

# Fonctionnalités textométriques : synthèse et innovations

Bénédicte Pincemin

CNRS, laboratoire ICAR, Lyon

*Séminaire du chantier Textopol, Université Paris 12 Créteil, 20 mars 2010.*

# Texte de référence

- Pincemin, Bénédicte, Heiden, Serge, Lay, Marie-Hélène, Leblanc, Jean-Marc, Viprey, Jean-Marie (2010) - "Fonctionnalités textométriques : Proposition de typologie selon un point de vue utilisateur", *10es Journées internationales d'Analyse statistique des Données Textuelles*, Rome, 9-11 juin 2010.

# Plan

- 1. Contexte de l'étude : plateforme TXM
- 2. Etat de l'art
- 3. Positionnement et objectif
- 4. Présentation de la typologie
  - organisation d'ensemble
  - métafonctionnalités
- 5. Bilan

# 1. Contexte : plateforme TXM

- **Projet ANR Textométrie (2007-2010)**
  - <http://textometrie.ens-lsh.fr/>
  - Fédération des recherches et développements en textométrie autour de la création d'une plateforme logicielle ouverte
    - patrimoine et innovation, pas réingénierie
    - "unifier sans uniformiser" : mutualiser avec modularité ; boîte à outils et applications

# 1. Contexte : plateforme TXM

- Fonctionnalités
  - synthèse des fonctionnalités implémentées et des propositions d'innovation
    - 2007-2008 : document de travail interne, grille comparative...
  - Spécifications
    - JADT 2010 : typologie
- En lien avec :
  - modélisation conceptuelle, terminologie,
  - ergonomie et interface, usages et méthodologie...

# 1. Contexte : logiciels considérés

- **Projet Textométrie :**
  - Hyperbase, Lexico 3, Trameur, Weblex
  - Astartex, DTM-Vic, Sato, Xaira
- **Connaissance plus théorique :**
  - Alceste, Taltac, le Sphinx-Lexica...

# 1. Logiciels : qq liens

- Hyperbase (Nice, Etienne Brunet)
- Lexico 3 (Paris 3, André Salem)
  - <http://www.cavi.univ-paris3.fr/ilpga/ilpga/tal/lexicoWWW/>
- Le Trameur (Paris 3, Serge Fleury)
  - <http://tal.univ-paris3.fr/trameur/>
- Weblex (ENS Lyon, Serge Heiden)
  - <http://weblex.ens-lsh.fr/wlx/>
  - <http://weblex.ens-lsh.fr/doc/weblex/index.html>

# 1. Contexte : Bibliographie

- Lebart & Salem (1994) - Statistique textuelle
  - <http://www.cavi.univ-paris3.fr/lexicometrica/livre/st94/st94-tdm.html>
  - <http://ses.telecom-paristech.fr/lebart/ST.html>
- Actes JADT
  - <http://www.cavi.univ-paris3.fr/lexicometrica/jadt/index.htm>
- etc.

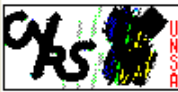


# 1. Contexte : Des comparatifs

- Deux grilles signalées sur le forum de Textopol (rubrique *Liens*) :
  - <http://textpol.free.fr/>
    - Françoise Deconinck
      - [http://www.cati.paris4.sorbonne.fr/centre/activites/colloques/synoptiq\\_11mars05.htm](http://www.cati.paris4.sorbonne.fr/centre/activites/colloques/synoptiq_11mars05.htm)
    - Philippe Gambette
      - <http://tinyurl.com/LogicielsTextometrie>
- Tableau comparatif réalisé dans le projet textométrie

## 2. Etat de l'art : Foisonnement

- Foisonnement de fonctionnalités
  - Hyperbase : interface, manuel
  - Le Trameur : interface, manuel



Index	Mot	Concordance	Contexte	Biblio	Exporter	Créer une base	Choix du corpus
				Lecture	Édition	Choix des bases	aucune

EXEMPLE

Changer la présentation

- RÉGLAGES
- Installer
- Réinitialiser
- Option Vers
- Option calcul
- Option Codes
- Option écran

- INFORMATION
- Info
- Aide

FONCTIONS DOCUMENTAIRES FONCTIONS DE SÉLECTION

anti-bulles

- FONCTIONS STATISTIQUES
- Spécificités
- Graphique
- Liste
- Distribution
- Factorielle
- Arborée
- Évolution
- Phrases-clés
- Excel
- Thèmes
- Associations
- Topologie
- Segments r.

- PROTECTION
- Mem+
- Clé+
- vider
- NAVIGATION
- Anar
- Internet
- Retour
- Quitter



**HYPERBASE**  
version 8.0  
(standard, non lemmatisée)



Cadre Ventilation Section Forme-Lemme Catégorie-Tag Segment Cooc Spécif Concordance Patron Graphe Sélection Rapport Param

**Chargement de la base :  
La Trame + Le Cadre**



**Options chargement :**

Etiquetage  Traceur  Dics

**Exporter la base :  
La Trame + Le Cadre**

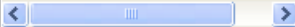


**Représentation du cadre :**

(choisir la/les partie(s))



Parties



(Ctrl-Clic : désélection partie)



**Recherche Forme dans  
le Cadre (regexp) :**

...

(Ecrire motif supra puis <Entrée>)

regexp

CADRE ACTIF : Clic sur partie : les formes | Clic-droit sur partie : les lemmes | Control+Clic sur partie : vocabulaire spécifique  
ARBRE XML : Clic sur noeud : les formes

Mb L. Shift-Clic : sélection | Clic-droit : édition | Ctrl-Clic : noeud | 2-Clic : graphe Partie : ? Annotation : 1 Aperçu : 50

1



## 2. Diversité terminologique et conceptuelle

- Exemples :
  - lecture, édition, (textploreur), (trame)...
  - dictionnaire, vocabulaire, mot...
  - histogramme, ventilation, spécificité...
  - répartition, topologie,...
- L'établissement de grilles comparatives en est d'autant plus périlleux

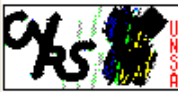
## 2. Conventions terminologiques pour l'exposé

- On compte des *unités* dans des *parties*
- unités comme parties sont représentées par des *propriétés*
- la *valeur* de la propriété considérée donne le *type* auquel rapporter l'*occurrence*

## 2. Structuration

- Fonctionnalités documentaires *vs* statistiques
  - documentaires : "sans calcul", lecture intuitive
  - statistiques : "avec calcul", lecture spécifique
  - cf. interface d'Hyperbase





Index	Mot	Concordance	Contexte	Biblio	Exporter	Créer une base	Choix du corpus
				Lecture	Édition	Choix des bases	aucune

EXEMPLE

Changer la présentation

- RÉGLAGES
- Installer
- Réinitialiser
- Option Vers
- Option calcul
- Option Codes
- Option écran

- INFORMATION
- Info
- Aide

FONCTIONS DOCUMENTAIRES FONCTIONS DE SÉLECTION

anti-bulles

- FONCTIONS STATISTIQUES
- Spécificités
- Graphique
- Liste
- Distribution

- PROTECTION
- Mem+
- Clé+
- vider

- NAVIGATION
- Arborée
- Évolution
- Phrases-clés
- Excel
- Thèmes
- Associations
- Topologie
- Segments r.

- Anar
- Internet
- Retour
- Quitter

**HYPERBASE**  
 version 8.0  
 (standard, non  
 lemmatisée)





## 2. Structuration

- Plan de présentation avec épaisseur historique ( [Hyperbase](#), [Lexico 3](#) ) :
  - texte, dictionnaire -points d'entrée
  - concordance, contextes
  - partitionnement et visualisation de la distribution d'une unité
  - AFC
  - spécificités
  - cooccurrences ; topologie

Lexico3 - [TextPloreur]

Fichier Traitement Fenêtre

Navigation Rapport Dictionnaire

Sélectionnez une couleur :  

Recherche :

Formes (ordre lexicométrique)	Fréquence
de	6130
les	4748
la	4275
et	3773
le	2764
à	2632
que	2496
qui	2079
des	1944
il	1825
l	1770
pour	1665
en	1309
qu	1298
d	1274
nous	1270
un	1267
est	1232
tous	1195
ils	1187
ne	1165
du	1118
foutre	1100
pas	1100
vous	1097
je	979
n	955
dans	924
on	856
a	834
plus	798
se	770
ce	747
...	677

11002 formes

Prêt

C:\Program Files\Lexico 3\Duchn.par

Feuille 1

<<mois=01><quinzaine=11><semaine=111>  
 <<Sda=1793><numero=260><edito=0><Epg=1>  
 <<Sat=0>  
 § la grande douleur du \*père \*duchesne  
 au sujet de la mort de \*marat assassiné à  
 coups de couteau par une garce du \*calvados ,  
 dont l' évêque \*fauchet était le directeur . ses  
 bons avis aux braves \*sans - culottes pour  
 qu' ils se tiennent sur leurs gardes ,  
 attendu qu' il y a dans \*paris plusieurs milliers  
 de tondu de la \*vendée qui ont la patte  
 graissée pour égorger tous les bons citoyens .  
 <<edito=1> § \*marat n' est plus , foutre . peuple , gémis ,  
 <<Epg=2> pleure ton meilleur ami ; il meurt martyr de  
 la liberté . c' est le \*calvados qui a vomi le  
 monstre sous les coups duquel il vient  
 de périr . une jeune fille , ou plutôt une  
 furie armée par les prêtres , et pénitente , dit on ,  
 du cafard \*fauchet , part de \*caen pour  
 exécuter cet horrible attentat . elle arrive à  
 \*paris , et après avoir acheté un large couteau  
 au \*palais que je ne cesserai d' appeler \*royal ,  
 puisqu' il est le rendez - vous de tous les fripons  
 de la terre , elle va trois jours consécutifs  
 frapper à la porte de \*marat et  
 demande à lui parler .  
 le pauvre bougre épuisé de travail , ne

## 2. Structuration : scénario

- Trameur (surtout le **manuel**) :
  - vue d'ensemble du corpus (cadre, carte des sections -cf. "navigation textométrique" dans Lexico 3) : rechercher, visualiser, modifier, spécificités ;
  - autres présentations : concordance, courbes, graphes de mots (SR, cooc, patrons), AFC.

## 2. Structuration : scénario

- Weblex
  - manuel :
    - non focalisé synthétique (dimensions, vocabulaire, rafales...)
    - focalisé sans statistiques (édition, concordance, CQP, répartition...)
    - lexicogramme
    - partition (spécificités)
  - interface :
    - CQP (focalisé + contrainte technologique -index utilisé)
    - non focalisé
    - voisinage d'un mot synthétique (lexicogramme, cooc., SR, termes)
    - spécificités ; édition ; divers

Weblex 3.0 [Frames] : Corpus dh - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

http://weblex.ens-lsh.fr/wlx/cgi/weblex-f?corpus=dh&script=D%E9marrer

Les plus visités Débuter avec Firefox À la une

# Weblex.

## Bienvenue !

© Serge Heiden   
[UMR5191 ICAR CNRS/ENS Lettres](mailto:serge.heiden@ens-lsh.fr)

*Thu Mar 18 15:49:06 WET 2010*

[Manuel de l'Utilisateur de Weblex 3.0](#)

Corpus  

Source A  

( [CQP](#) :       )

à

ou

[Général](#)  
[Index](#)  
[Concordances](#)  
[Vocabulaire](#)  
[Élagage quantitatif](#)

### Options générales de l'affichage

Afficher les résultats dans   Listes sous forme de  (forcer )

Terminé

## 2. Structuration : Scénario

- Lebart & Salem 1994
  - construction des unités
  - cartographie des parties
  - interprétation des parties sur la base des unités (spécificités)

## 2. Granularité

- Grilles
  - ex. concordance : références, tris...
  - illustrations
    - forum Textopol
      - grille Françoise Deconinck
      - grille Philippe Gambette
    - grille Textométrie

## 2. Transversalité : objets de la recherche

- Démultiplié :
  - Hyperbase (sauf Liste)
  - Trameur (sauf gestionnaire de sélection)
- Transversal :
  - Weblex (interface ; mais manuel : 4.4)
  - Lexico 3 (Tgen, drag-and-drop, mais manuel : 2.6)
  - grille projet Textométrie
- => focalisé vs panoramique
- idem : manipulation des sorties (annotation, tris...)



Retour



Forme

Lemme

Expr.

Initial

Final

Chain

Liste

Tout

Nb 162

CONCORDANCE

Trier

Notes



Sommaire

Cliquer sur une ligne pour voir la page

Ma 6d| être que j' avais besoin d' être aimable dans tout ce que je disais de  
 Ma 13c| e fille ; c' était toujours cette aimable enfant . Etait - il question  
 Ma 14b| , et au bout de six mois , cette aimable enfant ne fut plus qu' une pau  
 Ma 32c| s que j' étais d' une figure très aimable , et à cet âge où les grâces s  
 Ma 83a| enir , je ressemblais assez à une aimable petite fille , toute fraîche s  
 Pa 120d| on côté , Geneviève est une fille aimable , je protège ses parents , et  
 Za 178c| , ayant de la santé , une figure aimable , un esprit juste et modéré ,  
 Za 198a| quoi perdre un homme vertueux et aimable . Plein de cette cruelle joi  
 Za 210b| able , non pas parce qu' il était aimable , mais parce qu' il était prem  
 Za 211b| ministre ! La reine disait : L' aimable ministre ! et tous deux ajouta  
 Za 221b| la destinée déplorable de la plus aimable des femmes et de la première r  
 Za 256a| t ce que la nature a fait de plus aimable a péri peut - être d' une mani  
 Za 267c| mauvais que vous fussiez le plus aimable de tous les hommes ; et ce fut  
 Za 292c| t d' inclination pour un homme si aimable . Quand l' ermite et lui fur  
 Ca 349d| - il faire , disais - je , que l' aimable Candide et le sage Pangloss se  
 Hé 501a| ner plus belle encore ; mais plus aimable et plus digne du coeur d' un h  
 Hé 524d| e le plus vertueux , quand il est aimable , il vaut mieux être deux fill  
 Hé 530a| peut aimer ; nous avons perdu ton aimable frère ; je ne vois presque jam  
 Hé 532a| Oh que vous étiez bien plus aimable quand vous étiez moins belle !  
 Hé 541a| Je trouve je ne sais quoi d' aimable et d' attrayant dans cette sag  
 Hé 541c| reux ! Pour profiter d' un état aimable faut - il en négliger un meill  
 Hé 561c| int au Ciel de m' unir à un homme aimable , mais à un homme qui eût l' â  
 Hé 566b| i cette charmante amie , et toute aimable toute piquante qu' elle est ,  
 Hé 585c| e , il faudrait peindre un peuple aimable et galant . Mais toi , ma Ju  
 Hé 605b| eur sagesse égale et sûre , et l' aimable pudeur du sexe , et ses innoce  
 Hé 626b| et les miens sont au comble : l' aimable Julie est à l' extrémité et n'  
 Hé 633a| re . Je n' y vois que toi , mon aimable amie ; ne me prive pas d' une  
 Hé 643c| REPONSE . Il fut un tems , mon aimable Ami , où nos Lettres étoient f  
 Hé 646a| est de toi que je l' attends , mon aimable Ami . Depuis que je n' ose p  
 Hé 655d| utile au public . pénétrer l' aimable secret que tu me dérobes , il  
 Hé 672a| nt j' eusse été amoureux de cette aimable Cousine , si Julie n' eut pas  
 Hé 675c| peux . Pour moi qui connois l' aimable et bonne fille qui l' a écrite  
 Hé 688b| Je ne m' en défends point , mon aimable ami , je voudrais ajouter auta  
 Hé 723c| tant le voile qui les couvre , l' aimable pudeur n' y substitue - t -  
 Hé 731a| à mardi donc , mon aimable ami , mon maître , mon péniten  
 Hé 779c| IE . Faudra - t - il toujours , aimable Cousine , ne remplir envers to  
 Hé 787a| marier : elle est noble , jeune , aimable , riche ; elle ne cède à Julie  
 Em 848d| e de mérite , à l' homme vraiment aimable , et vouloir plaire à ces peti  
 Em 877c| Mais pensez - vous qu' une femme aimable et sage , ornée de pareils tal



**Chargement de la base :  
La Trame + Le Cadre**



**Options chargement :**

Etiquetage  Traceur  Dics

**Exporter la base :  
La Trame + Le Cadre**



**Représentation du cadre :**

(choisir la/les partie(s))



**Parties**

- dhl.1
- text
- front
- div
- p

(Ctrl-Clic : désélection partie)



**Recherche Forme dans  
le Cadre (regexp) :**

..|

(Ecrire motif supra puis <Entrée>)

regexp

CADEP ACTIF : Clic sur partie : les formes | Clic-droit sur partie : les lemmes | Control+Clic sur partie : vocabulaire spécifique  
ARBRE XML : Clic sur noeud : les formes

**Le Cadre Lexicométrique**

- <text(id\_x.c.hr.ce.1950.fr[pos=1])> ...
- <text(id\_x.c.hr.ce.1952.fr[pos=21548])> ...
- <text(id\_x.c.hr.ce.1963.fr[pos=23742])> ...
- <text(id\_x.c.hr.ce.1983.fr[pos=27051])> ...
- <text(id\_x.c.hr.ce.1984.fr[pos=29852])> ...
- <text(id\_x.c.hr.ce.2000.fr[pos=34549])> ...
- <text(id\_x.c.hr.ce.1961.fr[pos=37915])> ...
- <text(id\_x.c.hr.ce.1996a.fr[pos=60793])> ...
- <text(id\_x.c.hr.ce.1988.fr[pos=91661])> ...

Nb L. Shift-Clic : sélection | Clic-droit : édition | Ctrl-Clic : noeud | 2-Clic : graphe Partie : text(91661,99014) Annotation : 1 Aperçu : 50

```

1          1988-05-05      Strasbourg      Protocole additionnel à la Charte
2 sociale européenne      Les Etats membres du Conseil de
3 l'Europe, signataires du présent Protocole,      Résolus à prendre de nouvelles
4 mesures propres à élargir la protection des droits
5 sociaux et économiques garantie par la Charte sociale
6 européenne, ouverte à la signature à Turin le 18 octobre
7 1961 (ci-après dénommée "la Charte"),      Sont convenus
8 de ce qui suit:      Partie I Les Parties reconnaissent comme objectif d'une
9 politique qu'elles poursuivront par tous les moyens utiles,
10 sur les plans national et international, la réalisation de
11 conditions propres à assurer l'exercice effectif des droits
12 et principes suivants: Tous les travailleurs ont droit à
13 l'égalité de chances et de traitement en matière d'emploi et
14 de profession, sans discrimination fondée sur le sexe. Les
15 travailleurs ont droit à l'information et à la consultation
16 au sein de l'entreprise. Les travailleurs ont le droit de
17 prendre part à la détermination et à l'amélioration des
18 conditions de travail et du milieu du travail dans
19 l'entreprise. Toute personne âgée a droit à une protection
20 sociale.      Droit à l'égalité de chances
21 et de traitement en matière d'emploi et de profession,
22 sans discrimination fondée sur le sexe.
    
```

**Ventilation d'item :**

[sélection po]  RegExp

(Ecrire motif supra puis <Entrée>)

**Sélection Annotation :**

Forme  Lemme  Catégorie

**Annotation sélectionnée :**

1 Forme

**Type de graphe :**

Bar   Ligne

**Parties**

dhl.1  
text  
front  
div  
p

(Ctrl-Clic : désélection partie)

**Paramètres Graphique :**

**Taille de l'image**

800x600   1000x800

**Fréquence**

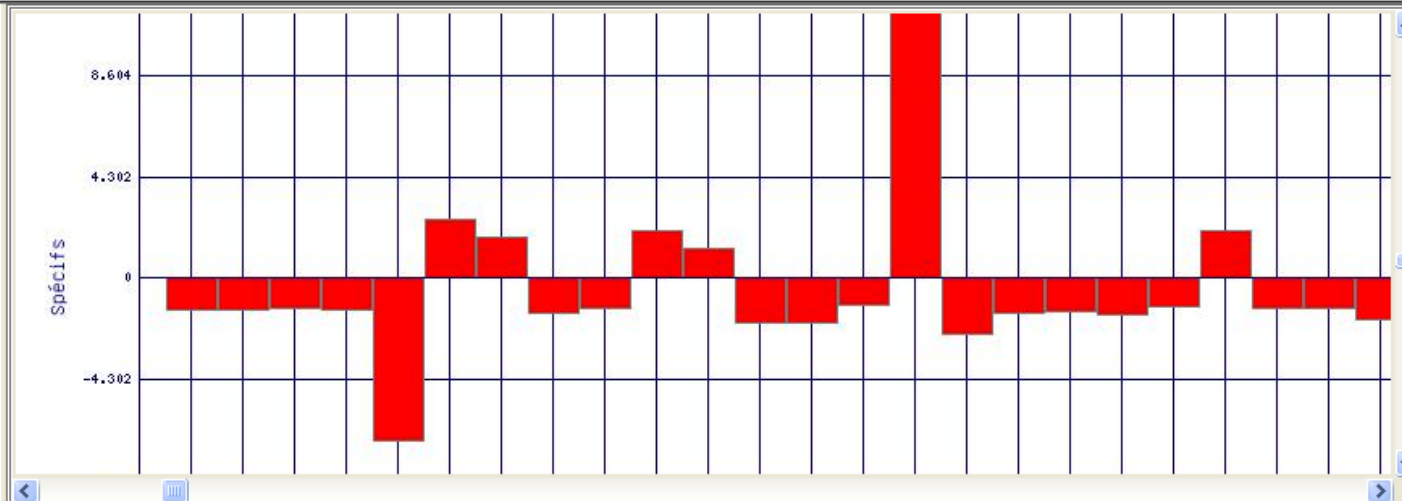
Absolues  Relatives  Spécifs

**Accroissement de vocabulaire :**



Trace Trame

**Ajouter au Rapport :**



**Chargement de la partie : text**

```
<Jeu de positions : {positions(égalité)} >
<id_di:t.dh.al.1969.fr>
Fréquence de {sélection-positions(égalité)} : 3
<id_x.c.hr.al.1948a.fr>
Fréquence de {sélection-positions(égalité)} : 1
<id_x.c.hr.al.1948b.fr>
Fréquence de {sélection-positions(égalité)} : 0
<id_x.c.hr.al.1985.fr>
Fréquence de {sélection-positions(égalité)} : 0
<id_x.c.hr.al.1988.fr>
Fréquence de {sélection-positions(égalité)} : 0
<id_x.c.hr.al.1990.fr>
Fréquence de {sélection-positions(égalité)} : 0
<id_x.c.hr.al.1994.fr>
Fréquence de {sélection-positions(égalité)} : 3
<id_x.c.hr.al.1998a.fr>
Fréquence de {sélection-positions(égalité)} : 1
<id_x.c.hr.al.1998b.fr>
Fréquence de {sélection-positions(égalité)} : 1
<id_x.c.hr.ce.1950.fr>
Fréquence de {sélection-positions(égalité)} : 0
<id_x.c.hr.ce.1952.fr>
Fréquence de {sélection-positions(égalité)} : 0
```

Weblex 3.0 [Frames] : Corpus dh - Mozilla Firefox

http://weblex.ens-lsh.fr/wlx/cgi/weblex-f?corpus=dh&script=D%E9marrer

## Index des occurrences de *[word="liberty"]* dans le corpus dh-lang-EN

ord	f	événement
1	4	liberty and security
2	3	liberty shall be
3	2	liberty and social
4	2	liberty by arrest
5	2	liberty of individuals
6	2	liberty of movement
7	2	liberty of parents
8	1	liberty ( other
9	1	liberty , duties

Corpus: dh-lang-EN

Source: [word="liberty"]

Recherche simplifiée

( CQP : Index | Concordances | Références | Contextes | Répartition | Spécificités )

Vocabulaire | Répartitions | Pareto | Zipf | Longueur des phrases | Dimensions

Lexicogramme | Lexicogramme récursif | Lexicogrammes récursifs | Cooccurrences | Segments répétés | Termes

Spécificités total alpha | Spécificités total fréq. | Spécificités parties

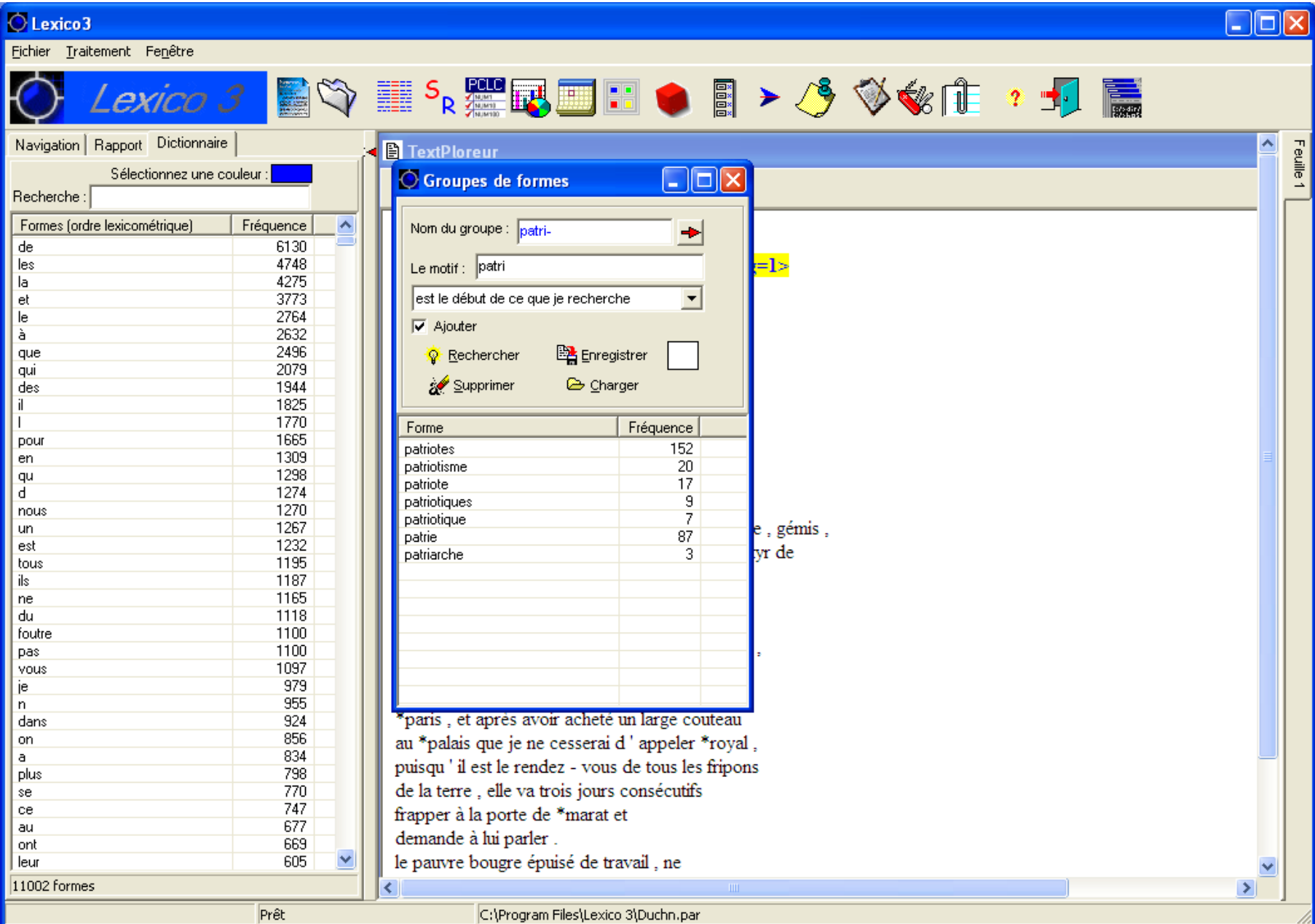
[Général](#)  
[Index](#)  
[Concordances](#)  
[Vocabulaire](#)

### Options générales de l'affichage

Afficher les résultats dans: output

Listes sous forme de: Tableau ( forcer

Terminé



**Groupes de formes**

Nom du groupe : patri-

Le motif : patri

est le début de ce que je recherche

Ajouter

Rechercher    Enregistrer   

Supprimer    Charger

Forme	Fréquence
patriotes	152
patriotisme	20
patriote	17
patriotiques	9
patriotique	7
patrie	87
patriarche	3

\*paris , et après avoir achete un large couteau  
au \*palais que je ne cesserai d ' appeler \*royal ,  
puisque ' il est le rendez - vous de tous les fripons  
de la terre , elle va trois jours consécutifs  
frapper à la porte de \*marat et  
demande à lui parler .  
le pauvre bougre épuisé de travail , ne

### 3. Objectif : vue unifiée

- Technique : état de l'art pour spécification d'une plateforme textométrique
- Applicatif : repères synthétiques pour comprendre les possibilités d'un logiciel
  - *mais attention à une comparaison trop directe, trop grossière*
- Théorie : logique d'ensemble, lieux éventuels d'innovation (ex. combinatoire, cf. Sato)

# 3. Objectif : vue orientée utilisateur

- Simple paramétrage, ou usages différents ?
  - retour au texte, attestations = une fonction :
    - mot, taille de contexte → contextes
  - KWOC / KWAC / KWIC / (KWUT) (Pincemin Albi 2006)
    - KWAC = concordance
      - effet de superposition par alignement vertical, tri, et affichage sur une seule ligne
      - analyse à l'écran, présentation dynamique
    - KWIC = contextes
      - confort de lecture (lignes pas trop longues), observation sur un contexte plus large
      - lecture et publication papier



Weblex 3.0 [Frames] : Corpus dh - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

http://weblex.ens-lyon.fr/wlx/cgi/weblex-f?corpus=dh&script=D%E9marrer

Les plus visités Débuter avec Firefox À la une

Weblex 3.0 [Frames] : Corpus dh

3 [DI T DH CE 1996A EN](#) workers and employers have the right to **freedom of** association in national or internationa

4 [DI T DH AL 1950 EN](#) declares recognition of the principle of **freedom of** association to be a means of improving

5 [DI T DH CE 1998A EN](#) t to freedom of peaceful assembly and to **freedom of** association with others , including the

6 [DI T DH ONU 1976B EN](#) thers . Everyone shall have the right to **freedom of** association with others , including the

7 [DI T DH CE 1998A EN](#) of expression Everyone has the right to **freedom of** expression . This right shall include f

8 [DI T DH ONU 1976B EN](#) rence . Everyone shall have the right to **freedom of** expression ; this right shall include f

9 [DI T DH ONU 1990 EN](#) law . The child shall have the right to **freedom of** expression ; this right shall include f

10 [DI T DH AL 1950 EN](#) claration of Philadelphia reaffirms that **freedom of** expression and of association are essen

11 [DI T DH ONU 1976A EN](#) be construed so as to interfere with the **liberty of** individuals and bodies to establish and

12 [DI T DH ONU 1990 EN](#) be construed so as to interfere with the **liberty of** individuals and bodies to establish and

13 [DI T DH AL 1992A EN](#) l be interpreted as a restriction of the **freedom of** individuals and entities to establish a

14 [DI T DH AL 1948 EN](#) mination . Every person has the right to **freedom of** investigation , of opinion , and of the

Corpus  Recherche simplifiée

Source

( [CQP](#) :       )

[Général](#)  
[Index](#)  
[Concordances](#)  
[Vocabulaire](#)

**Affichage du vocabulaire**

Terminé

Weblex 3.0 [Frames] : Corpus dh - Mozilla Firefox

http://weblex.ens-lsh.fr/wlx/cgi/weblex-f?corpus=dh&script=D%E9marrer

Les plus visités Débuter avec Firefox À la une

## Il y a 32 occurrences de *[word="freedom|liberty"]* dans le corpus dh-lang-EN

[DI T DH AL 1948 EN](#) o profess a religious faith , and to manifest and practice it both in public and in private . Right to freedom of investigation , opinion , expression and dissemination . Every person has the right to **freedom of** investigation , of opinion , and of the expression and dissemination of ideas , by any medium whatsoever . Right to protection of honor , personal reputation , and private and family life . Every person has the right to the protection of the law against abusive attacks upon his honor , his reputation , and his private and family life . Right to a family and to protection thereof . Every person ha

[DI T DH AL 1948 EN](#) or any other factor . Right to religious freedom and worship . Every person has the right freely to profess a religious faith , and to manifest and practice it both in public and in private . Right to **freedom of** investigation , opinion , expression and dissemination . Every person has the right to freedom of investigation , of opinion , and of the expression and dissemination of ideas , by any medium whatsoever . Right to protection of honor , personal reputation , and private and family life . Every person has the right to the protection of the law against abusive attacks upon his honor , his reputation

Corpus dh-lang-EN ? Recherche simplifi

Source [word="freedom|liberty"]

( [CQP](#) : Index Concordances Références **Contextes** Répartition Spécificités )

Vocabulaire Répartitions Pareto Zipf Longueur des phrases Dimensions

Lexicogramme Lexicogramme récursif Lexicogrammes récursifs Cooccurrences Segments répétés Termes

Spécificités total alpha Spécificités total fréq Spécificités parties

[Général](#)  
[Index](#)  
[Concordances](#)  
[Vocabulaire](#)

Tri : 1<sup>ère</sup> clé Référence 2<sup>ème</sup> clé Aucun  
3<sup>ème</sup> clé Aucun 4<sup>ème</sup> clé Aucun

Terminé

### 3. Objectif : vue orientée utilisateur

- Une même fonction, des usages différents :
  - idem : calcul des spécificités
    - étude d'une partition
    - une mesure de cooccurrence
- Des calculs différents mais qui répondent à un même questionnement :
  - répartition au fil du texte :
    - carte des sections, rafales, accroissement du vocabulaire, topologie...

# 4. Organisation de la typologie

- En amont : établissement des données
- Fonctionnalités de calcul textométrique
  - Lecture : on voit des *occurrences* ; retour au texte
  - Synthèses : on voit des *types*
- En aval : exploitation des résultats

# 4. Périmètre/Limites de la typologie

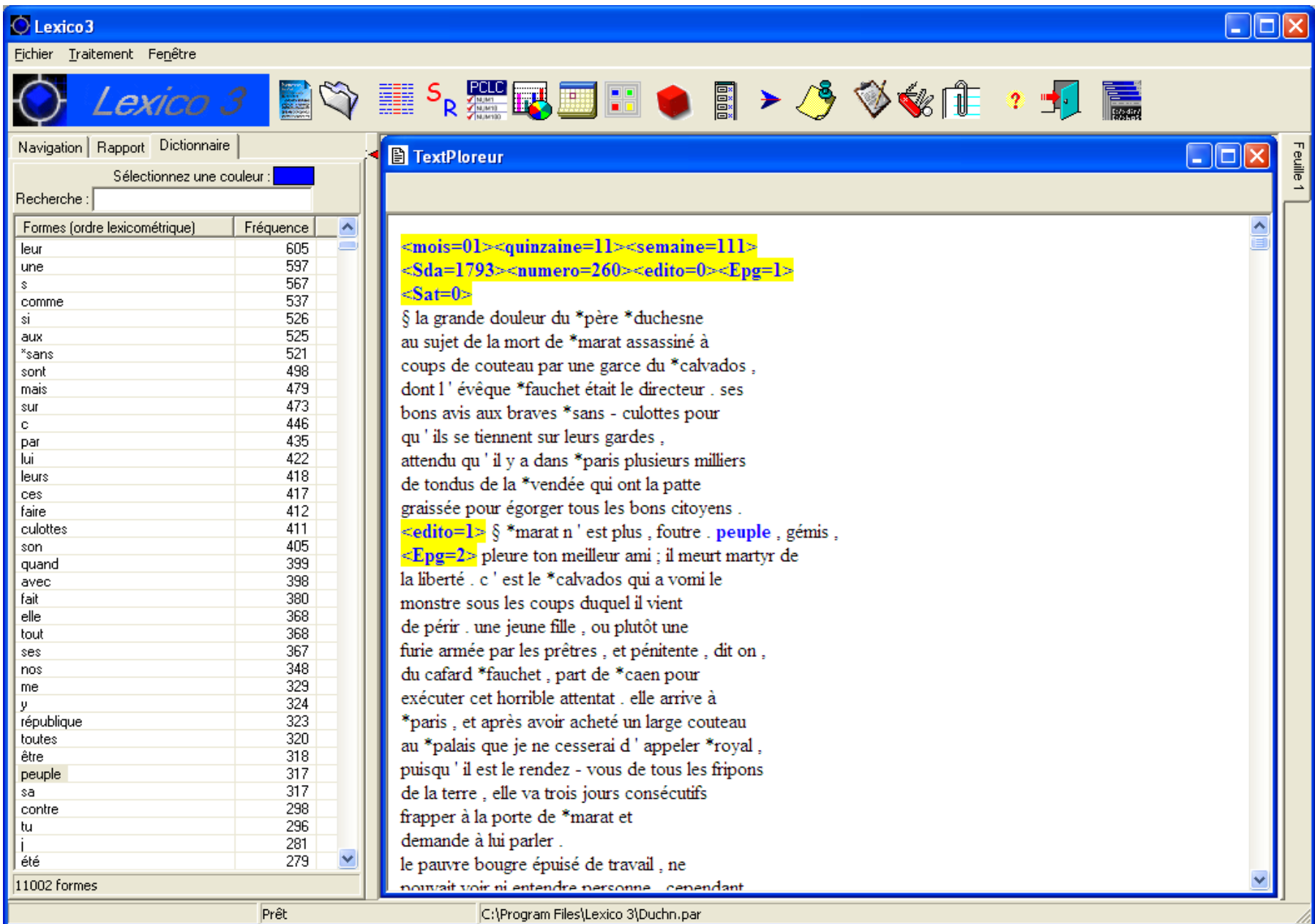
- Ergonomie
  - aide en ligne
- Dynamique d'usage
  - navigation hypertexte
- Description détaillée de chaque métafonctionnalité

# 4. Typologie : Lecture

- Texte
- Vue interne
- Extraits
- Concordance

# 4. Typologie : Texte

- Définition
  - affichage du texte
    - linéarité
    - structure logique (paragraphes, sections...) -vue globale/locale
    - typographie
  - possibilités de navigation
    - accès séquentiel et direct (page, chapitre, occurrence...)
    - parcours d'occurrence en occurrence
    - articulation de vues globale et locale





# 4. Typologie : Texte

- Usage
  - retour au texte
    - feuilletage
    - contexte non limité en taille
  - occurrences groupées, disposition des occurrences

## 4. Typologie : Vue interne

- Visualisation de la représentation sur laquelle opèrent les traitements
  - structure (segmentation, découpage...)
  - propriétés (toutes ou choix)
  - présentations adaptées : tableau, arborescence...
- Percevoir pour contrôler le traitement
  - conception des interrogations
  - interprétation des résultats

# Affichage tabulé des propriétés forme, P2, P3, P4, P5, P6, P7, P8 des occurrences de la phrase n°100 dans le corpus socio

◀ 100 ▶

ord	forme	P2	P3	P4	P5	P6	P7	P8
1	Il	Pp3msn	il	ilP	1	-	sujm	1
2	est	Vmip3s	être	êtreV	2	2	vm	1
3	temps	Ncm.	temps	tempsN	2	3	vm	1
4	d'	Sp	de	deS	6	-	saja	1
5	en	Pp3..-	en	enP	6	-	saja	1
6	tirer	Vmn--	tirer	tirerV	6	-	saja	1
7	les	Da-p-d	le	leD	8	-	codi	1
8	leçons	Ncfp	leçon	leçonN	8	-	codi	1
9	:	Yps	:	:Y	-	-	-	0

Corpus socio



Recherche simplifiée

Source

Spécificités total alpha   Spécificités total fréq   Spécificités parties

Édition   Édition phrase 100 à 100

Biblio   Crosstag   Confusion   ou   Ré-initialiser les paramètres

[Général](#)

[Index](#)

[Concordances](#)

[Vocabulaire](#)

● Composer l'index avec les champs :

word  p2  p3  p4  p5  p6  p7  p8  p9  p10

p11  p12  p13  p14  p15  p16  p17  p18  p19  p20

# 4. Typologie : Extraits

- Série de passages, avec localisation
  - tris ; regroupements
  - critère :
    - contextes (un focus, une cooccurrence), avec différents types de délimitation (fenêtre, délimiteur, structure)
    - mais aussi : propriété, position, score (cf. résumés d'Hyperbase)

C:\HYPERBAS\EXEMPLE.EXE

Refaire résumé **Chouans** Voir les mots spécifiques et cacher les passages spécifiques Trier Recherche Retour Sommaire **Pons** Cliquer sur un extrait pour voir le contexte

CLIC sur un mot: Recherche du mot dans les spécificités **corpus trié**

6.086 Qu'as-tu donc, **COMMANDANT**, est-ce que tu me connaissais ? REPRIT brusquement le **JEUNE HOMME**.

5.919 Au **MOMENT** où les **BAÏONNETTES** des quatre **SOLDATS** ne brillèrent plus, le **CAPITAINE** Merle revenait, après avoir accompli les **ORDRES** du **COMMANDANT** avec la **RAPIDITÉ** de l'éclair.

5.633 Est-ce que vous voulez les lire, **DEMANDA** impertinemment le **JEUNE MARIN** dont l'oeil bleu plein de **MALICE** étudiait alternativement la **SOMBRE FIGURE** du **COMMANDANT** et celle de Mlle de Verneuil.

5.562 Sur deux ou trois **ORDRES** donnés à **VOIX BASSE** par le **CHEF** et transmis par Marche - à - **TERRE** aux **Chouans** en bas **BRETON**, ils opérèrent leur retraite avec une habileté qui déconcerta les **Républicains** et même leur **COMMANDANT**.

5.259 Il y eut un **MOMENT** de **SILENCE**; mais il fut bientôt rompu par la **JEUNE Parisienne** qui marcha au **COMMANDANT**, lui tendit la **MAIN** et lui dit: Colonel, **QUOIQUE** votre barbe soit un peu longue, vous pouvez m'embarrasser, vous êtes un **HOMME**.

5.23 L'ordre de Hulot fut difficilement exécuté; car en **ENTENDANT** la **VOIX** de son adversaire, le **JEUNE CHEF** s'écria: Par sainte Anne d'Auray, ne les lâchez pas! égaillez-vous, mes **GARS**.

5.213 Mlle de Verneuil **REGARDA** fixement le froid et immobile **JEUNE HOMME** qui la dédaignait, mais elle ne vit rien en lui qui pût trahir un sentiment de crainte; elle l'instruisit par un sourire amer de la découverte qu'elle faisait en ce **MOMENT** du **SECRET** si traîtreusement gardé par lui; puis, d'une **VOIX** railleuse, les narines enflées de joie, la **TÊTE** de côté pour **EXAMINER** le **GENTILHOMME** et voir Merle tout à la fois, elle dit au **Républicain**: Ce **CHEF** - là, **CAPITAINE**, donne bien des inquiétudes au **Premier CONSUL**.

5.183 Cet **HOMME** - là, disait le **JEUNE MARIN** à l'**OREILLE** de l'hôtesse, est quelque **ESPION** de Fouché.

5.111 Monsieur le **MARQUIS**, dit Marche - à - **TERRE** en achevant son discours, c'est d'autant plus mal à eux de m'en vouloir que j'ai laissé en

6.807 Supposez aux époux **TRENTE MILLE FRANCS** de **RENTE**? dans le monde actuel, le **GARÇON RICHE** devient un **PAUVRE** diable qui regarde au **PRIX** d'une course à Chantilly.

6.732 Eh! mademoiselle, si **MONSIEUR** votre **COUSIN** voulait me vendre sa **COLLECTION**, j'en donnerais ce soir huit **CENT MILLE FRANCS**, et je ne ferais pas une mauvaise **AFFAIRE**.

6.563 Depuis vingt ans que le **VIEUX MUSICIEN DÎNAIT** chez son unique **COUSIN** Camusot, le **PAUVRE HOMME** attendait encore un mot sur ses **AFFAIRES**, sur sa vie, sur sa santé.

6.351 Si vous vouliez vendre votre **COLLECTION** de **TABLEAUX**, j'en donnerais bien **CINQ** à six **CENT MILLE FRANCS**...

6.232 Le **PRÉSIDENT** de Marville demeurait rue de Hanovre, dans une maison achetée depuis dix ans par la **PRÉSIDENTE**, après la mort de son père et de sa mère, les **SIEUR** et dame Thirion, qui lui laissèrent **ENVIRON CENT MILLE FRANCS** d'économies.

5.917 Elle ne pardonnait pas au **RICHE MARCHAND** droguiste, **ANCIEN PRÉSIDENT** du **TRIBUNAL** de commerce d'être devenu successivement **DÉPUTÉ**, **MINISTRE**, comte et **PAIR**.

5.688 Il se trouve en ce moment pour **DOUZE CENT MILLE FRANCS** de fermes et d'herbages à vendre, dit le **PRÉSIDENT**.

5.685 Donner **CINQ CENT MILLE FRANCS** à son compagnon d'infortune! oh! maman, j'aurai voiture et **LOGE** aux Italiens.

5.622 Elle alla jusqu'à regarder comme une bonne **AFFAIRE** un **HOMME** de **QUARANTE** - huit ans, pourvu qu'il eût vingt **MILLE FRANCS** de **RENTE**.

5.419 Dans ce siècle, où le luxe a pénétré jusque dans les loges de **CONCIERGE**, les jeunes **GENS** hésitent à joindre leur sort à celui de la **FILLE** d'un **PRÉSIDENT** à la Cour **ROYALE** de Paris, quand on ne lui constitue que **CENT MILLE FRANCS** de dot.

5.392 Cette maison, objet de la convoitise du coiffeur, lui fut vendue moyennant

Menu déroulant pour le choix des textes

# 4. Typologie : Extraits

- Usages
  - recherche d'un exemple, d'une attestation
  - analyse systématique, en classant les extraits

# 4. Typologie : Concordance

- Heuristique visuelle mettant en évidence les régularités de voisinage
  - alignement vertical en colonne
  - présentation sur une seule ligne
  - tri (multiple) sur le focus, à gauche, à droite...
  - perspectives : zones (Pincemin et al. JADT 2006)
- Importance de la localisation

Retour



Forme

Lemme

Expr.

Initial

Final

Chain

Liste

Tout

Nb 305

CONCORDANCE

Trier



Notes



Trier



Notes



Sommaire



Cliquer sur une ligne pour voir la page

Po2463c| erte de raison sur l' arrangement musical des choses en ce bas monde .  
M13445a| ausses selon les règles de l' art musical le rendent intraduisible . M  
Tm8376a| reste , - n' était pas seulement musical et reptile , il avait jusqu' à  
Sw7493b| thme , devenant fluide , sonore , musicale , innombrable , universelle :  
Bo4130a| parut prise d' une grande ardeur musicale . Un soir que Charles l' éc  
Sw7955f| me qui a reçu une forte éducation musicale , battant la mesure avec sa t  
Bo4132a| z pour plus tard sur l' éducation musicale de votre enfant ! Moi , je  
Sw7477a| sions rares , cette même effusion musicale , cette même philosophie idéa  
Sw7970d| edoublant de vitesse , l' émotion musicale était à son comble , un domes  
Po2795c| uelques instants , l' inspiration musicale , excitée par le tremblement  
Sw7706c| rée , il avait entendu une oeuvre musicale exécutée au piano et au viol  
St7093b| aisir et de passion . La partie musicale avait été confiée à un remarq  
Po2354e| ne ou la réplique de cette phrase musicale .  
Sw7710c| Même cet amour pour une phrase musicale sembla un instant devoir amor  
Tm8559e| : le petit sillon qu' une phrase musicale ou la vue d' une église a cre  
Sw7701d| la Patronne et de sa sensibilité musicale .  
Po2361b| i veut de sérieuses connaissances musicales . Schmucke , par le consei  
Po2795d| de Liszt , les deux organisations musicales qui se rapprochent le plus d  
Sw8000b| ons qu' il faudra que ces phrases musicales , ces notions qui existent p  
Tm8515c| matinale enveloppant des phrases musicales qui en émergent partiellemen  
Tm8278c| ailleurs par des raisons purement musicales : Dame , c' est que la musiq  
Sw7707c| être pourtant les seules purement musicales , inévitables , entièrement o  
Sw7567e| tendus comme certains intervalles musicaux , elles m' offraient indéfini  
Sw7997b| en lui , Swann tenait les motifs musicaux pour de véritables idées , d'  
Po2341b| On devenait poète , écrivain , musicien à si peu de frais ! Pons ,  
Po2339a| ssemblances . M . Sauvageot , musicien comme Pons , sans grande fort  
Hé 708c| contrainte que donne chez nous au musicien le perpétuel combat du chant  
Sw7998a| musique , que le champ ouvert au musicien n' est pas un clavier mesquin  
Po2898b| rse quinze louis et les tendit au musicien . C' est à vous , c' est si  
Sw7715c| aisaient profession d' admirer ce musicien ( quand Swann avait dit que s  
Bo3788b| Aussi je ne m' étonne plus de ce musicien célèbre qui , pour exciter mi  
Po2656a| précisément par votre bonhomme de musicien qui a voulu déshonorer , par  
Po2448c| de France , qu' il prit le digne musicien à part . Ah ça , mon vieil  
Po2624a| , de l' artiste du comédien , du musicien , sont égayées par les jovial  
Po2366d| , maison où le père et la mère du musicien étaient commanditaires après  
Ca 365b| très grand succès , et j' ai été musicien de la chapelle de madame la p  
Po2333c| at à Rome , pour devenir un grand musicien , Sylvain Pons en avait rappo  
Po2333a| re , le peintre , le graveur , le musicien obtenus par ce procédé mécani  
Po2493c| ront , le but sera le même . Le musicien descendit avec rapidité l' es

Extrait 1 Suite →



# 4. Typologie : Synthèses

- Relevés
  - Vocabulaire
  - Mesures
- Positions
- Associations

# 4. Typologie : Vocabulaire

- Liste des unités avec leur fréquence
  - tri (hiérarchique, alphabétique...)
  - propriété d'analyse
  - panoramique ou focalisée
- Atouts
  - simplicité d'interprétation
  - réduction
- Usages
  - point d'entrée
  - ajustement d'une sélection

Lexico3 - [Concordances]

Fichier Traitement Fenêtre

Navigation Rapport Dictionnaire

Recherche : Sélectionnez une couleur : [bleu]

Forme : [ ] Tri : Après Regroupement : <Aucun> Largeur : 40

Expression rationnelle Type de documentation : Concordance Délimiteurs de séquence : [?/\_\^"000\$]

Formes (ordre lexicographique)	Fréquence
plaire	4
plaisir	11
plaisirs	13
plait	5
planche	2
plans	3
planta	2
plantant	1
plante	3
planté	2
plantée	1
plantées	1
planter	2
planteront	2
plantes	5
plantez	1
plat	6
plate	4
plates	1
plâtre	4
plats	7
plein	20
pleine	4
pleines	9
pleins	2
pleurant	1
pleure	4
pleurent	3
pleurer	13
pleurera	2
pleurerons	1
pleureurs	1
pleureuses	1
pleurez	1
pleut	2
pleuvaient	1

11002 formes

ils sont destinés au crime , comme telle **plante** , à nous empoisonner . il est aussi natu  
n pas pour s ' amuser à la moutarde , en **plantant** de ville en ville l ' arbre de la libe  
remier fruit de la vigne qu ' elle avait **plantée** , et d ' avoir fait , avec le petit avo  
disait le jean - foutre , les \*français **planteront** l ' arbre de la liberté à \*vienne ,  
is rentreront dans leurs murs , et ils y **planteront** l ' arbre de la liberté , mais , foug  
ui vous endormez comme le patriarche qui **planta** la vigne , après qu ' il eut trouvé le se  
e . j ' aime de tout mon coeur celui qui **plante** la vigne , celui qui élève et engraisse  
plus également à ses enfants ; celui qui **plante** la vigne doit avoir la plus belle grappe  
tres , enfants bâtards du patriarche qui **planta** la vigne , les \*normands qui vous ont va  
on , achever votre ouvrage ? après avoir **planté** la vigne , vous vous endormez comme le pa  
tre , qu ' une bonne loi les envoie tous **planter** leurs choux ; qu ' il soit établi à \*pa  
comme un champignon , le chêne que vous **plantez** ne vous ombragera pas de ses rameaux ,  
sont comme les arbres ; celui qui a été **planté** par un bon cultivateur , qui a été greffé  
sont glissés à la montagne ; à côté des **plantes** salutaires croissent les plus mortels po  
, et si trois cents potences avaient été **plantées** sur la terrasse des \*tuileries pour y  
and je proposai , il y a trois mois , de **planter** trois cents potences sur la terrasse de  
eul épi ne mûrira à son ombre ; mais les **plantes** venimeuses , c ' est à dire la noblesse  
ts les plus doux , si nous arrachons les **plantes** venimeuses ; mais , foutre , si nous no  
une main salulaire a éloigné toutes les **plantes** venimeuses qui auraient dévoré sa sève .  
es ronces et les épines . détruisons les **plantes** venimeuses qui croissent encore sur son

Nombre de contextes : 20

Prêt C:\Program Files\Lexico 3\Duchn.par

# 4. Typologie : Mesures

- Caractérisations quantitatives
  - taille du corpus, des parties, des phrases
  - nombre de types, d'occurrences, de hapax
  - Zipf-Pareto, richesse du vocabulaire...
  - perspective : construction de tableaux de mesures

**Lexico3**  
 Fichier Traitement Fenêtre

Navigation Rapport Dictionnaire

Sélectionnez une couleur : █

Recherche :

Formes (ordre lexicométrique)	Fréquence
leur	605
une	597
s	567
comme	537
si	526
aux	525
*sans	521
sont	498
mais	479
sur	473
c	446
par	435
lui	422
leurs	418
ces	417
faire	412
culottes	411
son	405
quand	399
avec	398
fait	380
elle	368
tout	368
ses	367
nos	348
me	329
y	324
république	323
toutes	320
être	318
peuple	317
sa	317
contre	298
tu	296
i	281
été	279

11002 formes

Prêt

C:\Program Files\Lexico 3\Duchn.par

---

**Principales caractéristiques de la partition : mois**

Nombre d'occurrences: 142177    Nombre de formes: 10988  
 Nombre d'hapax: 5056    Fréquence maximale: 6130

	Num	Partie	Occurenc	Formes	Hapax	Fmax	Forme
✓	1	01	19256	3460	2054	886	de
✓	2	02	20342	3795	2233	875	de
✓	3	03	19277	3586	2121	853	de
✓	4	04	18004	3571	2155	753	de
✓	5	05	17814	3389	2027	746	de
✓	6	06	17545	3413	2084	757	de
✓	7	07	16295	3313	2031	669	de
✓	8	08	13644	2996	1846	591	de

Princ. Caract. Lexicométriques

---

**TextPloreur**

<mois=01><quinzaine=11><semaine=111>  
 <Sda=1793><numero=260><edito=0><Epg=1>  
 <Sat=0>

§ la grande douleur du \*père \*duchesne  
 au sujet de la mort de \*marat assassiné à  
 coups de couteau par une garce du \*calvados

# 4. Typologie : Synthèses

- Relevés
- Positions : *caractérisation des localisations d'unités dans le corpus*
  - Déroulement
  - Distribution
  - Evolution
- Associations

# 4. Typologie : Déroulement

- Définition :
  - Au fil du texte (linéarité, continuité) vs partition (parties individualisées)
    - courbe plutôt qu'histogramme
  - Disposition, régularité
- Usages
  - analyse intratextuelle
  - topologie, topographie



Liste Graphique Aide

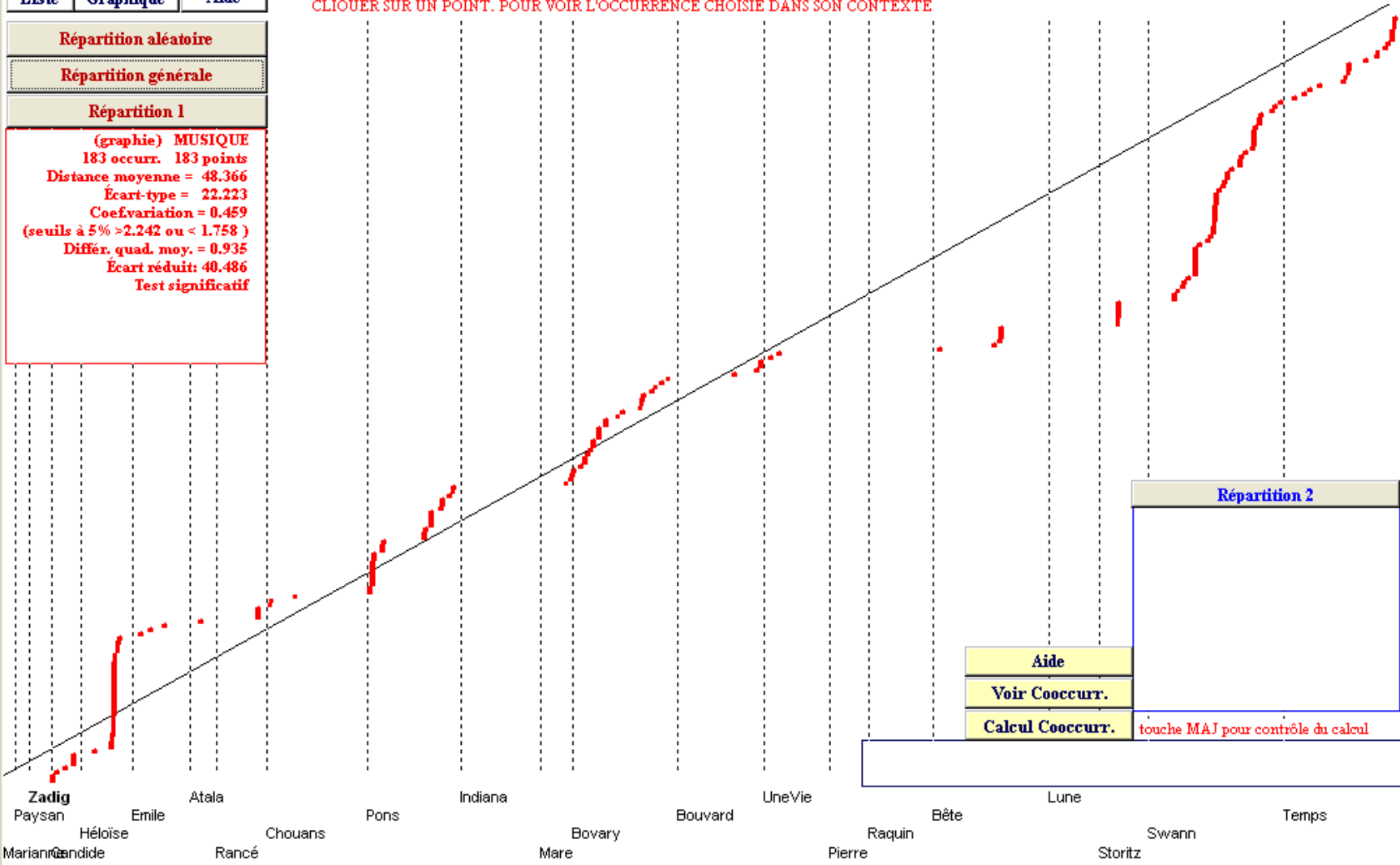
Répartition aléatoire

Répartition générale

Répartition 1

(graphie) MUSIQUE  
 183 occurr. 183 points  
 Distance moyenne = 48.366  
 Écart-type = 22.223  
 Coefvariation = 0.459  
 (seuils à 5% >2.242 ou < 1.758 )  
 Différ. quad. moy. = 0.935  
 Écart réduit: 40.486  
 Test significatif

Cliquer sur les boutons "Répartition 1" ou "Répartition 2" ou "Répartition générale". Dans ce dernier cas, le corpus entier est examiné et les résultats consignés dans un fichier. Déplacer les champs (en les draguant) si leur position est gênante. Le champ "Cooccurrences" n'est rempli que lorsque deux répartitions sont comparées.  
**CLIQUEZ SUR UN POINT. POUR VOIR L'OCCURRENCE CHOISIE DANS SON CONTEXTE**



Répartition 2

Aide  
 Voir Cooccurr.  
 Calcul Cooccurr.

touche MAJ pour contrôle du calcul



**Lexico 3**  
 Fichier Traitement Fenêtre

Navigation Rapport Dictionnaire

Sélectionnez une couleur : ■

Recherche : ami

Formes (ordre lexicométrique)	Fréquence
t	138
bougres	137
fois	136
ni	136
*france	133
aristocrates	133
eux	132
ma	132
notre	132
tant	132
roi	130
autre	129
républicains	129
va	129
ainsi	128
étaient	127
sera	126
mettre	125
sous	125
révolution	124
jusqu	123
traîtres	122
depuis	121
main	120
mes	120
soit	119
donner	118
te	118
avons	115
brigands	112
maintenant	109
là	107
terre	107
scélérats	105
amis	104
beau	104

11002 formes

Prêt

C:\Program Files\Lexico 3\Duchn.par

Sections découpées d'après le délimiteur: § - vue n° 1

Partition : 10 <Aucune>

seuillage peuple+ amis+ Spécifs

Section :  
 Occurrence :

\*marat n' est plus , foutre . **peuple** , génis , <Epg=2>pleure ton meilleur **ami** ; il meurt martyr de la liberté . c' est le \*calvados qui a vomi le monstre sous les coups duquel il vient de périr . une jeune fille , ou plutôt une furie armée par les prêtres , et pénitente , dit on , du cafard \*fauchet , part de \*caen pour exécuter cet horrible attentat . elle arrive à \*paris , et après avoir acheté un large couteau au \*palais que j tous les fripons de la terre , elle va trois jours consécutifs bougre épuisé de travail , ne pouvait voir ni entendre p oreille ; il croit que c' est une infortunée qui vient implor

jusqu' à la porte . les voisins accourent au bruit et saisissent la scélérate . la garde accourt ; toute la section de \*marseille entoure aussitôt la porte avec ses canons . § cette fatale nouvelle est bientôt répandue dans tout \*paris . les aristocrates sont au comble de la joie : les bons citoyens au désespoir

**Groupes de formes**

Nom du groupe : ami+

Le motif : ami

est le début de ce que je recherche

Ajouter

Rechercher  Enregistrer

Supprimer  Charger

Forme	Fréquence
amis	104
ami	32
amitié	4

Feuille 1

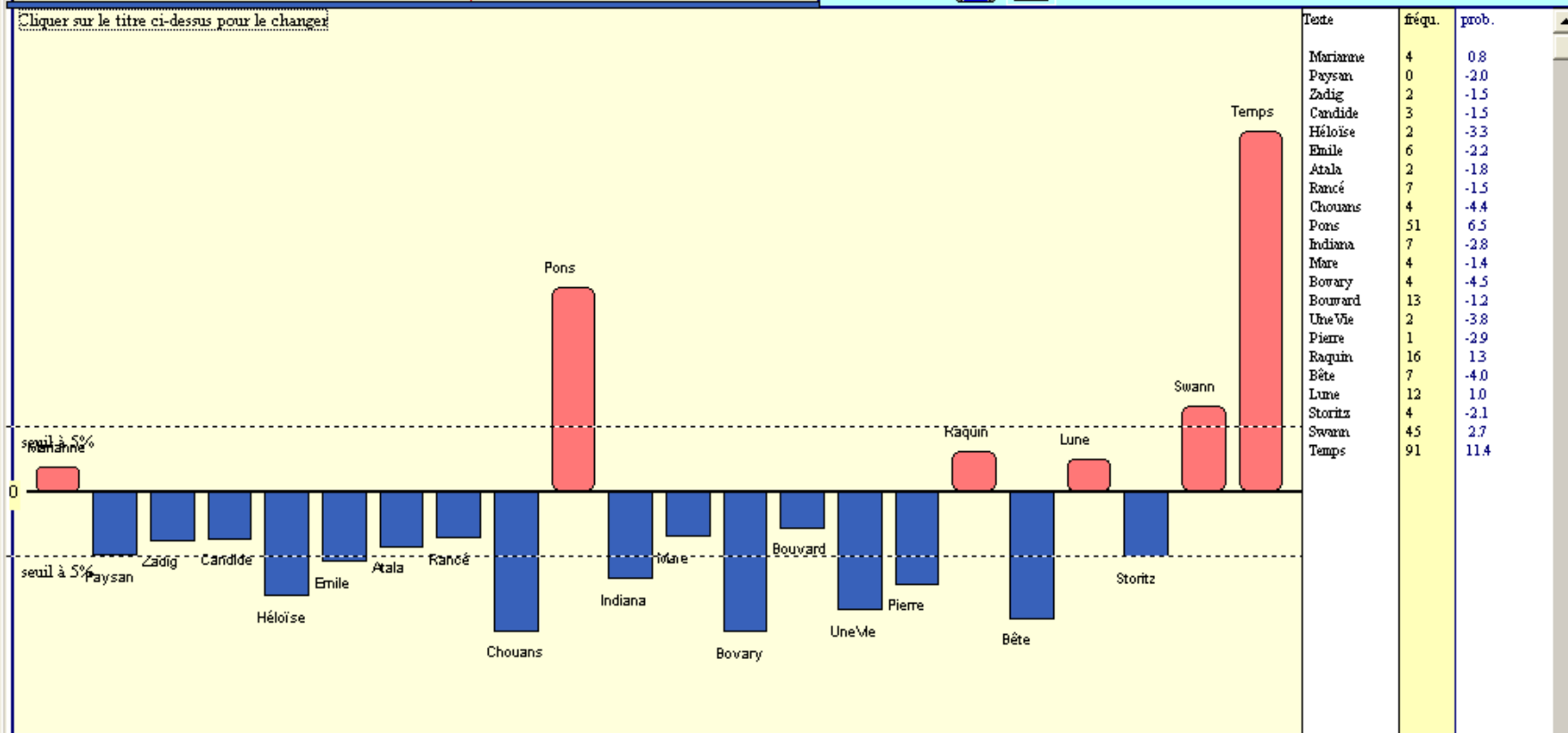
# 4. Typologie : Distribution

- Répartition contrastée d'unités dans les parties d'une partition
  - typiquement : calcul des spécificités, histogramme
  - deux voire trois entrées classiques : unité, partie, corpus (banalité)





oeuvre 287 occ.

■ déficits ■ excédents  Corrél. +0.322  
Seuil à 5% 0.42  

Cliquer sur le titre ci-dessus pour le changer





Zone réservée à une légende supplémentaire (cliquer)

Sommaire   simple  double  Retour

**Courbe sur quotient** Modifier les noms

Traiter une ligne de la liste:  
CLIC + MAJ sur un bouton  
Voir les contextes: CLIC  
un "bâton" du graphique

 Sous-ensembl  Mettre noms en couleurs

Marivaux, Voltaire, Rousseau, Chateaubriand, Balzac, Sand, Flaubert, Maupassant, Zola, Verne, Proust,

C:\HYPERBAS\EXEMPLE.EXE

Refaire résumé **Chouans** Voir les passages spécifiques et cacher les mots spécifiques

CLIC sur un mot: Recherche du mot dans les spécificités **corpus trié**

Trier Cherche Retour Sommaire **Pons** Cliquer sur un extrait pour voir le contexte

N°	écart	corpus	texte	mot	N°	écart	corpus	texte	mot
9	37.3	504	322	mille	10	37.3	601	600	pons
9	37.3	342	339	verneuil	10	37.3	516	512	cibot
9	37.3	294	259	marquis	10	37.3	495	495	schmucke
9	37.3	206	206	hulot	10	37.3	209	208	fraisier
9	37.3	199	198	chouans	10	33.8	158	157	présidente
9	35.4	172	172	francine	10	30.9	132	132	rémonencq
9	34.2	162	162	corentin	10	29.0	556	250	francs
9	32.9	189	172	commandant	10	28.9	120	119	poulain
9	28.4	370	207	marche	10	26.4	10332	1333	vous
9	28.4	117	117	gua	10	24.2	169	121	allemand
9	28.0	159	137	gars	10	23.9	84	84	marville
9	25.1	94	94	miche	10	23.2	940	269	monsieur
9	24.9	96	95	pille	10	23.2	80	80	brunner
9	23.7	85	85	chouan	10	22.9	78	78	magus
9	23.6	98	92	fougères	10	22.3	97	86	che
9	22.4	77	77	chopine	10	22.0	73	73	cécile
9	21.5	81	77	galope	10	21.1	68	68	camusot
9	20.8	353	152	marie	10	20.4	130	91	portière
9	20.8	68	68	gudin	10	20.3	4412	636	dit
9	20.1	150	98	soldats	10	20.2	63	63	gaudissard
9	20.0	68	66	merle	10	19.5	59	59	popinot
9	19.8	62	62	montauran	10	18.4	1928	346	m
9	19.4	65	63	gérard	10	18.4	118	79	testament
9	18.4	168	95	bleus	10	17.8	1681	310	mme
9	17.9	4412	626	dit	10	17.5	139	81	tableaux
9	17.7	806	206	reprit	10	17.5	68	58	musicien
9	16.9	101	69	république	10	17.5	49	49	topinard
9	15.3	978	208	terre	10	15.9	698	169	pauvre
9	14.7	1277	240	jeune	10	15.5	331	110	président
9	14.3	35	35	orgemont	10	15.3	912	192	mille
9	14.0	34	34	barbette	10	15.3	207	86	théâtre
9	13.9	38	36	pèlerine	10	15.3	41	40	élie
9	13.7	551	136	écria	10	15.1	134	69	cousin
9	13.7	39	36	sulpice	10	14.9	37	37	cantinet
9	13.5	91	53	vallée	10	14.5	446	122	docteur
9	13.4	208	79	comte	10	14.0	74	49	succession
9	13.2	10332	1074	vous	10	14.0	33	33	fritz
9	13.1	68	45	bretagne	10	13.7	148	66	notaire

Titre du texte de gauche (un clic permet de montrer alternativement les mots ou les phrases spécifiques de ce texte)

# 4. Typologie : Evolution

- Partition avec structure chronologique
  - cf. Lexico 2 et 3 :
    - spécificités chronologiques
    - accroissements spécifiques

**Lexico 3 - [Principales caractéristiques de la partition : mois]**

Fichier Traitement Fenêtre

**Lexico 3**

Navigation Rapport Dictionnaire

Spécifs évolutives - mois Spécifs - Part : mois

Corpus de référence : 02, 03, 04, 05, 06, 07, 08  
 Parties sélectionnées : 06,  
 Spécificités  positives  négatives

Terme	Frq Tot.	Frq P...	Spécif
patriotes	152	60	18
*phélipotin	13	13	13
nouvelle	46	21	9
*vincent	15	10	7
conspirateurs	48	18	6
les	4748	687	6
ai	202	48	6
mes	120	30	5
accuser	10	7	5
contre	298	60	5
diviser	16	9	5
dis	34	13	5
aristocrates	133	32	5
généraux	33	11	4
rues	19	8	4
intrigants	73	18	4
je	979	152	4
cartes	32	11	4
talons	19	8	4
meilleurs	40	13	4
réjouissance	13	6	4
connaît	13	6	4
est	1232	190	4
alors	84	20	4
victoire	30	11	4
là	107	26	4
comité	57	17	4
brouiller	21	8	4
on	856	137	4
salut	27	10	4
chat	55	15	4
fripons	96	22	4
déjà	64	14	3
revient	15	5	3
furent	15	5	3

Formes/SR	F	f	coeff	Partie(s)
patriotes	152	60	18	06
constitution	72	53	16	01 - 02
*paris	171	78	14	01 - 02
*phélipotin	13	13	13	06
vous	1097	705	13	01 - 04
avez	171	55	12	04
*custine	70	23	12	01 - 02
foutre	1100	784	10	07 - 08
phélipotins	14	14	10	07 - 08
*vincent	15	14	9	06 - 07
tu	296	200	9	03 - 06
*courtille	13	3	9	08
nouvelle	46	21	9	06
républicains	129	91	9	05 - 08
on	856	288	9	05 - 06
elle	368	90	9	01
nobles	42	21	9	01
*marat	65	27	9	01
je	979	561	8	04 - 07
sociétés	15	10	8	08
*cordeliers	17	6	8	08
ni	136	63	8	04 - 05
ai	202	114	8	04 - 06
procès	20	17	8	02 - 03
çà	8	8	8	02
argent	47	29	8	04 - 05
*ami	8	8	8	01
*gorsas	21	13	7	03
c	446	308	7	04 - 08
esclavage	20	9	7	08
*montagne	14	4	7	01
ton	63	43	7	03 - 05
*mayence	14	13	7	01 - 02

Princ. Caract. Lexicométriques Spécificités Chronologiques

Prêt C:\Program Files\Lexico 3\Duchn.par

Feuille 1

Spécifs  
 Sp Diac.  
 AFC  
 CAH  
 Pareto  
 Ac  
 Sp. Tot

# 4. Typologie : Synthèses

- Relevés
- Positions
- Associations : *liens entre unités, entre parties, entre propriétés*
  - Séquences
  - Cooccurrences
  - Analogies
  - Corrélations

# 4. Typologie : Séquences

- Repérage d'enchaînements récurrents, de figements. Exemples :
  - segments répétés
  - inventaire distributionnel
- Usage
  - construction d'unités endogènes
  - motifs



**Lexico3 - [Concordances]**

Fichier Traitement Fenêtre

**Lexico 3**

Navigation Rapport Dictionnaire

Spécifs évolutives - mois

Spécifs - Part : mois Segments répétés

Sélectionnez une couleur :

Lg	Segment	Frq
5	grande colère du *père *duchesne	55
5	les ennemis de la république	12
5	ses bons avis aux braves	23
5	les brigands de la *vendée	10
5	tous les ennemis de la	11
5	perdre le goût du pain	12
5	tous les jean foutres qui	26
5	foutre la grande joie du	32
5	de la poudre aux yeux	15
5	de la tête aux pieds	12
5	sa grande joie de voir	12
5	donner la clef des champs	11
5	*sans culottes pour qui ils	14
5	donner le coup de grâce	13
5	c est ainsi foutre que	10
5	du *père *duchesne au sujet	19
5	le vert et le sec	17
5	le manche après la cognée	13
5	du *père *duchesne de voir	43
5	la liberté et l égalité	11
5	à chien et à chat	29
5	le coup de grâce à	13
5	à feu et à sang	10
5	la tête à la fenêtre	16
5	c est vous qui avez	10
6	jeter de la poudre aux yeux	10
6	la grande colère du *père *duches...	54
6	il n y a pas de	15
6	l âge d un petit chien	13
6	la pluie et le beau temps	16
6	la grande joie du *père *duchesne	36
6	le vert et le sec pour	15
6	du *père *duchesne au sujet de	10
6	du *père *duchesne de voir que	35
7	la grande colère du *père *duches	14

Forme : braves Tri : Après Regroupement : <Aucun> Largeur : 40

Expression rationnelle Type de documentation : Inventaire distributio Délimiteurs de séquence : ;:1?/\_-^""000\$

25 ----- braves \*montagnards

4 ----- braves \*montagnards qui

2 ----- braves \*montagnards qui avez

48 ----- braves \*sans culottes

2 ----- braves \*sans culottes ce

2 ----- braves \*sans culottes de

5 ----- braves \*sans culottes des

2 ----- braves \*sans culottes des départements pour

13 ----- braves \*sans culottes pour

3 ----- braves \*sans culottes qui

2 ----- braves \*sans culottes vous

30 ----- braves bougres

19 ----- braves bougres qui

4 ----- braves bougres qui combattent

3 ----- braves bougres qui combattent pour

2 ----- braves bougres qui combattent pour la

2 ----- braves bougres qui composent maintenant le comité de sûr

10 ----- braves bougres qui nous

10 ----- braves défenseurs

6 ----- braves défenseurs de la

3 ----- braves défenseurs de la patrie

3 ----- braves défenseurs de la république

2 ----- braves défenseurs il est

7 ----- braves gens

4 ----- braves lurons

2 ----- braves lurons des départements

6 ----- braves républicains

5 ----- braves soldats

5 ----- braves volontaires

29 segments

C:\Program Files\Lexico 3\Duchn.par

Feuille 1

# 4. Typologie : Cooccurrences

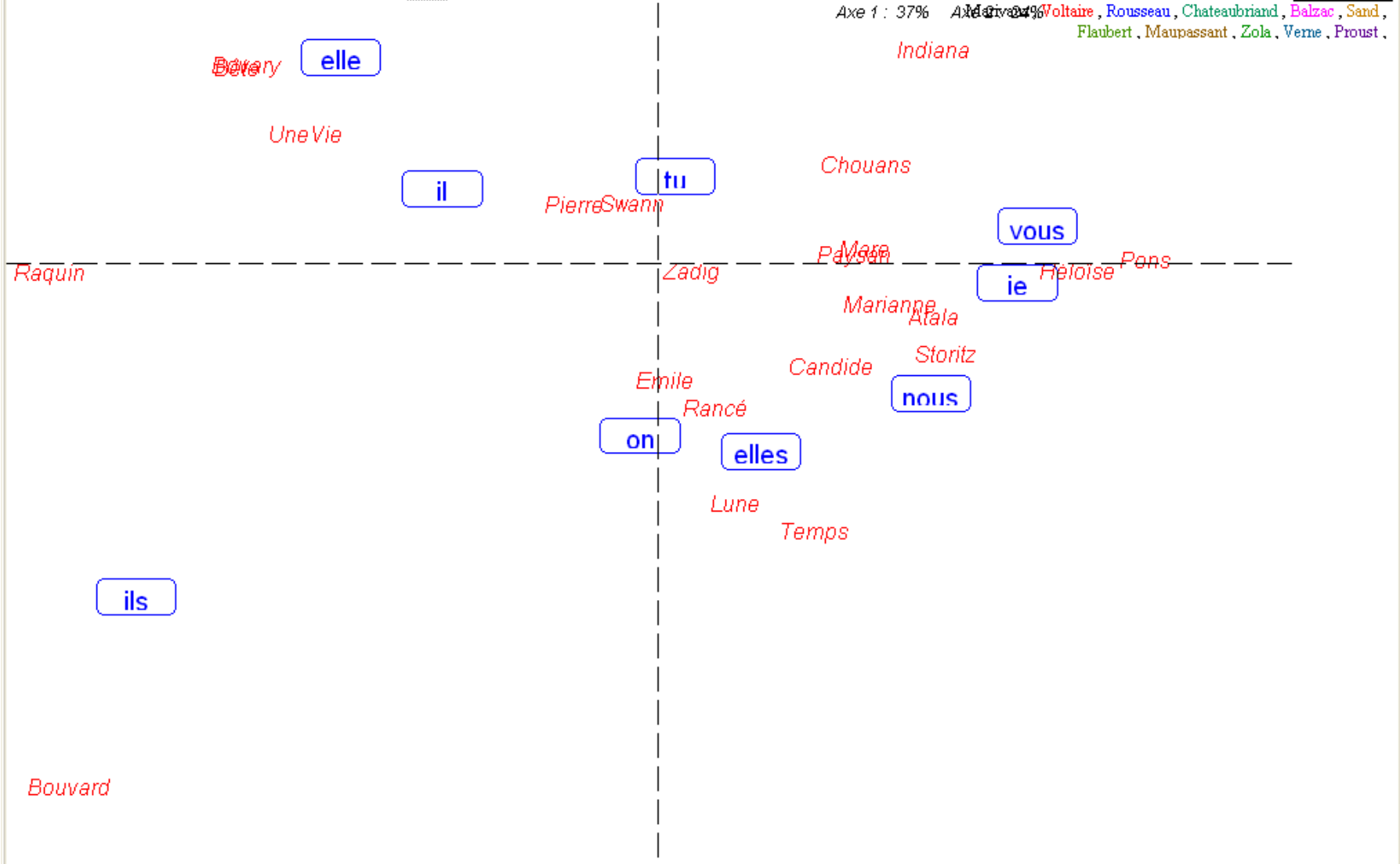
- Coprésence d'unités au sein d'un même contexte
  - orientée ou non
  - multiples formules (cf. aussi linguistique de corpus)
- Usages
  - associations plus souples que les Séquences

Clic sur un mot pour voir les contextes  
Clic +MAJ pour effacer un mot

HIERARCHIQUE				ALPHABETIQUE			
écart	corpus	texte	mot	écart	corpus	texte	mot
10.11	287	290	oeuvre	3.73	46	5	accomplir
10.11	270	38	art	4.02	84	7	admiration
10.11	105	43	chefs	2.99	84	5	allais
10.11	86	89	oeuvres	3.58	80	6	amours
10.10	446	52	chef	2.56	262	8	années
7.26	22321	370	d'	10.11	270	38	art
5.32	50	8	artistes	5.00	136	11	artiste
5.08	14	5	constamment	5.32	50	8	artistes
5.00	136	11	artiste	3.01	50	4	attendu
4.92	45	7	éventail	3.24	366	12	aujourd'
4.88	90	9	vinteuil	2.18	377	9	aurais
4.84	8	4	testateur	2.49	122	5	auteur
4.69	75	8	peinture	2.80	61	4	auxquelles
4.49	61	7	impressions	3.72	1090	27	avais
4.37	13	4	bravo	2.31	602	13	beau
4.29	70	7	écrivain	3.46	241	10	beauté
4.29	47	6	musée	2.98	164	7	belles
4.18	1782	41	vie	2.72	65	4	bergotte
4.17	3668	70	ces	3.47	57	5	bibliothèque
4.13	78	7	magus	2.09	110	4	bonnes
4.11	17	4	littéraire	4.37	13	4	bravo
4.06	56	6	collection	2.77	284	9	celles
4.05	18	4	génies	2.20	201	6	certain
4.02	84	7	admiration	2.27	142	5	certaines
3.95	61	6	piano	4.17	3668	70	ces
3.90	21	4	peintres	2.57	74	4	charité
3.81	43	5	niveau	10.10	446	52	chef
3.73	72	6	sentais	10.11	105	43	chefs
3.73	46	5	accomplir	4.06	56	6	collection
3.72	1090	27	avais	2.40	7267	102	comme
3.70	47	5	ressemblance	2.42	176	6	compris
3.70	47	5	desquels	2.64	70	4	conditions
3.64	108	7	peintre	5.08	14	5	constamment
3.58	80	6	amours	2.58	114	5	cousine
3.47	57	5	bibliothèque	7.26	22321	370	d'
3.46	241	10	beauté	2.77	99	5	dernières
3.42	437	14	mettre	2.04	114	4	derniers
3.40	164	8	êtres	2.33	16883	219	des
3.34	170	8	genre	3.70	47	5	desquels
3.33	37	4	toiles	2.11	580	12	donner
3.30	458	14	nature	4.29	70	7	écrivain
3.30	38	4	mémoires	3.22	41	4	élie
3.28	39	4	sieur	3.01	50	4	essence
							352 guerre
							082 victoire
							1011 pendant
							068 allemagne
							071 ministre
							297 france
							196 général
							259 françoise
							065 brichot
							050 début
							085 ennemi
							208 charlus
							911 avant
							139 croix
							206 hulot
							500 nouveau
							810 saint
							041 jardinier
							262 années
							077 allemands
							046 rappelant
							1096 depuis
							086 révolution
							087 journaux
							051 voulaient
							8752 avait
							097 bloch
							097 armes
							060 courte
							199 chouans
							101 république
							103 anglais
							151 lesquels
							803 assez
							105 chefs
							106 circonstance
							211 mauvais
							530 longtemps
							335 affaire
							070 venise
							077 robert
							184 peuple
							042 condamné

# 4. Typologie : Analogies

- Similarités
  - entre unités, entre parties
  - mesure et caractérisation qualitative
- Configuration
  - cartographie : AFC, arborée, kohonen...
  - classification (non supervisée) : hiérarchique ou non, ascendante ou descendante...



# 4. Typologie : Corrélations

- Force et orientation du lien entre deux propriétés, caractérisées par leurs valeurs au fil des occurrences
  - cf. analyse statistique
  - perspective

# 4. Typologie : Données

- Initialisation
  - Profil : *paramétrage par défaut, traces...*
  - Interfaçage : *importation, exportation*
  - Session : *ouverture/fermeture avec état mémorisé*
- Constructions
- Vue courante

# 4. Typologie : Données

- Initialisation
- Constructions
- Vue courante : des paramètres généraux ?
  - Fond
  - Structure
  - Localisations
  - Plan d'analyse
  - Propriétés d'affichage
  - Focus
    - unifié, persistance



## 4. Typologie : Vue courante

- Fond : sous-corpus, référence pour les calculs
- Structure : partition, regroupements, parallélisme
- Localisations : constitution et forme des adresses
- Plan d'analyse : unités : segmentation, propriété définissant les types
- Propriétés d'affichage
- Focus : objet de la recherche

# 4. Typologie : Données

- Initialisation
- Constructions
  - Propriétés : *dynamique de la description, ex. combinaisons d'étiquettes*
  - Sélections : *moteur de recherche, gestionnaire de sélection*
  - Scénarios : *cf. scripts Sato*
- Vue courante

# 4. Typologie : Analyse

- Gestion de la charge
  - progression du calcul (jauge), possibilité d'interruption
  - information et possibilité d'échantillonnage (ex. Xaira)
- Visualisations
- Organisation
- Annotation
- Archive

# 4. Typologie : Analyse

- Gestion de la charge
- Visualisations
  - tableaux
  - graphiques
  - diagrammes
  - cartes
- Organisation
- Annotation
- Archive

# 4. Typologie : Visualisations

- Tableaux
  - lignes, colonnes : déplacement, suppression, fusion...
  - valeurs : recodage, normalisation...
- Graphiques
  - Courbe, diagramme en bâtons, camembert
- Diagrammes
  - graphes, dendogrammes...
  - + Aide à la lecture, styles d'édition
- Cartes

# 4. Typologie : Analyse

- Gestion de la charge
- Visualisations
- Organisation
  - filtrage, regroupements, tris
  - synopse et comparaison : projection, mouvement
- Annotation
- Archive

# 4. Typologie : Analyse

- Gestion de la charge
- Visualisations
- Organisation
- Annotation
  - commentaire : libre, signet -non indexé
  - édition du corpus : correction, enrichissement
- Archive

# 4. Typologie : Analyse

- Gestion de la charge
- Visualisations
- Organisation
- Annotation
- Archive
  - journal : systématique ; pour scénarios
  - rapport : édition sélective, documentée



# 4. Typologie : Analyse

- Gestion de la charge
- Visualisations
- Organisation
- Annotation
- Archive
  - journal : systématique ; pour scénarios
  - rapport : édition sélective, documentée

# Bilan : innovations ?

- Données :
  - modélisation textométrique des textes :  
construction de propriétés, fond, plan d'analyse...
  - dynamique : corpus, propriété
  - transversalité

# Bilan : innovations ?

- Lecture, Synthèses
  - structuration différentielle
    - optimiser
  - unifier
  - localement :
    - mesures
    - évolution
    - corrélations

# Bilan : innovations ?

- Analyse
  - transversalité des visualisations
  - intégration des procédures d'exploitation des résultats (du résultat à la réponse)
  - annotation : Sato ; CorTeCS, Analog ; ...
  - rapport *et* journal

# 5. Retour sur les objectifs

- Technique - développement de TXM
  - priorités
  - poids des usages : on implémente "les spécificités"
  - maturité technologique : zones
- Applicatif
  - grille et descriptifs en ligne

# 5. Retour sur les objectifs

- Théorie
  - organisation, mais pas de clôture ?
  - innovations
    - on les positionne
    - pas (encore) d'induction ?