elle n'est dans aucun dictionnaire... (D'où la bizarrerie du concept institutionnel de "langue"...)

Or, à force d'être usités — et c'est hélas le devenir constaté de bien des métaphores —, ces mots perdent le lien avec leur référent culturel, tout en conservant les traits ciblés : ils deviennent des clichés. Ainsi *gigantesque*, ou *grotesque* renvoient respectivement à un schème de /grandeur/ ou de /bizarrerie/extravagance/, sans plus avoir de lien avec les géants ou les grottes — à moins qu'un auteur ne veuille réveiller la métaphore qui sommeille dans tout cliché, par un environnement textuel qui la ramène à son étymon :

«l'état (temps, espace, froid) où devait être le monde à l'époque des cavernes, des mammouths, de bisons, et autres bêtes gigantesques chassées par des hommes gigantesques pour prendre leurs fourrures, boire leur sang chaud, au sein de gigantesques et inépuisables forêts.» (Claude Simon, *Les Géorgiques*, p. 119)

L'étape suivante du devenir cliché a été, pour certains mots, qu'ils ne purent plus qualifier qu'eux-mêmes. Ils se sont substantivés : un dessin *arabesque* est ainsi devenu une *arabesque* ; ou bien un parler spécifique, brusque et direct, a édifié une catégorie humaine, qui est maintenant indépendante, la *soldatesque*.

Ce glissement métaphorique, comme par ailleurs cette substantivation, s'expliquent par la coïncidence catalytique d'un facteur grammatical et d'un facteur sémantique.

Grammaticalement. C'est la maniabilité du procédé qui facilite son usage : la comparaison est discrète, maniable, de moins en moins sensible au fur et à mesure que chaque locuteur la pratique, puisqu'elle tient dans un seul mot — la copule immigrée et agglutinée ("esque") s'étant fait oublier. De plus, l'emploi adjectival a lieu presque exclusivement en position d'épithète (de sorte qu'un substantif générique et redondant -sémantiquement faible- a vite fait de disparaître dans une ellipse !...)⁶

Sémantiquement. Sachant que l'on sait que tout le monde sait que "chaplinesque" renvoie à Chaplin-acteur, on pourra l'utiliser pour qualifier un gag (champ sémantique du "jeu" cinématographique ou autre), puis une personne pour l'ensemble de son comportement (glissement vers le champ individuel), puis un simple geste typé, même très bref... Plus le mot "vieillit", est usité, plus on y a donc recours à des fins métaphoriques⁷, éloignées de son champ d'action originel.

De plus, un facteur pragmatique non négligeable renforce la découverte que l'on peut faire de métaphores et d'ellipses⁸. La position d'épithète est plus souvent le fait du texte écrit, voire littéraire, puisqu'en rassemblant un nombre de sèmes important au sein d'un groupe nominal, elle enrichit et complique la compréhension, crée des liens plus complexes avec les autres parties de la phrase (et du texte dans son ensemble, éventuellement). Il ne sera donc pas étonnant, venant d'un écrivain, d'y trouver des ellipses, ou que l'emploi du mot glisse vers une métaphore que seul le contexte rendra sensible.

On remarquera, pour clore cette brève étude, que la formation en "esque", non contente de se diversifier par la métaphore ou l'ellipse, est aussi capable d'assimiler des mots-valise (Jarry fabrique : bouffre(esque) = gouffre + bouffon ?) et des mots étrangers (Queneau emploie *tailoresque* venant de "taylor", mais aussi, plus intertextuellement et humoristiquement comme c'est toujours le cas chez Queneau, de "my taylor is rich" !)

Patrick REBOLLAR
CNRS-INaLF
Paris

NOTES

1. L'une et l'autre bien améliorées dans la version 2.
2. Des menus appropriés facilitent cette sélection.
3. Ces difficultés sont grandement allégées dans la nouvelle formule.
4. Version informatique du *Grand Robert*, disponible sur CD-ROM, consultable dans certaines bibliothèques universitaires, ou achetables pour utilisation avec un traitement de texte. Marque déposée.
5. Un récent colloque sur l'humour et les arts du rire s'est intitulé "humoresques", et a donné lieu à une publication des actes en deux volumes sous le même titre.
6. Toutes proportions gardées on pourrait comparer cette ellipse à l'élision dont est victime, dans l'évolution de la langue, une syllabe faible qui voisine une syllabe forte : la vigueur sémantique de l'adjectif affaiblit le substantif au point que le locuteur l'enlève - ne le *voyant* plus.

7. En cela "caricaturesque" constitue un pléonisme, mais aussi une *mise en abyme énonciative*, selon la définition de L. Dällenbach.
8. Ce qui revient à devoir définir un autre corpus pour une autre étude, incluant par exemple du discours oral (où l'on rencontrera plus d'adjectifs en esque en position d'attribut) et des textes journalistiques (avec plus de néologismes contemporains, mais aussi plus d'emplois de mots figés), par exemple.

II - UTILISATION D'UN EXTRAIT DE FRANTEXT:

Un corpus de textes relatifs à la Révolution Française ou "minibase révolutionnaire"

L'objet de ce travail était une présentation de ce corpus FRANTEXT et de ses fonctionnalités dans le cadre du séminaire "Études Littéraires et Lexicales Assistées par Ordinateur", animé par Henri Béhar et Danielle Bonaud-Lamotte à l'Université Paris III.

Cette minibase, dite "minibase révolutionnaire", est constituée de 56 textes (1755-1830) choisis par l'INaLF pour des interrogations interactives dans le cadre des expositions du Bicentenaire. Elle ne dispose que de deux fonctions FRANTEXT : *biblio* et *cherche*.

Il s'agit donc d'une version simplifiée de FRANTEXT, version de démonstration sur disquettes destinée à un public de visiteurs.

Pour rester dans les limites des conditions d'utilisation, notre hypothèse d'interrogation a été la suivante : à partir des choix offerts par la fonction *biblio* pour la sélection des textes, faire émerger les mots ou concepts-clés de "l'imaginaire révolutionnaire" d'un visiteur d'une exposition du Bicentenaire en utilisant la fonction *cherche* qui affiche les contextes d'utilisation et comptabilise le nombre d'occurrences d'un mot comme par exemple "esclavage" dans *L'Esprit des Loix* (64 occ.) et *Le Contrat Social* (14 occ.) ou d'une expression comme "progrès des Lumières" dans *Esquisse d'un Tableau Historique...* de Condorcet (12 occ.).

_____ Solution 1 _____

Il cite

l'esclavage en exemple.

ROUSSEAU J.-J. / DU CONTRAT SOCIAL / 1762/p. 176

_____ Solution 2 _____

Aristote avant eux tous avoit

dit aussi que les hommes ne sont point naturellement

égaux, mais que les uns naissent pour l'esclavage et

les autres pour la domination

ROUSSEAU J.-J. / DU CONTRAT SOCIAL / 1762/p. 177

_____ Solution 3 _____

Tout homme né dans l'esclavage naît pour

l'esclavage, rien n'est plus certain.

ROUSSEAU J.-J. / DU CONTRAT SOCIAL / 1762/p. 177

_____ Solution 4 _____

Tout homme né dans l'esclavage naît pour

l'esclavage, rien n'est plus certain

ROUSSEAU J.-J. / DU CONTRAT SOCIAL / 1762/p. 177

_____ Solution 5 _____

De l'esclavage

Puisqu'aucun homme n'a une autorité naturelle sur son

semblable, et puisque la force ne produit aucun droit,

restent donc les conventions pour base de toute

autorité légitime parmi les hommes

ROUSSEAU J.-J. / DU CONTRAT SOCIAL / 1762/ LIVRE 1 CHAPITRE 3/

p.181

On notera qu'une même citation est imprimée et comptabilisée autant de fois qu'elle comporte d'occurrences du mot cherché. Une lecture attentive des sorties en contexte est donc nécessaire pour un comptage précis.

Les composants 1,2 3 de cette même fonction *cherche* permettent les recherches de cooccurrences comme celles de "liberté"- "égalité"- "fraternité" sur un ensemble de textes et/ou sur une période sélectionnée.

L'utilisation de la fonction *biblio* pour la constitution des corpus de travail est simple. Dans cette version de démonstration la sélection des textes est accessible directement au moyen des touches de fonction. Le principe de la constitution d'un corpus étant le même que dans FRANTEXT on peut obtenir à partir de la liste chronologique des oeuvres affichées à l'écran une sélection d'un seul texte, de tous les textes d'un même auteur, de tous les textes d'une date ou d'une période donnée. À tout moment de la sélection on peut modifier, ajouter ou supprimer un texte ou un ensemble de textes en fonction des hypothèses de l'interrogation.

Une fonction **trier les textes** (ordre alphabétique des auteurs ou des oeuvres) permet de procéder à des regroupements par genre et/ou par thème : par exemple l'ordre alphabétique des titres fait apparaître un classement par genre pour les oeuvres d'idées, nombreuses dans un corpus du 18e comme

Considérations sur

Lettres de

Essai sur

L'interrogation par la fonction **cherche** suppose l'établissement d'une liste de références, de concepts-clés, mots ou expressions, que l'on appellera "Liste thématique du 18e siècle".

Termes choisis dans les index des ouvrages et manuels d'histoire littéraires du 18e et classés ici en sous-groupes comme représentatifs par période d'un "imaginaire du 18e siècle".

Groupe 1 : "Thématique des Lumières"

Bonheur	Nature	Sensibilité	Raison	Vertu	Progrès	Progrès des
Lumières	Commerce	Science	Histoire	Utopie	Démocratie	Optimisme
Religion	Vérité	Tolérance	Esclavage	Contrat	Social.	

Groupe 2 : "Thématique révolutionnaire"

Liberté	Égalité	Fraternité	Justice	Droits de l'Homme	Citoyen(s)
Révolution	Constitution	Tyran	Despote	Ennemis de la Liberté.	

Recherche des cooccurrences de Liberté-Égalité-Fraternité, Liberté-Esclavage, Tyran-Despote, Peuple-Nation, Religion-Liberté

Utilisation de la fonction *biblio*

Examen du corpus de travail :

La version dont nous disposons groupait 56 textes représentant tous les genres³ qui, se présentant dans l'ordre chronologique, pouvaient se prêter à un premier groupement de type périodique :

- les textes fondateurs : Montesquieu, Rousseau ;
- les textes des auteurs-acteurs de la Révolution : Sieyes, Marat, Robespierre, Condorcet, et de la période révolutionnaire : Laya, *L'Ami des Loix*, Volney, *Les Ruines*, Chateaubriand, *Essai sur les Révolutions* ;
- d'autres textes concernant la Révolution écrits ou publiés postérieurement à 1815.

D'autres groupements par genre et/ou par thème étaient possibles : le roman de l'émigration, les correspondances, la poésie de Chénier avec *L'Invention* et *L'Amérique*, textes moins connus des non-spécialistes, la comédie de Laya ont fait l'objet d'interrogations ponctuelles. De même la présence de textes scientifiques des savants de la Révolution : Lavoisier, Monge, Geoffroy pouvaient se prêter à des recherches sur la langue.

Solution 1

le charbon au lieu de s'y consommer paisiblement comme dans l'air ordinaire, y brûloit avec flamme et une sorte de décrépitation, à la manière du phosphore, et avec une vivacité de lumière que les yeux avoient peine à supporter.

LAVOISIER / TRAITÉ ÉLÉMENTAIRE DE CHIMIE / 1789 / p. 3

Solution 2

de la composition du gaz oxygène par le soufre, le phosphore et le charbon, et de la formation des acides en général

LAVOISIER / TRAITÉ ÉLÉMENTAIRE DE CHIMIE / 1789 / p. 57

Solution 3

Le charbon, que tout jusqu'à présent porte à faire regarder comme une substance combustible simple, a également la propriété de décomposer le gaz oxygène et d'enlever sa base au calorique : mais l'acide qui résulte de cette combustion ne se condense pas au degré de pression et de température dans lequel nous vivons ;

LAVOISIER / TRAITÉ ÉLÉMENTAIRE DE CHIMIE / 1789 / p. 67

Solution 4

On peut opérer la combustion du charbon, comme celle du phosphore, sous une cloche de verre A, *planche IV, figure 3*, remplie de gaz oxygène, et renversée dans du mercure : mais comme la chaleur d'un fer chaud et même rouge, ne suffiroit pas pour l'allumer, on ajoute par-dessus le charbon, un petit fragment d'amadou et un petit atome de phosphore.

LAVOISIER / TRAITÉ ÉLÉMENTAIRE DE CHIMIE / 1789 / p. 67

Sous-corpus choisis pour la manipulation en fonction de l'hypothèse de travail :

- *Sur une oeuvre* : Condorcet et la "thématique des Lumières"

Esquisse d'un tableau historique des progrès de l'esprit humain (texte indexé par ailleurs dans le cadre de la BDHL [Banque de Données d'Histoire Littéraire (M. Béhar, Paris III)] et dont il était intéressant de comparer concepts-clés et indexation thématique),

- Sur des ensembles de textes correspondant au découpage périodique proposé plus haut pour la "thématique révolutionnaire,

- Mais pas de recherches possibles sur l'ensemble du corpus pour des concepts-clés très représentés comme "liberté". Le programme ne le permettait pas (limitation en mémoire du programme, d'ordre technique). Par contre on a obtenu 138 occurrences de "démocratie" sur l'ensemble des textes.

Utilisation de la fonction *cherche*

À condition de définir des sous-corpus "maniabiles", on obtient des résultats rapides de sorties en contextes (mots ou expressions) et le nombre d'occurrences (ou solutions trouvées).

À titre d'exemples, quelques extraits des résultats obtenus en nous limitant à quelques concepts-clés de la liste thématique. Ces résultats dans le cadre d'une manipulation rapide constituent la première étape de la recherche et fournissent avec les graphiques obtenus des éléments de comparaison. (Voir tableau, p. 154).

Remarques d'utilisation

L'utilisation des composants 1, 2, 3 n'est pas précisée et peut poser des problèmes de manipulation pour la recherche des cooccurrences. Pour la lisibilité des résultats il est conseillé de faire cette recherche texte par texte et non sur un ensemble. (Voir en *Annexe* des exemples de sorties des cooccurrences de "religion" et "liberté" dans les textes de Chateaubriand et Madame de Staël).

Ces premiers résultats obtenus, après analyse des contextes, peuvent servir d'indicateurs d'utilisation des concepts-clés qui caractérisent le 18e siècle. Des résultats exhaustifs pourraient être utilisés pour un essai de classement thématique (évolution des concepts-clés du vocabulaire des Lumières au vocabulaire révolutionnaire).

En résumé, l'obtention des résultats est rapide mais pour pouvoir apprécier et comparer ceux-ci, il faudrait disposer sur le même produit d'un tableur-grapheur permettant leur visualisation immédiate (ce que permet FRANTEXT dans sa deuxième version, avec la commande *distributions de fréquences*).

Cette version volontairement réduite de FRANTEXT ne comportant pas de fonction *index* ni de fonction *fréquences*, il n'y a pas de mesures possibles, simplement des observations de contextes et des comptages qui permettent dans un premier temps de préciser des acceptions ou d'infléchir les hypothèses

	Montesq. <i>Exp. Loix</i>	Rousseau <i>Contr. Soc.</i>	Sieyès <i>Tiers-Ét.</i>	Marat <i>Pamphlets</i>	Laya <i>Ami Loix</i>	Robesp. <i>Discours</i>	Condorcet <i>Esquisse</i>	Chateaubr. <i>Essai sur</i>	Mme de Staël <i>Considérat.</i>	St Mart. <i>L'H. de D.</i>	Vohney <i>Ruines</i>
<i>liberté</i>	230	56	22	64	12	286	84	52	222	15	9
<i>égalité</i>	40	24	13	3	2	33	39	1	9	1	17
<i>fraternité</i>	0	0	0	2	0	4	5	2		1	1
<i>justice</i>	36	15	1		6	49	17				
<i>progrès des lumières</i>			1			1	12		5		
<i>religion et liberté</i>								10	15		

de départ en dégageant des tendances. On peut cependant utiliser *les partitions* FRANTEXT dans la saisie des textes pour observer *les spécificités* par partie. Dans la minibase les textes de :

Montesquieu, *De L'Esprit des Loix*, t. 1.2.3. - 3 partitions

Mirabeau, *Lettres écrites du Donjon* - 3 partitions

Chateaubriand, *Essai sur les Révolutions* - 2 partitions

Mme de Staël, *Considérations sur la Révolution Française* - 4 partitions

se prêtaient à ce type de comparaisons.

L'étape suivante de la recherche consistera donc dans l'utilisation de FRANTEXT avec toutes ses fonctions et ses améliorations de consultation. En tant que version de démonstration cette minibase provoquait donc avant tout la curiosité et l'intérêt pour FRANTEXT en permettant d'accéder rapidement et simplement à un corpus étendu sur un thème donné..

Liliane POUZADOUX
CNRS-INaLF
Paris

ANNEXES

1. Sorties en contexte de "poète" dans *L'Invention et L'Amérique* de A. CHÉNIER

Solution 1

Il faut dans cet ouvrage, soit quand le poète parlera, soit par la bouche des personnages, soit dans les discours prophétiques des êtres surnaturels, décrire de côte en côte absolument toute la géographie du globe aujourd'hui connue.
CHÉNIER A ./ L'AMÉRIQUE / 1794 / p. 83

Solution 2

le poète Alphonse, à la fin d'un repas nocturne en plein air, prié de chanter, chantera un morceau astronomique : quelles étoiles conduisirent Christophe Colomb.
CHÉNIER A ./ L'AMÉRIQUE / 1794 / p. 104

Solution 3

Il ne sera pas mal que le poète raconte allégoriquement quelque part l'histoire physique du tonnerre.
CHÉNIER A ./ L'AMÉRIQUE / 1794 / p. 110

Solution 4

Mais le poète les admettra tous.

CHÉNIER A ./ L'AMÉRIQUE / 1794 / p. 111

Solution 5

Un jeune héros-poète dira que dans sa jeunesse il ne chantait que les amours ;

CHÉNIER A ./ L'AMÉRIQUE / 1794 / p. 123

Solution 6

Il faut mettre ceci dans la bouche du poète (qui n'est pas moi).

CHÉNIER A ./ L'AMÉRIQUE / 1794 / p. 124

Solution 7

Le poète divin, tout esprit, tout pensée,
ne sent point dans un corps son âme embarrassée.

CHÉNIER A ./ L'AMÉRIQUE / 1794 / p. 124

Solution 8

Tel que le taon envoyé par Junon va
tourmenter Io... descript... ainsi le poète,
tourmenté par son génie, tourne... descript...
magnum si pectore possit...
et secouer le dieu qui tourmente son sein,
et bientôt ce dieu s'exhale de sa bouche en vers
brûlants...

ô si je puis un jour... Mais quel est ce murmure,
quelle attaque nouvelle et plus forte et plus dure ?

CHÉNIER A ./ L'INVENTION / 1794 / p. 7-8

Solution 9

Tel le bouillant poète, en ses transports brûlants,
le front échevelé, les yeux étincelants,
erre, tourne à grands pas, seul, en d'épais bocages,
s'il pourra de sa tête apaiser les orages,
et secouer le dieu qui fatigue son sein.

CHÉNIER A ./ L'INVENTION / 1794 / p. 9-10

Solution 10

Qui que tu sois enfin, ô toi, jeune poète,
travaille ;

CHÉNIER A ./ L'INVENTION / 1794 / p. 7

Solution 11

tel le bouillant poète, en ses transports brûlants,
le front échevelé, les yeux étincelants,
s'agite, se débat, cherche en d'épais bocages,
s'il pourra de sa tête apaiser les orages,
et secouer le dieu qui fatigue son sein.

CHÉNIER A ./ L'INVENTION / 1794 / p. 23

2. Cooccurrences de "religion" et "liberté"
chez Chateaubriand, *Essai sur les Révolutions*
et Madame de Staël, *Considérations sur la Révolution Française*

_____ Solution 1 _____

si j'ai parlé avec enthousiasme de tout ce qui me paroît beau, et touchant sur la terre, la religion, la vertu, l'honneur, la liberté, l'infortune, il faudra convenir que ma passion supposée pour la célébrité, sort du moins d'un principe excusable : on pourra me plaindre, il sera difficile de me condamner.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XXXV

_____ Solution 2 _____

Se conformer en tout à l'esprit d'élévation et de douceur de l'évangile, marcher avec le temps, soutenir la liberté par l'autorité de la religion, prêcher l'obéissance à la charte comme la soumission au roi, faire entendre du haut de la chaire des paroles de compassion pour ceux qui souffrent quels que soient leur pays et leur culte, réchauffer la foi par l'ardeur de la charité, voilà selon moi ce qui pouvoit rendre au clergé la puissance légitime qu'il doit obtenir : par le chemin opposé, sa ruine est certaine.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLIV

_____ Solution 3 _____

C'est donc une religion de liberté : c'est la mienne.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

_____ Solution 4 _____

Il n'est donc pas vrai que la religion protestante soit plus favorable à la cause de la liberté que la religion catholique.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

_____ Solution 5 _____

Il n'est donc pas vrai que la religion protestante soit plus favorable à la cause de la liberté que la religion catholique.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

Solution 6

Renfermons-nous dans cet axiome : il n'y a point de véritable religion sans liberté, ni véritable liberté sans religion.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

Solution 7

Renfermons-nous dans cet axiome : il n'y a point de véritable religion sans liberté, ni de véritable liberté sans religion.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

Solution 8

Renfermons-nous dans cet axiome : il n'y a point de véritable religion sans liberté, ni de véritable liberté sans religion.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

Solution 9

Renfermons-nous dans cet axiome : il n'y a point de véritable religion sans liberté, ni de véritable liberté sans religion.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

Solution 10

de l'autre, ce sont ces esprits bornés et violents qui disent la religion en péril, parce que nous avons une charte, parce que les divers cultes chrétiens sont reconnus par l'état, et surtout parce que nous jouissons de la liberté de la presse.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. XLVI

Solution 11

Le bandeau royal, celui de la religion, le bonnet de la liberté, peuvent déformer plus ou moins la tête des hommes, mais leur cœur reste toujours le même.

CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
p. 125

Solution 12

le plus honnête homme est celui qui ne falsifie point sa drogue, et qui la débite toute pure, sans en déguiser l'amertume avec

de la liberté, du patriotisme, de la religion.
 CHATEAUBRIAND F. DE / ESSAI SUR LES RÉVOLUTIONS / 1797
 p. 364

_____ Solution 13 _____

L'ordre social, qui admet tous nos
 semblables à l'égalité devant la loi comme devant
 Dieu, est aussi bien d'accord avec la religion
 chrétienne qu'avec la véritable liberté : l'une et
 l'autre, dans des sphères différentes, doivent suivre
 les mêmes principes.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 7

_____ Solution 14 _____

La religion
 protestante, et surtout la liberté de la presse, ont
 depuis créé dans le Danemark une opinion indépendante
 qui sert de limite morale au pouvoir absolu.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 13

_____ Solution 15 _____

Ils ont toujours été et seront
 toujours des amis de la liberté : l'esprit d'examen en
 matière de religion conduit nécessairement au
 gouvernement représentatif, en fait d'institutions
 politiques.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 23

_____ Solution 16 _____

La religion
 fondée sur l'examen, l'instruction publique, les
 élections et la liberté de la presse, sont des sources
 de perfectionnement qui avaient agi depuis plus
 de cent ans en Angleterre.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 161

_____ Solution 17 _____

L'Assemblée constituante, en proclamant la parfaite
liberté des cultes, remplaçait la religion dans
 son sanctuaire, la conscience ;

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 222

Solution 18

Cette même police, pour laquelle nous n'avons pas de termes assez méprisants, pas de termes qui puissent mettre assez de distance entre un honnête homme et quiconque pouvait entrer dans une telle caverne, c'était elle que Bonaparte avait chargée de diriger l'esprit public en France : et en effet, dès qu'il n'y a pas de liberté de la presse, et que la censure de la police ne s'en tient pas à réprimer, mais dicte à tout un peuple les opinions qu'il doit avoir sur la politique, sur la religion, sur les mœurs, sur les livres, sur les individus, dans quel état doit tomber une nation qui n'a d'autre nourriture pour ses pensées que celle que permet ou prépare l'autorité despotique !

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
p. 117

Solution 19

Ils veulent un roi absolu, une religion exclusive et des prêtres intolérants, une noblesse de cour, fondée sur la généalogie, un tiers état affranchi de temps en temps par des lettres de noblesse, un peuple ignorant et sans aucun droit, une armée purement machine, ces ministres sans responsabilité, point de liberté de la presse, point de jurés, point de liberté civile, mais des espions de police, et des journaux à gages, pour vanter cette oeuvre de ténèbres.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
p. 165

Solution 20

Ils veulent un roi absolu, une religion exclusive et des prêtres intolérants, une noblesse de cour, fondée sur la généalogie, un tiers état affranchi de temps en temps par des lettres de noblesse, un peuple ignorant et sans aucun droit, une armée purement machine, ces ministres sans responsabilité, point de liberté de la presse, point de jurés, point de liberté civile, mais des espions de police, et des journaux à gages, pour vanter cette oeuvre de ténèbres.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
p. 165

Solution 21

Le débat entre lui
 et le parti qui lui avait fait la révolution porta sur
 les querelles religieuses : la nation anglaise
 voulait la réformation, et considérait la religion
 catholique comme inconciliable avec la liberté.
 STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 170

Solution 22

Car il en est de la
liberté comme de la religion :
 STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 222

Solution 23

La religion et la liberté président à l'acte
 imposant qui permet à l'homme de condamner
 à mort son semblable.
 STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 318

Solution 24

Grâce à la tolérance, aux institutions politiques
 et à la liberté de la presse, il y a plus de respect
 pour la religion et pour les moeurs en Angleterre
 que dans aucun autre pays de l'Europe.
 STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 355

Solution 25

Car les amis de la liberté sont
 aussi les frères en religion du peuple anglais.
 STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 401

Solution 26

il n'existe presque
 pas de crimes dans le monde que leurs auteurs
 n'aient attribués à l'honneur,
 à la religion, ou à la liberté.
 STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817
 p. 454

Solution 27

On a trouvé le secret de présenter les amis de la
liberté comme des ennemis de la religion : il y a
 deux prétextes à la singulière injustice qui voudrait

interdire au plus noble sentiment de cette terre
l'alliance avec le ciel.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817

p. 457

_____ Solution 28 _____

Certes, ce n'est que parce que
les français n'ont pas uni la religion à la liberté
que leur révolution a sitôt dévié de sa direction
primitive.

STAEL G. DE / CONSIDÉR. RÉVOL. FRANÇAISE / 1817

p. 487

III - DISCOTEXT

DISCOTEXT est la version sur Disque Optique Numérique de la base FRANTEXT. Le DON¹ est un support qui a de quoi séduire l'utilisateur de micro-ordinateur : 546 textes littéraires, entre 1827 et 1923, sur un seul disque, quelque chose comme une demi-tonne de livres accessible à partir d'un poste de travail personnel, sans se soucier de la facturation horaire propre au serveur télématique.

Nous avons mis en oeuvre DISCOTEXT à l'occasion d'un travail sur la notion de décadence à la fin du XIX^e siècle. Le découpage chronologique proposé par DISCOTEXT se prête parfaitement à une étude sur cette période.

La première opération consiste à sélectionner le corpus sur lequel on souhaite travailler. En l'occurrence, la totalité de la base a été choisie, ce qui permet de voir évoluer le thème sur un siècle environ². Il faut ensuite constituer une liste des termes relevant du champ lexical à explorer. L'ordre alphabétique permet de sélectionner une dizaine de mots dérivés de *décadence* (*décadent*, *décadentisme*,...). Il est évident qu'en dehors de ce regroupement autour d'une racine commune, l'utilisateur doit faire appel à d'autres outils pour compléter cette liste. Des termes comme *fin de siècle*, *barbare* ou *dégénérescence* ne peuvent être trouvés qu'intuitivement, ou à l'aide d'un dictionnaire analogique, ou encore en consultant les contextes des premiers mots proposés pour y trouver des corrélats.

L'étape suivante, la recherche des mots dans les textes, est remarquablement rapide. On voit s'afficher à l'écran non seulement le nombre de citations trouvées mais aussi les premiers textes. Il est ainsi facile de faire un choix. Le problème que nous avons rencontré est celui de toute recherche documentaire : le bruit. Le nombre d'occurrences trouvées dépassant la centaine, il est impossible d'en lire tous les contextes.

Un coup d'oeil sur les premiers extraits nous permet de voir que le terme *décadence*, qui représente la grande majorité des occurrences, est polysémique et fournit des contextes où il est question de décadence personnelle et non sociale. Nous le retirons donc de la liste, ce qui nous permet d'obtenir un nombre de textes plus raisonnable, de l'ordre d'une soixantaine.

Recherche dans le corpus	Corpus actif = Base complète
Résultats : 63 occurrence(s) dans 36 texte(s)	
*11. Huysmans, J.-K. - À rebours - 1884 - p. 216	
<p>dont les idées et le style présentassent ces faisandages, ces taches morbides, ces épidermes talés et ce goût blet, qu'il aimait tant à savourer parmi les écrivains décadents, latins et monastiques, des vieux âges.</p> <p>Xiii</p>	
*12. Péladan, J. - Le Vice suprême (1) - 1884 - p. 49	
<p>Le prince de Condé se laissant battre, se laissant pendre par la Feuchères, était vieux ; mais la sénilité du vouloir est de tous les âges. Pantin cassé aux ficelles pendantes, le décadent n'a pas même le ressort qu'il faut pour déplacer son vice et changer de fumier ; il pourrit sur place, heureux</p>	

F2 Zoom F3 (Dé)sélection F4 Voir numéro F5 Sauvegarde Ins Désélectionne tout
Positionne Pgdn Écran suivant Pgup Écran précédent Esc Retour

La lecture des extraits est facilitée par la possibilité de zoomer sur un texte. Malheureusement, le contexte le plus large est d'une page environ, ce qui ne permet pas, dans certains cas, de se rendre compte de l'environnement d'un extrait. En particulier, il est parfois impossible de lire l'en-tête d'une lettre, et donc d'en connaître le destinataire et la date précise³. Les passages qui ne correspondent pas au but de la recherche peuvent être éliminés.

Le plus gros problème que nous avons rencontré est celui de la récupération des données. DISCOTEXT semble n'avoir été prévu que pour la lecture et

l'impression sur papier. Or, le chercheur et l'enseignant ont besoin de récupérer certains extraits pour illustrer une démonstration, un cours, un article. Il est difficile d'admettre qu'un outil aussi perfectionné que DISCOTEXT impose à son utilisateur de recopier les textes qu'il a imprimés. DISCOTEXT ne permet d'ailleurs que l'impression page par page, ce qui interdit à peu près la sortie sur papier d'un corpus qui représenterait une soixantaine de pages et près de deux heures d'impression⁴.

Pour donner une idée plus précise du travail que DISCOTEXT permet d'effectuer, voici la comparaison entre le document fini par l'ordinateur et un bref extrait du résultat final :

Recherche dans le corpus	Corpus actif =
Base complète	
Résultats : 1 occurrence(s) dans 1 texte(s)	
*1.	Goncourt, E. et J. - Journal : 1851-1863 : t. 1 (3) - 1863 - p. 656 À la fin des sociétés troublées, quand il n'y a plus de doctrines, d'écoles, que l'art est entre une tradition perdue et une tradition qui s'inaugure, il se trouve des décadents singuliers, prodigieux, libres, charmants, des aventuriers de la ligne et de la couleur, qui mêlent tout, risquent tout et marquent toutes choses d'un cachet singulier, corrompu, rare ; brouillons de bonne foi, d'élan, d'abondance, de génie, qui semblent un grand artiste manqué, une imagination qui déborde. Tel Fragonard, le plus merveilleux des improvisateurs parmi les peintres. Je m'imagine Fragonard sorti du même moule que Diderot. Chez tous deux, même feu, même verve. Une page de Fragonard, c'est comme une peinture de Diderot. Même ton polissonnant et ému ; tableaux

F2 Zoom Positionne F3 (Dé)sélection Pgdn Écran suivant F4 Voir numéro F5 Sauvegarde Pgup Écran précédent Ins Désélectionne tout Esc Retour

Recherche dans le corpus	Corpus actif =
Base complète	
Résultats : 1 occurrence(s) dans 1 texte(s)	
<p>*1. Bourget, P. - Nouveaux essais de psychologie contemporaine - p. 189</p> <p>littérature. Il s'amuse alors aux bizarreries de la syntaxe, aux curiosités du néologisme. Comme Baudelaire, il se sait décadent et le proclame : " s'il est vrai que les langues aient une décadence, dit un des hommes de lettres amis de Charles Demailly, mieux vaut être Lucain que le dernier imitateur de Virgile qui n'a pas de nom. " peut-être les Goncourt, en citant l'auteur de la Pharsale, n'ont-ils pas très adroitement choisi leur homme ; mais, à coup sûr, ils ont nettement exprimé une théorie, pratiquée par la jeune école, avec un tel parti pris, qu'elle ne semblera une hardiesse à aucun des écrivains nouvellement venus dans la littérature.</p> <p>J'ai rapproché le cas des Goncourt de celui</p>	

F2 Zoom F3 (Dé)sélection F4 Voir numéro F5 Sauvegarde Ins Désélectionne tout
 Positionne Pgdn Écran suivant Pgup Écran précédent Esc Retour

En 1863, ce sont les Goncourt qui, dans leur *Journal*, consigneront le terme "décadent" dans un sens positif, au sujet de Fragonard et Diderot : "À la fin des sociétés troublées, quand il n'y a plus de doctrines, d'écoles, que l'art est entre une tradition perdue et une tradition qui s'inaugure, il se trouve des décadents singuliers, prodigieux, libres, charmants, des aventuriers de la ligne et de la couleur, qui mêlent tout, risquent tout et marquent toutes choses d'un cachet singulier, corrompu, rare ; brouillons de bonne foi, d'élan, d'abondance, de génie, qui semblent un grand artiste manqué, une imagination qui déborde." (t. 1, p. 656). Ils semblent parler à l'intention des jeunes qui vont leur succéder. D'ailleurs, dans ses *Nouveaux Essais de psychologie contemporaine* (1885), Paul Bourget, faisant allusion à leur roman de 1860, *Charles Demailly*, leur attribue la responsabilité d'avoir "nettement exprimé une théorie, pratiquée par la jeune école", celle que Huysmans impulsera.⁵

On voit qu'il y a deux aspects dans l'élaboration qui va d'un état de l'information à l'autre. Le premier, qui nécessite le savoir-faire du chercheur et qui constitue son travail de création, consiste à sélectionner les contextes les plus

intéressants, à les mettre en perspective, à les articuler en une démonstration. Le deuxième consiste à imprimer les résultats, la réutilisation de ceux-ci pour d'autres traitements informatiques n'ayant pas été prévue dans ce CD-ROM.

La première opération dessine l'avenir de la recherche assistée par ordinateur, la deuxième appartient à sa préhistoire⁶.

Michel BERNARD
CNRS-INaLF
Paris

NOTES

1. Ou CD-ROM. Ou DOC (Disque Optique Compact).
2. Il aurait été possible, comme sur STELLA, de réduire le corpus par date, auteur, genre, etc.
3. Même problème pour les journaux intimes. Sauf à se livrer à un exercice fastidieux qui consiste à chercher le contexte des premiers mots de la page et à remonter ainsi de proche en proche jusqu'à l'endroit voulu.
4. On peut toujours utiliser la fonction "impression d'écran" de l'ordinateur mais l'on est alors pénalisé par les encadrements, qui ralentissent l'opération. DISCOTEXT ne permet pas non plus, sauf erreur, de configurer l'impression. Un papier qui ne serait pas à la taille A4, par exemple, pose des problèmes insurmontables.
5. Extrait de l'étude sur *A Rebours* de Huysmans, citée *supra*.
6. Il est à noter que les logiciels de *Zyromys* et du *Robert électronique* permettent d'exporter facilement des informations vers un traitement de texte ; les autres CD-ROM issus de FRANTEXT le permettront également.

