

SYLED - CLA2T

Université de la Sorbonne Nouvelle - Paris 3



Outils de statistique textuelle

Cédric Lamalle

William Martinez

Serge Fleury

André Salem

Manuel d'utilisation abrégé

(Dix premiers pas avec Lexico3)

Andrea Kunцова

Aude Maisondieu

Installation

1 - Avertissement

Il se peut que malgré tout le soin que nous avons apporté à la préparation de cette version, quelques erreurs subsistent encore. Nous vous demandons de signaler les éventuelles anomalies à :

Lexico3 / ILPGA

19, rue des Bernardins
75005 Paris, France

Joignez à votre envoi : le corpus de texte sur lequel vous aurez constaté des dysfonctionnements ainsi que le fichier *atrace.txt* créé au moment de l'exploitation (dans le répertoire où se trouve le corpus analysé).

Configuration minimale

Windows 95

Processeur 486, 4Mo de mémoire vive

3 Mo libres sur le disque dur

Configuration conseillée

Windows 98

Pentium II, 8Mo de mémoire vive

3 Mo libres sur le disque dur

Lexico3 fonctionne sous Windows 95, et Windows NT 3.51 et 4.0.

Nous conseillons vivement de regrouper programme et corpus sur le disque dur.

2 - Installer le logiciel

Pour installer LEXICO3

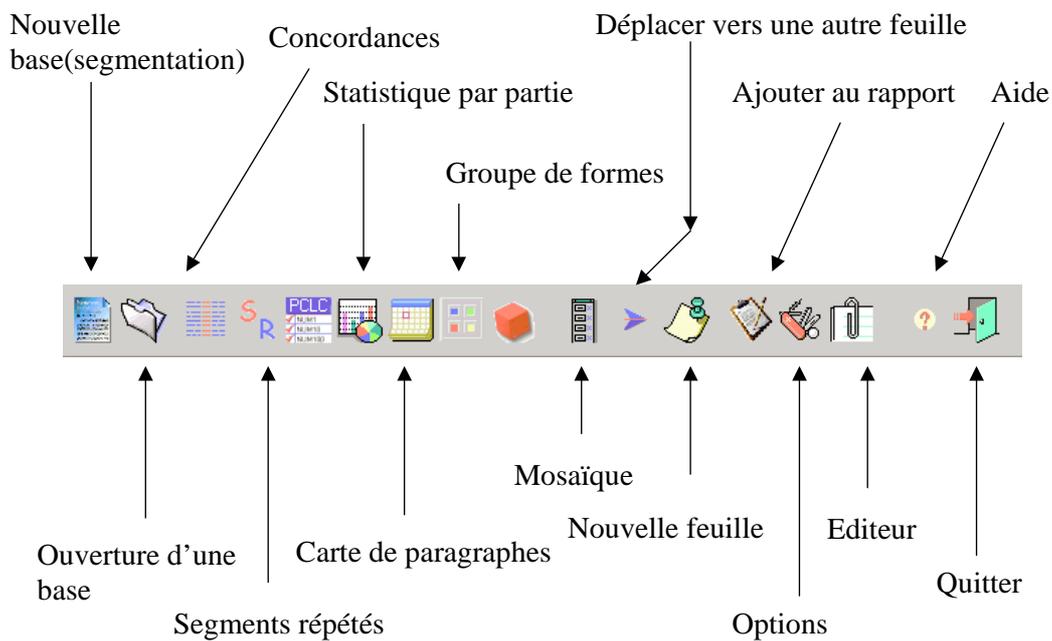
Insérer le CD ROM dans votre lecteur.

Double-cliquez sur l'icône du fichier **SETUP.EXE** qui se trouve sur ce CD-ROM.

Laissez vous guider par le programme d'installation

Le message : **Lexico3 a été installé** signale la fin de l'installation.

Barre des outils



Les 10 premiers pas avec Lexico3

Cette section est une initiation rapide à **Lexico3**.

Les différentes fonctions sont décrites en détail dans la partie **4**.

Ouvrir le logiciel en cliquant sur l'icône →



Pas n° 1 : Intégrer un texte dans la base



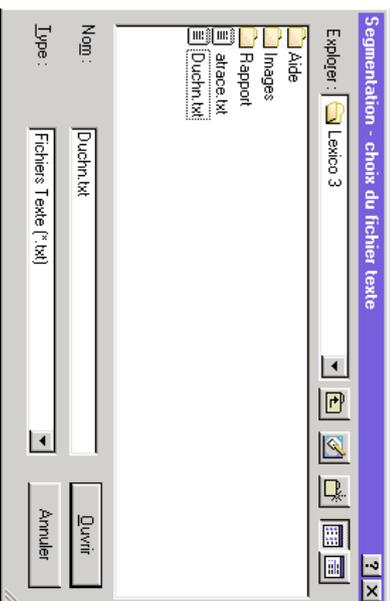
1.1... Pour segmenter le fichier-texte, cliquez sur l'icône

Nouvelle base →

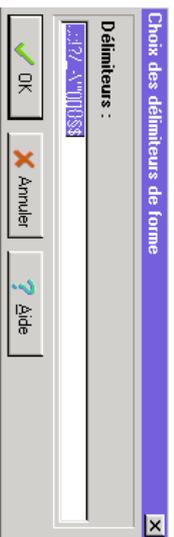


une boîte de dialogue s'ouvre pour vous permettre de sélectionner un texte.

1.2 Double-cliquez sur le fichier-texte *Duchn.txt*



1.3 **Lexico3** propose une liste de délimiteurs par défaut



1.4 Acceptez cette liste en cliquant sur le bouton **OK**
Lexico3 segmente le texte et affiche la liste des formes du corpus (dans la fenêtre de gauche).

Pas n° 2 : Réaliser une concordance

- 2.1 Cliquez sur l'icône *concordance* →
- 2.2 Sélectionnez la forme (dictionnaire) ou le segment répété dont vous souhaitez étudier les contextes,



glissez-la dans la fenêtre de droite, lâchez.

Glisser/Déposer
 Sélectionnez une forme - clic gauche de la souris.
 Maintenez le bouton gauche enfoncé et faites glisser la forme sélectionnée vers l'endroit souhaité puis déposez (lâchez le bouton gauche).



Pas n° 3 : Créer une partition du corpus

3.1 Cliquez sur l'icône de *partition de corpus* → une boîte de dialogue apparaît pour vous permettre de sélectionner la clé qui servira à partitionner le corpus



- 3.2 Sélectionnez, par exemple, la clé *semaine*
- 3.3 Validez cette partition en cliquant sur le bouton **Créer**
- 3.4 Sélectionnez dans le dictionnaire la forme dont vous souhaitez étudier les contextes, glissez-la dans la fenêtre de droite, puis lâchez-la dans la fenêtre du graphique de ventilation

Pas n° 4 : Constituer un groupe de formes

4.1 Cliquez sur l'icône *groupe de forme* →



4.2-a Entrez la forme que vous désirez chercher dans la zone de texte *Nouvelle entrée* (par exemple la séquence *patr* pour rechercher les formes qui commencent par ces quatre lettres) et cliquez sur le bouton *Rechercher* (ou retour-chariot).

Forme	Fréquence
patrolles	152
patrice	87
patroisme	20
patroide	17
patroiques	9
patroique	7
patiarche	3
patillon	2
patronne	2
patres	1

4.2-b Vous pouvez également glisser plusieurs formes à partir du dictionnaire ou d'une autre fenêtre de texte

4.3 **Lexico3** affiche la liste des formes du groupe.

4.4 Glissez sur le graphique la flèche rouge située à droite

du nom de groupe pour voir la ventilation du groupe.

Pas n° 5 : Créer une carte des sections

5.1 Cliquez sur l'icône *la carte des sections* →



une boîte de dialogue s'ouvre pour vous permettre de sélectionner des délimiteurs : paragraphe, point... qui partageront le texte en sections.

5.2 Sélectionnez, par exemple, le délimiteur paragraphe.

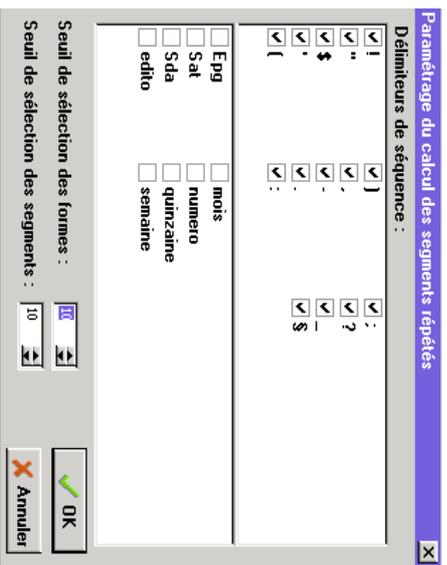
5.3 Cliquez ensuite sur le bouton **Créer**. Une fenêtre apparaît sur la droite dans laquelle chaque carré représente une section du texte (vous pouvez visualiser le découpage en parties en sélectionnant la partition *semaine* dans la boîte des partitions).

5.4 Sélectionnez (dans le dictionnaire) la forme dont vous souhaitez étudier les contextes, glissez-la dans la fenêtre de droite, puis lâchez.

Pas n° 6 : Créer un inventaire des segments répétés



- 6.1 Cliquez sur l'icône *Segments répétés* → une boîte de dialogue s'ouvre pour vous permettre de sélectionner les délimiteurs de segments (délimiteurs de formes et clés).



- 6.2 **Lexico3** vous propose des seuils de sélection et des délimiteurs par défaut.
- 6.3 Acceptez ces seuils en cliquant sur le bouton **OK**
- 6.4 **Lexico3** affiche la liste des segments répétés du corpus (dans une fenêtre de gauche).

Pas n° 7 : Calculer les spécificités d'une partie



- 7.1 **Rappel**: une Partition en semaine a été créée au pas 3
- 7.2 Cliquez sur l'icône *PCLC* → la fenêtre des "*Principales caractéristiques de la partition*" s'ouvre sur la droite.
- 7.3 Sélectionnez une partie du corpus dans la fenêtre



Partie	Document	Formes	Hapeax	Formax	Forme
✓ 111	3888	1255	840	116	de
✓ 112	5930	1532	1000	239	de
✓ 121	4909	1415	938	220	de
✓ 122	4778	1387	929	232	de
✓ 211	4314	1320	894	177	de
✓ 212	4857	1473	993	188	de
✓ 221	4756	1431	946	216	de
✓ 222	6415	1745	1140	294	de
✓ 311	4655	1361	897	211	de
✓ 312	4475	1395	979	214	de
✓ 321	4313	1329	911	189	de
✓ 322	3824	1639	1097	239	de
✓ 411	4352	1413	1003	192	de
✓ 412	4109	1231	816	169	de
✓ 421	4656	1432	970	188	de
✓ 422	4887	1523	1060	204	de
✓ 511	4333	1297	888	193	de
✓ 512	2934	965	680	117	de
✓ 521	6730	1644	1108	212	de
✓ 522	4817	1404	934	224	de



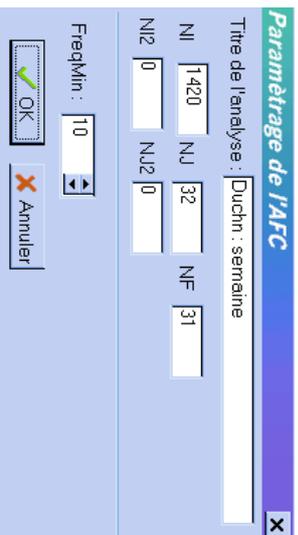
- 7.4 Cliquez sur le bouton *Spécif* → une boîte de dialogue s'ouvre qui vous permet de fixer un seuil et une fréquence minimale.
- 7.5 Acceptez le seuil et la fréquence proposés en cliquant sur le bouton **OK**
- 7.6 **Lexico3** affiche dans une colonne à gauche les spécificités par semaine de la forme sélectionnée

Pas n° 8 : Réaliser une Analyse Factorielle des Correspondances

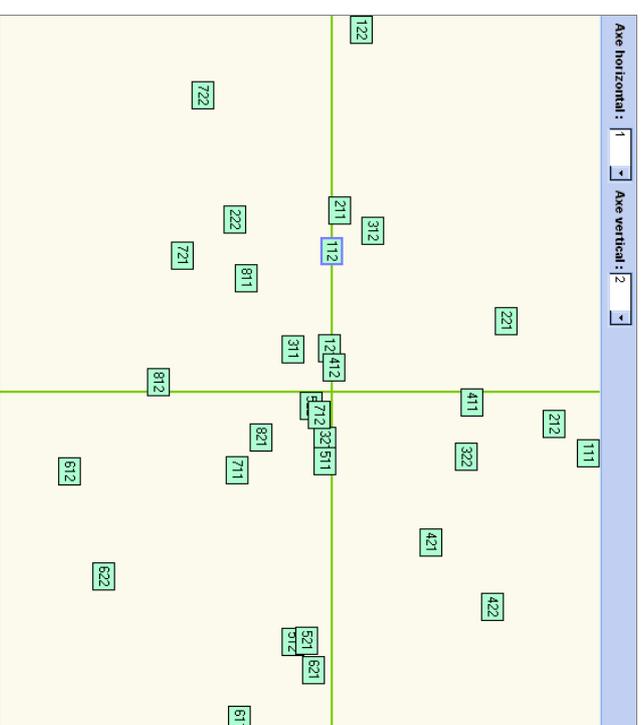
8



- 8.1 **Rappel:** une Partition en semaine a été créée au pas 3
- 8.2 Cliquez sur l'icône *PCIC* → une fenêtre de *Principales caractéristiques de la partition* s'ouvre
- 8.3 Cliquez sur l'icône *AFC* (à droite sur l'écran)
- 8.4 **Lexico3** propose un paramétrage de l'AFC par défaut



- 8.5 Acceptez ce paramétrage en cliquant sur le bouton **OK**
- 8.6 **Lexico3** affiche la typologie obtenue à partir de l'Analyse factorielle des correspondances sur le plan des deux premiers facteurs.

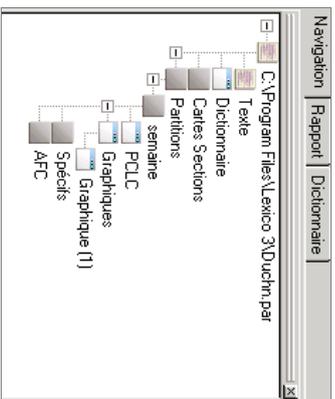


Pas n° 9 : Naviguer parmi les résultats

9



L'onglet *Navigation* vous permet de naviguer parmi les différents résultats produits par **Lexico3**. Un clic sur l'une des unités de l'arborescence provoque l'affichage du graphique correspondant dans la fenêtre droite.

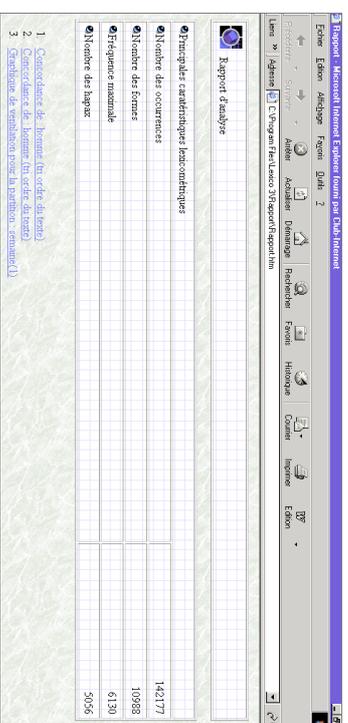


Pas n° 10 : Enregistrer des résultats dans le rapport

10



- 10.1 Cliquez sur l'onglet *Rapport* dans la fenêtre gauche
- 10.2 Sélectionnez et maximisez la fenêtre que vous désirez intégrer au rapport
- 10.3 Cliquez sur l'icône *Ajouter au rapport* →
- 10.4 Dans l'onglet *Rapport* cliquez sur le bouton *Enregistrer*.
- 10.5 Après avoir enregistré le rapport, vous pouvez le consulter en cliquant sur l'icône *Editeur* →
- 10.6 Le navigateur invoqué (Netscape ou Internet Explorer) vous permet ensuite d'imprimer chacun des graphiques sauvegardés dans le rapport.



Fin des dix premiers pas avec Lexico3.

Vous pouvez envisager une promenade plus longue.