



# FRANTEXT

Base de données  
en littérature française

Guide d'utilisation

Service Commun de la Documentation  
Université d'Avignon et des Pays de Vaucluse

[www.bu.univ-avignon.fr](http://www.bu.univ-avignon.fr)

Mise à jour : mai 2006

## Présentation et Accès

Frantext est un important corpus électronique de textes français s'échelonnant du XVI<sup>ème</sup> au XX<sup>ème</sup> siècle. Il contient environ 3.500 œuvres : essentiellement des œuvres littéraires, mais aussi quelques ouvrages techniques illustrant les principales disciplines scientifiques.

Il est édité par l'ATILF (Analyse et Traitement Informatique de la Langue Française), un laboratoire du CNRS.

Pour accéder à Frantext depuis les postes du SCD et de l'UAPV :  
Bases de données => Sciences sociales et humaines => Littérature  
=> Frantext => [Frantext version non catégorisée](#) => [Entrée dans Frantext](#)

Vous pouvez effectuer dans Frantext des recherches simples ou avancées.

Utilisez toujours **Menu de Frantext**, un rectangle bleu en haut à gauche de l'écran pour effectuer vos requêtes.

## La recherche simple

### a. Définir le corpus de travail

Sélectionnez votre **corpus** : l'ensemble des textes sur lesquels vous voulez travailler. Vous pouvez faire une sélection par : auteur, titre, date, genre (correspondance, pamphlet, poésie, récit de voyage, roman, théâtre...).

[Corpus de travail](#) => [Définir le corpus de travail](#)

Exemple : Auteur = *flaubert*, Titre = *bovary* + [Enregistrer la sélection](#).

Résultat de la sélection : nombre de textes sélectionnés : 1

Si vous désirez connaître les références du/des texte(s) sélectionné(s) :

[Corpus de travail](#) => [Visualiser le corpus de travail](#) => [Affichage détaillé de la bibliographie](#) : FLAUBERT : Gustave – Madame Bovary – 1857 – roman – Paris : Les Belles Lettres, 1945\*.

\*L'édition numérisée dans Frantext n'est pas forcément l'édition originale...

### b. Lancer une recherche

Frantext vous permet de rechercher les **occurrences** d'un mot ou d'une expression dans votre corpus de textes, c'est-à-dire toutes les fois où ce mot ou cette expression apparaît dans les textes que vous avez sélectionnés.

[Recherche dans les textes](#) => [Lancer une recherche](#)

Rentrez au moins un critère dans le formulaire de recherche.

Exemple : Séquence 1 : *bottes* + [Cliquez](#).

Nombre de résultats trouvés : 16

Vous pouvez trier les résultats par ordre chronologique (croissant/décroissant) ou par ordre alphabétique (des auteurs/des titres) : ce tri est nécessaire quand vous obtenez un grand nombre de résultats.

[Visualiser les résultats](#) : les 16 phrases contenant le mot *bottes* (en rouge) s'affichent à l'écran avec leurs références précises.

Cliquez sur [Zoom](#) pour obtenir des extraits plus larges (mais ne respectant pas le découpage en phrases), et sur [Rapatriement des résultats](#) pour les enregistrer sur votre disque dur, sur une disquette ou sur « Mes fichiers ».

### c. Recherche pour le linguiste ou l'historien de la langue française

Cherchez le mot *ramentevoir* dans l'intégralité de la base Frantext:

1- Définissez le corpus de travail : cliquez sur [Enregistrer la sélection](#) sans avoir entré un seul critère de recherche.

2- [Recherche dans les textes](#) => [Lancer une recherche](#) =>

Séquence 1 : *ramentevoir*

Nombre de résultats trouvés : 58

En consultant les résultats selon l'ordre chronologique croissant (par défaut), vous remarquerez que ce mot est utilisé du XVI<sup>ème</sup> siècle (Rabelais, Montaigne...) au milieu du XVII<sup>ème</sup> siècle (Cyrano de Bergerac...), puis qu'ensuite il tombe en désuétude. Pour en savoir plus long sur ce mot : le H vert signale la présence d'un lien hypertexte. Vous pouvez également double-cliquer, dans un des résultats, sur le mot *ramentevoir*, puis, dans la colonne de gauche, sur [Recherche de « Ramentevoir » dans...](#) : l'hyperlien vous donne accès à des bases de données linguistiques comme le TLFi (Trésor de la Langue Française Informatisé) ou le Dictionnaire de l'Académie française : vous y découvrirez le contexte du mot (*vieux, littéraire, rare*), sa définition (*remettre en mémoire, rappeler au souvenir*) et quelques exemples.

### d. Quelques astuces pour optimiser vos recherches

1- [Définir le corpus de travail](#) : Auteur = *maupassant*

2- [Recherche dans les textes](#) => [Lancer une recherche](#) =>

Séquence 1 :

➤ **Rechercher les formes fléchies (conjuguées) d'un verbe & verbe à l'infinitif**

Exemple : *&cinterroger* qui permettra d'obtenir *interroge, interrogeais, interrogeant, interrogeons...*

➤ **Rechercher les formes fléchies d'un substantif (singulier/pluriel) ou d'un adjectif (singulier/pluriel),**

## masculin/féminin)

&m substantif au singulier, ou adjectif au masculin singulier

Exemple :

&mcheval qui permettra d'obtenir *cheval, chevaux*

&mblanc permettra d'obtenir *blanc, blancs, blanche, blanches*

### ➤ Expression d'un choix

Exemple : *Robe (blanche | bleue)* qui permettra d'obtenir *robe blanche, robe bleue*

NB : le signe de relation est un « prompt » : vous l'obtiendrez en appuyant simultanément les touches [Alt Gr] et [6].

### ➤ Deux mots séparés par 0, 1 ou plusieurs autres mots

Exemple : un **&q(0,2)** homme qui permettra d'obtenir *un et homme* séparés par 0, 1 ou 2 autres mots : *un homme, un jeune homme, un grand jeune homme...*

### ➤ Symbole de négation

Exemple : *cheval ^ blanc* qui permettra d'obtenir tout cheval qui n'est pas blanc : *un cheval fougueux, un cheval triste...*

## La recherche avancée

Il existe plusieurs possibilités de recherches avancées.

### a. Recherche de co-occurrences

On appelle **co-occurrence** l'apparition simultanée de deux ou de trois séquences (mots ou expressions) dans un texte. Les séquences doivent être, soit dans la même phrase (option par défaut), soit à une distance maximale de 300 mots.

On peut aussi définir la position relative des co-occurrences : l'une avant ou après l'autre, ou encore position indifférente (par défaut).

Exemple :

1- Définir le corpus de travail : Auteur = *balzac*

2- Recherche dans les textes => Lancer une recherche =>

Séquence 1 = *charbon*, Séquence 2 = *réchaud* Vous + Cliquez

### b. Recherche de listes

#### ➤ Listes thématiques

En créant des listes de mots apparentés par leur sens, cette recherche vous permet d'étudier les champs sémantiques, donc la thématique d'une oeuvre.

Exemple :

1- Définir le corpus de travail : Auteur = *flaubert*, Titre = *salammô*.

2- Listes de mots => Création manuelle d'une liste :

Tapez les mots de votre liste à raison d'un mot par ligne : *serpent / python / boa / dragon*.

Nom de la liste : *reptiles* + Sauvegarde de la liste.

3- Recherche dans les textes => Lancer une recherche =>

Séquence 1 = *&lreptiles*. On obtient d'un seul coup les occurrences des quatre mots de la liste au lieu de faire quatre interrogations successives.

#### ➤ Listes de troncatures

Il est très ennuyeux de devoir recourir aux listes pour effectuer de simples troncatures !

Listes de mots => Création d'une liste à partir des mots du corpus de travail

#### • Mots finissant par...

Critère de sélection : *.\*toire*, Nom de la liste à créer : *toire* Cliquez sur Sauvegarde de la liste : la liste *toire* a été créée.

Recherche dans les textes => Lancer une recherche =>

Séquence 1 : *&ltoire* permettra de trouver les mots finissant par *toire* : *histoire, victoire, méritoire...*

#### • Mots commençant par...

Critère de sélection : *.para.\**, Nom de la liste à créer : *para*.

Recherche dans les textes => Lancer une recherche =>

Séquence 1 : *&lpara* permettra de trouver les mots finissant par *parabole, parader, paradis, paradoxe...*

#### • Mots dont le noyau est...

Critère de sélection : *.\*form.\**, Nom de la liste à créer : *form*.

Recherche dans les textes => Lancer une recherche =>

Séquence 1 : *&lform* permettra de trouver *conforme, déformation, formalités, forme, formelle, formules, réformer, transformer, uniforme...*

### c. Calculs de fréquences

La **fréquence absolue** d'un mot dans un corpus est le nombre d'occurrences de ce mot dans le corpus.

La **fréquence relative** d'un mot dans un corpus est égale à la fréquence absolue de ce mot, divisée par le nombre total des mots du corpus. Ainsi, si le mot *maison* a deux occurrences dans un corpus contenant un million de mots, sa fréquence relative est de deux millièmes.

La fréquence absolue ne présente guère d'intérêt car vous l'obtenez aussi bien en recherche simple (et avec le contexte du mot).

La fréquence relative est plus intéressante : elle mesure l'importance d'un mot ou d'une expression chez un auteur ou pendant une période donnée ; pour qu'elle soit significative, il faut réserver cette recherche à un corpus important.

Exemple :

1- Définir le corpus : tout le corpus

2- Calculs de fréquence => Répartition de la fréquence de mots donnés dans le corpus de travail :

cadre 1 – case 1 (mot voulu) = *liberté*

cadre 2 – fréquences relatives

cadre 3 – choix 3 : (distribution des fréquences) par tranches de temps ; durée d'une tranche de temps (en années) : 10 ; résultats triés par ordre chronologique. Cliquez sur Execution.

Vous obtenez le graphique de l'importance relative du mot *liberté* chez les écrivains français de 1500 à 1999 : à vous de le comparer avec les grandes dates de l'Histoire de France pour mieux saisir l'évolution politique et philosophique de cette notion...

### d. Grammaires

Ces recherches complexes intéresseront plus particulièrement les linguistes experts en théorie du langage.

### e. Etude du voisinage d'un mot

Dans le cadre d'une étude thématique, il peut être intéressant de lister de façon systématique les mots utilisés dans la même phrase ou à proximité d'un mot donné.

Vous obtenez une liste de « mots voisins » qui, à la différence de la recherche simple, sont sortis de leur contexte et triés par ordre alphabétique ou par ordre croissant/décroissant des fréquences. Réservez cette procédure aux recherches donnant de nombreux résultats, afin d'ordonner ceux-ci de façon systématique.

Exemple :

1- Définir le corpus : Auteur = *baudelaire*

2- Etude du voisinage d'un mot => Lancer l'étude de vocabulaire :

cadre 1 – case à remplir pour l'étude d'un mot = *chat*

cadre 2 – cas 1 : le voisinage est la phrase contenant le pivot (le mot donné)

cadre 3 – résultats triés par ordre décroissant des fréquences

Chez Baudelaire, le mot *chat* se voit donc le plus souvent associé à : *coeur, beau, ange, archet, chanter, corde, harmonieux, instrument, mystérieux, parfait, royalement, subtil, sraphique, vibrante, voix, étrange*.

*Bonnes recherches dans Frantext !!*

*Source : aide en ligne de la base*