

ATHIS. L'historien, le texte et l'ordinateur. Les logiciels de traitement informatique du texte.

Alain Dallo
Université Paris I. LAMOP

1. Des logiciels qui ont une histoire. L'utilisation de corpus.	4
1.1. Plusieurs centres de recherche	4
1.2. Des avancées récentes	4
1.3. Corpus, logiciels, approches	4
2. L'approche thématique	5
2.1. La forme et sa distribution dans le corpus	5
2.1.1. Index avec Hyperbase	6
2.1.2. Dictionnaire des formes avec Lexico 3	7
2.1.3. Dictionnaire des lemmes avec Weblex	8
2.1.4. Nooj : liste des formes et dictionnaire	9
2.1.5. Nooj : Liste des formes et dictionnaire	10
2.1.6. Calcul des rafales dans weblex	11
2.1.7. Graphique de la répartition de je	12
2.1.8. Rafale, topologie : Europe (Vème République)	13
2.1.9. Rafale, topologie : Europe chez Mitterand	14
2.1.10. Rafale, carte des sections : Europe	15
2.2. Les concordances	16
2.2.1. Concordances de la forme je chez Mitterand	17
2.2.2. Concordances du lemme croire	18
2.2.3. Concordances d'une liste de quelques formes avec Lexico 3	19
2.2.4. Les concordances de deux lemmes éloignés avec Weblex. Utilisation de CQP = Corpus Query Processor	20
2.3. Les segments répétés	22
2.3.1. Les segments répétés avec Lexico 3	23
2.3.2. Les segments répétés du corpus voeux. Weblex	24
2.4. Les cooccurrences	25
2.4.1. Les cooccurrences : Unité France	26
2.4.2. Les contextes courts de je	27
2.4.3. L'environnement de la forme je	28
2.4.4. Extrait de la liste de toutes les cooccurrences significatives	29
2.4.5. Cooccurrences droites et gauches de France	30
2.4.6. Graphique des cooccurrences de je	31
2.4.7. Résumé du graphique des cooccurrences de je	32

3.	<i>L'approche statistique ou lexicométrique</i>	33
3.1.	Les caractéristiques d'un corpus	33
3.1.1.	L'étude des formes en fonction de leur fréquence	33
3.1.2.	Le nombre d'occurrences du corpus et de ses composantes	33
3.1.3.	Fréquence de nous	34
3.1.4.	Les mots outils dans le tableau lexical entier	35
3.1.5.	Les vrais et les faux hapax	36
3.2.	Les spécificités d'une partie	37
3.2.1.	J.C. Spécificités positives	38
3.2.2.	J.C. Spécificités négatives	39
3.2.3.	J.C. Spécificités phrastiques	40
3.3.	Rapprochement et éloignement des textes d'un corpus	41
3.3.1.	L'analyse factorielle du dictionnaire	42
3.3.2.	AFC sur une liste de mots	43
3.3.3.	Le classement automatique hiérarchisé	44
4.	<i>Une autre approche en gestation</i>	45
4.1.	Lemmes et catégories grammaticales	45
4.1.1.	formes et lemmes	46
4.1.2.	formes et catégories	47
4.1.3.	concordance dnav n=France	48
4.1.4.	concordance (N, V, A) nom = France, verbe = être,	49
5.	<i>Quelques liens</i>	50
5.1.	Logiciels	50
5.2.	Corpus	50

1. Des logiciels qui ont une histoire. L'utilisation de corpus.

1.1. Plusieurs centres de recherche

- Charles Muller
- Étienne Brunet. Université de Nice
- ENS de Saint-Cloud
- Maurice Gross, Paris VII. Max Silberstein

1.2. Des avancées récentes

- Corpus étiquetés
- Préparation des corpus et balisage morpho-syntaxique
- Les moteurs de recherche, la fouille de textes et l'indexation fine de documents

1.3. Corpus, logiciels, approches

Partir du mot pour aller jusqu'au positionnement des textes d'un corpus, en passant par des études thématiques

- ♦ Un corpus : les vœux des présidents de la V^{ème} République.
 - Corpus mis en place sur Weblex dans le cadre de la thèse de Jean-Marc Leblanc
 - Corpus intégré dans les autres logiciels à partir des textes fournis par Damon Mayaffre
- ♦ Les logiciels : Hyperbase, Lexico3, Weblex et Nooj
- ♦ L'approche thématique ou documentaire
- ♦ L'approche statistique

2. L'approche thématique

De la forme isolée à son environnement

- L'étude du dictionnaire
- Les concordances d'une forme
- Les segments
- Les cooccurrences

2.1. La forme et sa distribution dans le corpus

- Un index avec Hyperbase
- Un dictionnaire des formes avec Lexico 3
- Dictionnaire des lemmes avec avec Weblex
- Le dictionnaire des lemmes repérés : Nooj. Images 1 et 2.
- Calcul des rafales des vœux dans Weblex et graphique de la répartition de je
- Rafales et distribution d'une forme avec Hyperbase. Images : 1 et 2
- Rafales et distribution d'une forme avec Lexico3

2.1.1. Index avec Hyperbase

C:HYPERBASVOEUX.EXE

Les sous-
fréquences dans
les textes sont
réparties en
binômes, dont le
premier élément
indique le
numéro d'ordre
du texte, le
second la
fréquence du
mot dans ce
texte.

Cliquer sur un
mot pour l'activer

Index hiérarchique

Chercher un mot

Graphique

Éditer

Quitter

Retour

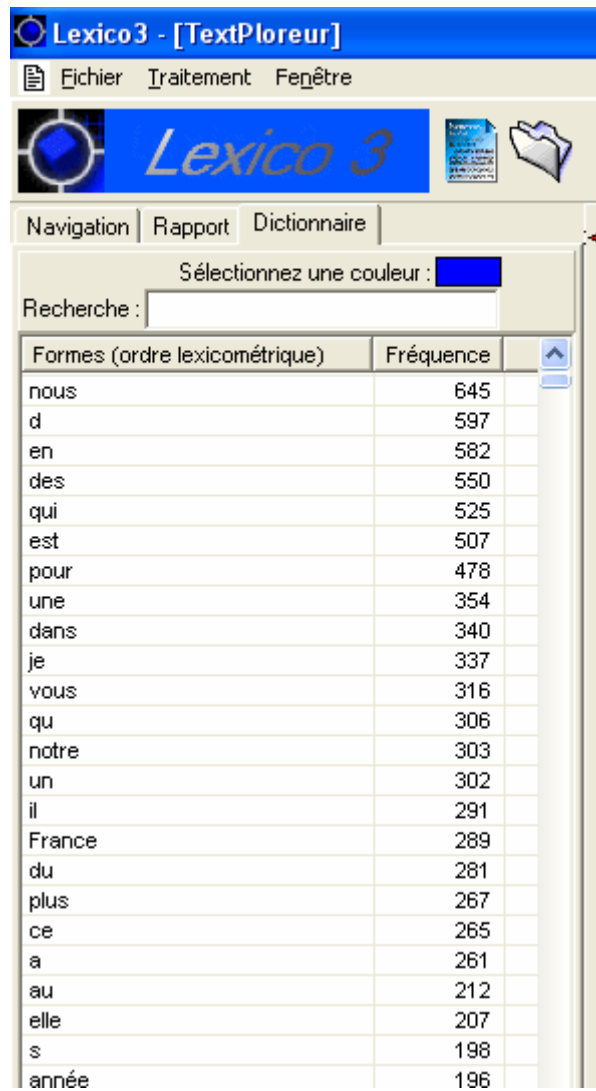
86 j'
4 jadis
1 jalousement
38 jamais
10 janvier
4 Japon
346 je
1 jeter
1 jettera
4 jeu
3 jeune
32 jeunes
11 jeunesse
1 jeux
11 joie
2 joies
1 joindre
1 joindront
1 joli
1 jouaient
2 joue
1 joué
3 jouer
1 jouera
1 joug
2 jouons
34 jour
1 journal
1 journées
10 jours
1 joyeusement
1 joyeux

N° 1 DG 31
N° 3 VG 80
N° 5 M2 64
TOUS LES TEXTES

N° 2 GP 28
N° 4 M1 62
N° 6 JC 81

je **fréquence totale: 346**
CLIQUEZ SUR UN TEXTE (ou sur TOUS) pour y repérer les contextes du mot "je"
Cliquer AILLEURS dans cette fenêtre pour l'effacer

2.1.2. Dictionnaire des formes avec Lexico 3



The screenshot shows the Lexico 3 software window titled "Lexico3 - [TextPloreur]". It has a menu bar with "Fichier", "Traitement", and "Fenêtre". Below the menu is a toolbar with icons for a search, a folder, and a document. The main area has tabs for "Navigation", "Rapport", and "Dictionnaire". The "Dictionnaire" tab is active, showing a search interface with a "Sélectionnez une couleur" dropdown and a "Recherche:" text box. Below this is a table of word forms and their frequencies.

Formes (ordre lexicométrique)	Fréquence
nous	645
d	597
en	582
des	550
qui	525
est	507
pour	478
une	354
dans	340
je	337
vous	316
qu	306
notre	303
un	302
il	291
France	289
du	281
plus	267
ce	265
a	261
au	212
elle	207
s	198
année	196

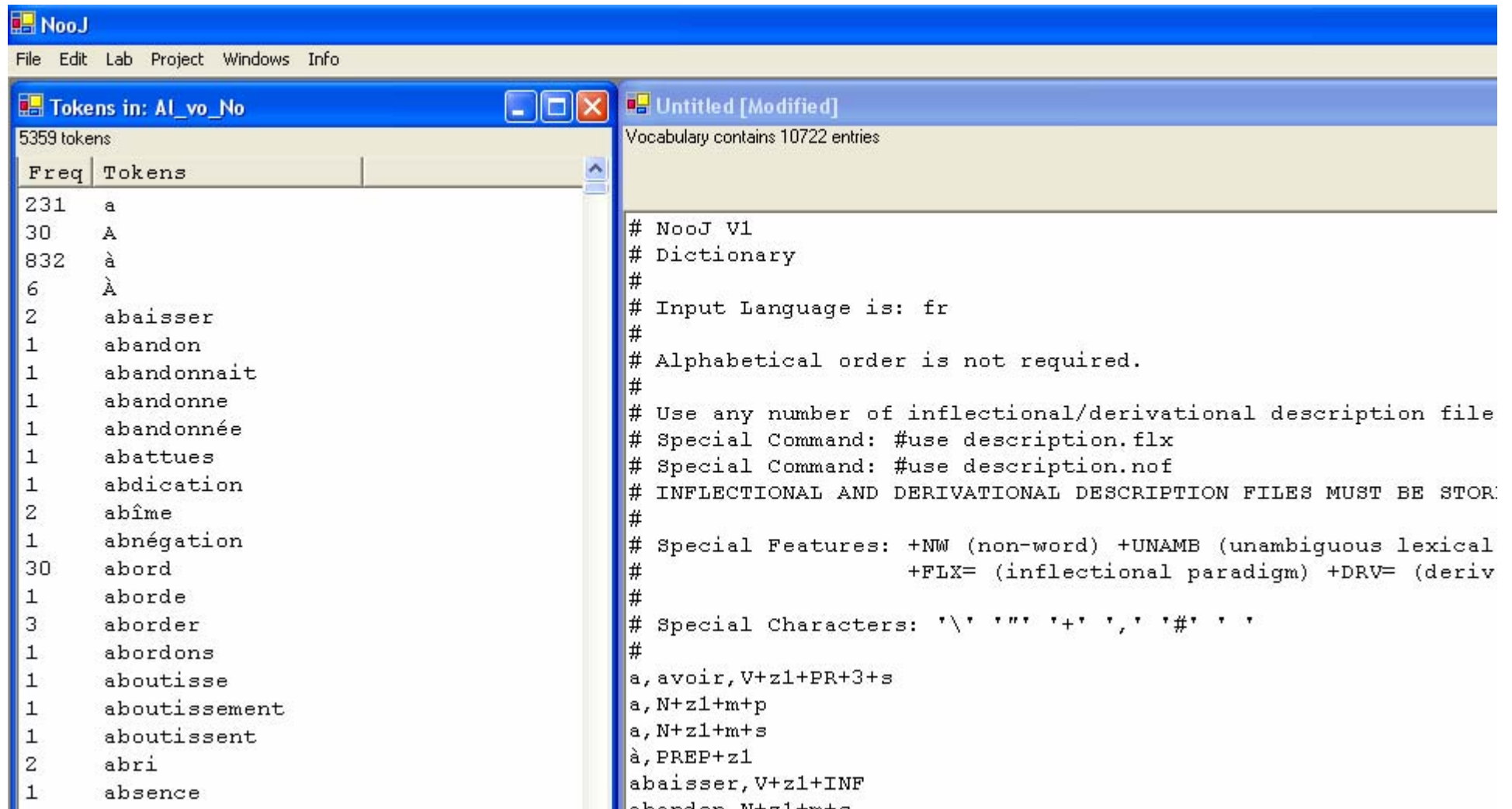
2.1.3. Dictionnaire des lemmes avec Weblex

Vocabulaire «élagué» du 1

p3s de fréquence supérieure ou égale à 3.

Liste alphabétique			Liste hiérarchique		
ord	p3	fréq.	p3	fréq.	
1	%	18	nous	654	
2	Afrique	14	je	426	
3	Airbus	4	France	300	
4	Algérie	21	français	219	
5	Allemagne	12	année	202	
6	Amérique	15	faire	183	
7	Angleterre	3	pouvoir	149	
8	Asie	6	devoir	125	
9	Berlin	3	pays	122	
10	Bosnie	3	on	113	
11	Canada	3	monde	102	
12	Chine	4	Europe	99	
13	Douze	3	vouloir	92	
14	Europe	99	voeu	88	
15	France	300	bon	87	
16	Gaulle	3	grand	85	
17	Gorbatchev	3	nouveau	85	
18	Grande-Bretagne	3	dire	82	
19	Guerre	3	peuple	82	
20	Irak	4	an	80	
21	Israël	3	souhaiter	79	
22	Italie	4	savoir	77	
23	Japon	4	aller	75	
24	Koweït	5	falloir	68	
25	Liban	4	vivre	68	
26	Moyen-Orient	4	état	67	
27	Nations-Unies	5	cher	66	
28	Nouvel	12	vie	66	
29	Noël	3	paix	64	
30	Outre-Mer	9	compatriote	63	
31	Paris	8	politique	61	
32	Pologne	4	République	60	
33	Russie	5	sociét	59	

2.1.4. Nooj : liste des formes et dictionnaire



The screenshot shows the NooJ software interface. The top menu bar includes File, Edit, Lab, Project, Windows, and Info. Below the menu bar, there are two main panes. The left pane, titled 'Tokens in: Al_vo_No', displays a list of 5359 tokens. The right pane, titled 'Untitled [Modified]', shows the contents of a dictionary file.

Token List (Left Pane):

Freq	Tokens
231	a
30	A
832	à
6	À
2	abaisser
1	abandon
1	abandonnait
1	abandonne
1	abandonnée
1	abattues
1	abdication
2	abîme
1	abnégation
30	abord
1	aborde
3	aborder
1	abordons
1	aboutisse
1	aboutissement
1	aboutissent
2	abri
1	absence

Dictionary File (Right Pane):

```
# NooJ V1
# Dictionary
#
# Input Language is: fr
#
# Alphabetical order is not required.
#
# Use any number of inflectional/derivational description file
# Special Command: #use description.flx
# Special Command: #use description.nof
# INFLECTIONAL AND DERIVATIONAL DESCRIPTION FILES MUST BE STOR
#
# Special Features: +NW (non-word) +UNAMB (unambiguous lexical
#                   +FLX= (inflectional paradigm) +DRV= (deriv
#
# Special Characters: '\ ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '
#
a, avoir, V+z1+PR+3+s
a, N+z1+m+p
a, N+z1+m+s
à, PREP+z1
abaisser, V+z1+INF
abandon, N+z1+m+s
```

2.1.5. Nooj : Liste des formes et dictionnaire

NooJ											
File Edit Lab Project Windows Info DICTIONARY											
Untitled [Modified]											
Dictionary contains 10722 entries											
Entry	S-Lemma	Category	Genre	Nombre	P...	Pers	Sem	T...	Val	SynSem	
abaisser	abaisser	V	-	-		-		INF		z1	
abandon	abandon	N	m	s			-			z1	
abandonnait	abandonner	V	-	s		3		I		z1	
abandonne	abandonner	V	-	s		2		IP		z1	
abandonne	abandonner	V	-	s		1		PR		z1	
abandonne	abandonner	V	-	s		3		PR		z1	
abandonne	abandonner	V	-	s		1		S		z1	
abandonne	abandonner	V	-	s		3		S		z1	
abandonnée	abandonné	A	f	s						z1	
abandonnée	abandonné	N	f	s			-			z1	
abandonnée	abandonner	V	f	s		-		PP		z1	
abattues	abattre	V	f	p		-		PP		z1	
abattues	abattu	A	f	p						z1	
abattues	abattue	N	f	p			-			z3	
abdication	abdication	N	f	s			-			z1	
abîme	abîmer	V	-	s		2		IP		z1	
abîme	abîmer	V	-	s		1		PR		z1	
abîme	abîmer	V	-	s		3		PR		z1	
abîme	abîmer	V	-	s		1		S		z1	
abîme	abîmer	V	-	s		3		S		z1	

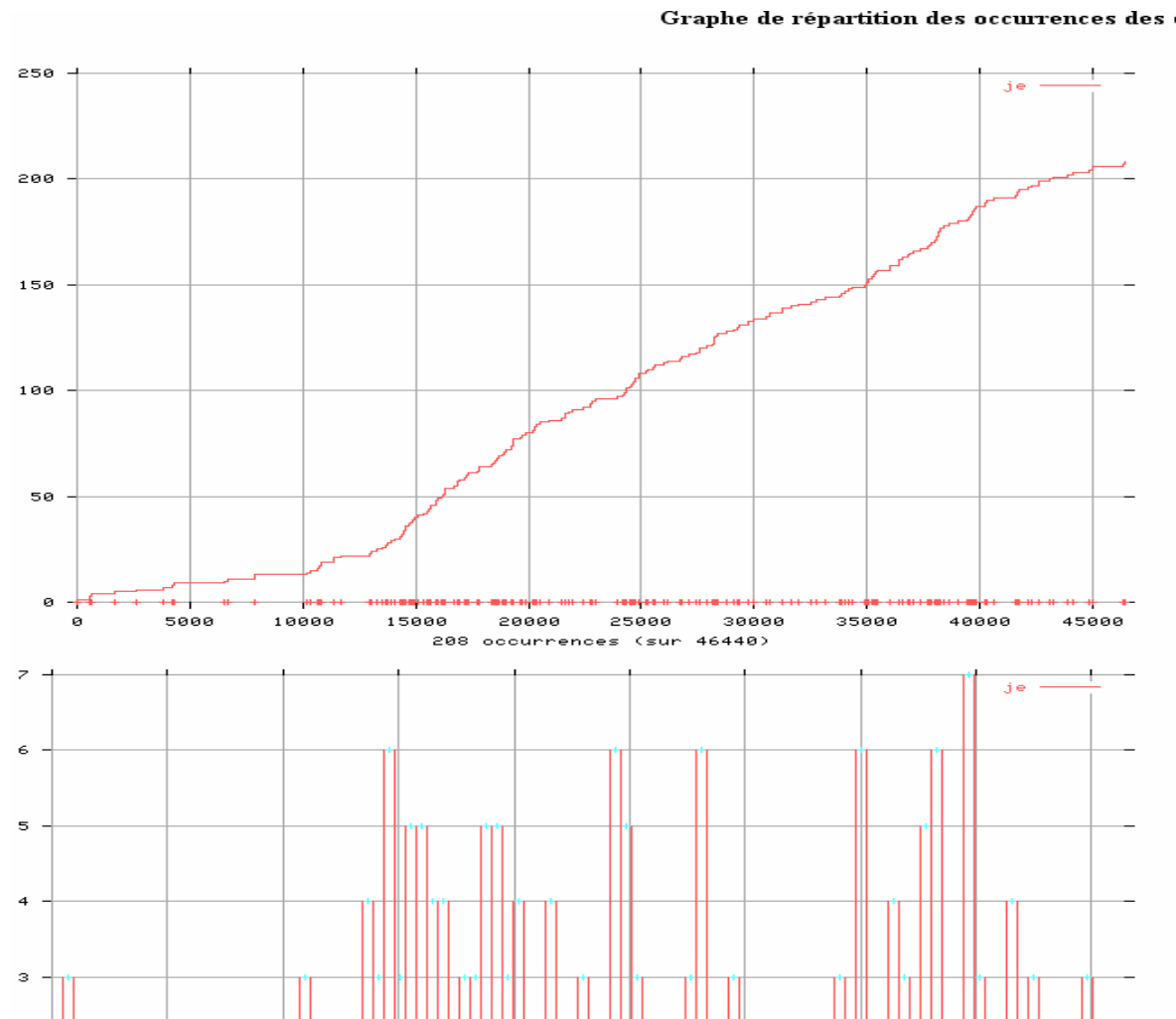
2.1.6. Calcul des rafales dans weblex

Répartition du vocal

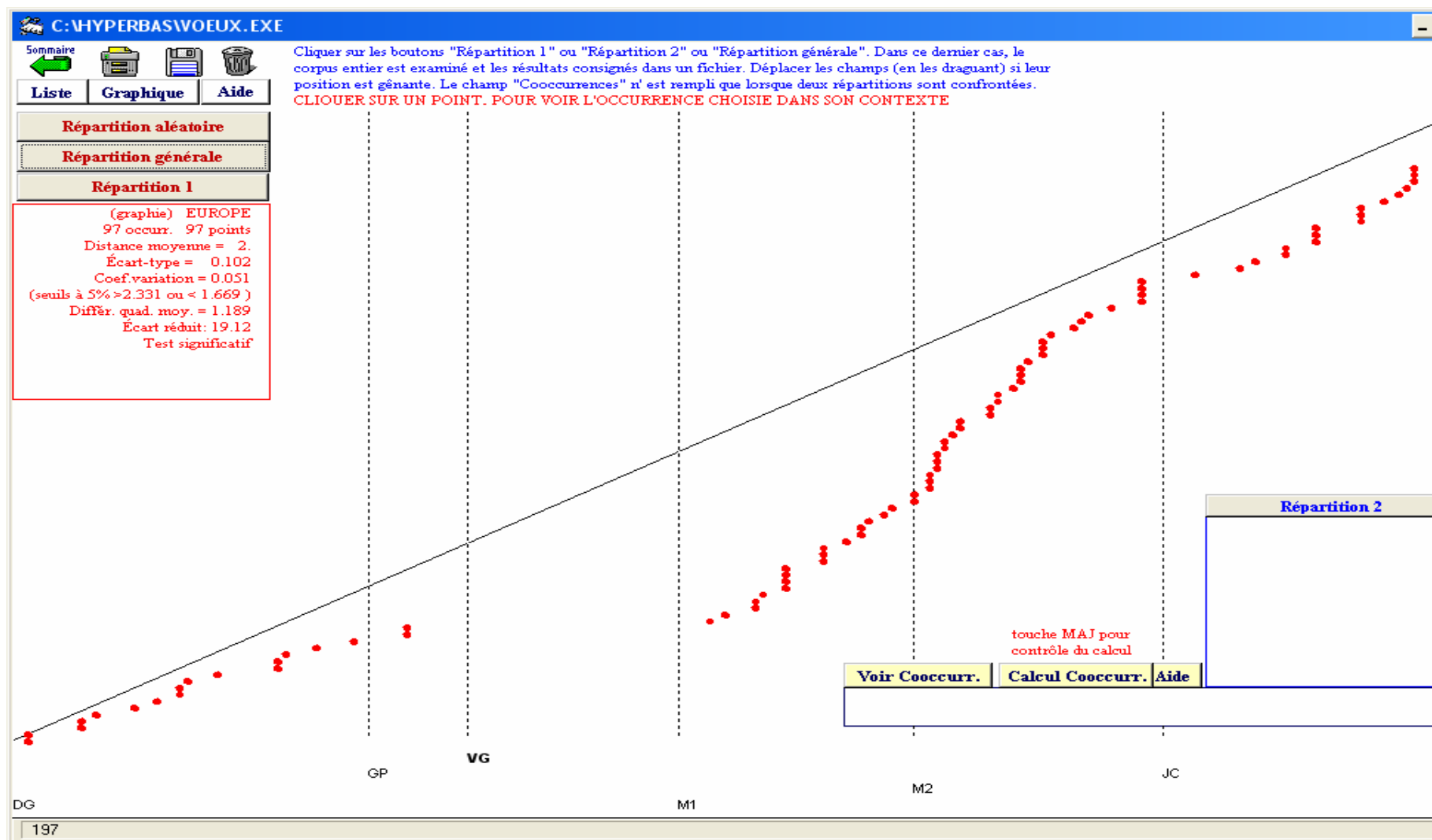
Z décroissant

forme	Z	fp
1 compatriote	20.8582	63
2 euro	20.145	13
3 je	17.1565	376
4 Europe	14.4225	91
5 Algérie	14.1149	19
6 nier	13.8835	10
7 peur	13.8149	8
8 cher	13.0681	65
9 scientifique	12.8823	10
10 nous	12.0972	483
11 but	11.8479	10
12 progrès	11.6406	51
13 ami	11.5533	11
14 développement	10.6219	22
15 difficultés	10.6133	19
16 soldat	9.92321	10
17 vive	9.28211	53
18 apporter	9.22639	23
19 hausse	8.88236	6
20 au-dehors	8.38362	7
21 %	8.18859	10
22 bonheur	8.11869	20
23 valeurs	8.08423	11
24 démocratie	8.01422	16
25 fête	7.25487	7

2.1.7. Graphique de la répartition de je



2.1.8. Rafale, topologie : Europe (Vème République)



2.1.9. Rafale, topologie : Europe chez Mitterrand

Sommaire	Retour	N°	Mots	254	Lettres	1338	Page		Ecart	Textes	Cherche	Notes		page	CLIC sur un mot pour voir les contextes
		5		M2			135								

Depuis quelques mois , nous avons dans les yeux les images terribles des combats que se livrent les peuples de Yougoslavie , hier encore associés sous un même drapeau . Comment arrêter cette guerre ?

La France soutient les efforts de négociation et d' arbitrage de la Communauté et des Nations - Unies . Elle reconnaît le principe de l' autodétermination . Mais il lui paraît urgent que soient mises en place des structures inter - européennes , où le droit à l' indépendance ne se confondra pas avec l' anarchie des tribus d' autrefois . Ce sera , je le pense , l' un des enjeux majeurs de 1992 .

Car l' inquiétude gagne l' **Europe de l' Est où l' on redoute** la contagion . Comment cette inquiétude nous épargnerait - elle , nous qui , à l' Ouest , avons pourtant la chance de vivre en paix et d' avoir dépassé nos propres divisions ? Raison de plus de se réjouir des récents accords de Maastricht . Une monnaie commune , l' amorce d' une diplomatie , d' une défense et d' une armée communes à l' Europe des Douze , une charte sociale , l' exemple de stabilité offert aux peuples qui se déchirent , bientôt 350 à 360 millions d' Européens solidaires sur la scène du monde - en attendant les autres - , bref , l' Europe qui se fait , voilà un grand dessein , capable d' enthousiasmer , de rassembler et de justifier l' espérance .

2.1.10. Rafale, carte des sections : Europe

Navigation | Rapport | Dictionnaire

Sélectionnez une couleur : ■

Recherche :

Formes (ordre lexicographique)	Fréquence
Caire	1
Calédonie	1
Cambodge	2
Charles	1
Constantine	1
Edimbourg	1
Edouard	1
Erika	1
Etats	10
Europe	97
Finul	1
France	289
France	7
Galatée	1
Genève	2
Gorbatchev	3
Grande	1
Guadeloupe	1
Helsinki	1
Hollande	2
Indochine	1
Isère	1
Isle	1
Ismailia	1
Israël	3
Italie	4
Kiev	1
Kinshasa	1
Kosovo	2
Koweït	4
Laurent	1
Liban	4
Lille	1
Londres	1
Louvre	1
Luxembourg	1
Maastricht	1
Marseille	1
Mikhaïl	1
Moscou	2
Moscowa	1
Nam	1
Nations	6
Nice	1
Nouvelle	1
ONU	1
Pankov	1

Partition : 10 <Aucune> ☐ seuillage Europe ☐ Spécifs

Section :

l' audace est celle de l' **Europe** , puisque demain , premier janvier 1993 , les frontières dis communauté .

2.2. Les concordances

L'environnement proche d'une forme

- **Les concordances d'une forme : Lexico 3**
- **Les concordances d'un lemme avec Hyperbase**
- **Les concordances d'une liste de formes**
- **Les concordances de deux lemmes éloignés avec Weblex. Utilisation de CQP = Corpus Query Processor**

2.2.1. Concordances de la forme je chez Mitterrand

Lexico 3

Navigation | Rapport | Dictionnaire

Sélectionnez une couleur : je

Recherche : je

Formes (ordre lexicographique) | Fréquence

interventions	1
interviendrai	1
interview	1
intimider	2
intolérance	1
intoxication	1
intuition	1
inutile	3
inutilement	2
inutiles	2
invasion	1
invectives	2
invente	1
inventé	1
inventer	4
inventer	1
inventerons	1
inventif	1
inventifs	1
invention	1
inventions	1
inventons	1
investir	3
investissement	3
investissements	2
investissent	1

Forme : Tri : Après Regroupement : psdt Largeur : 40

☐ Expression rationnelle Type de documentation : Concordance Délimiteurs de séquence : .:!?/_\''"000\$

Partie : 4m1, Nombre de contextes : 54

1 - eh bien retrouvons cet élan ! - mais **je** considère , quant à moi , que l ' aide nos cadres , de nos agriculteurs , quand **je** constate le rôle de la France sur la s e qu ' on gagne les victoires et , moi , **je** crois de toutes mes forces à la France des années , ne se sont pas produites ; **je** crois pouvoir dire : au contraire . l nature ont accablés . c ' est mon rôle , **je** crois , que d ' exprimer au nom de tou tre unis quand notre avenir est en jeu . **je** désire , précisément , réfléchir avec urité et notre indépendance . aussi ai - **je** donné l ' ordre - toute politique est ' opposer , parfois même à se déchirer . **je** forme des vœux , simplement , pour qu l an , je vous présente ce soir ceux que **je** forme pour la France . comme la plupar retour d ' un être cher . les vœux que **je** forme pour vous ne varient pas avec le es français aux français . cette mission **je** la remplirai . nous avons traversé bie e le souhaitable et le possible . mais , **je** le répète ici , il n ' y aura pas de r perpuissances s ' arranger ou surarmer . **je** m ' en tiens à cette règle d ' or : l possible , dans un esprit de tolérance . **je** n ' établirai pas ici la liste de ces fois désespérés attendent un emploi et , **je** n ' hésite pas à le dire , trop d ' in u des conditions de notre indépendance . **je** n ' insisterai ce soir que sur un poin milles . en dépit de leurs divergences , **je** ne me laisserai jamais d ' espérer ni d otre pays sortir peu à peu de la crise , **je** ne promets rien d ' autre à personne q t dangereux que les deux puissances dont **je** parle puissent coexister sur la base d ses , français , avant de vous quitter , **je** pense à celles et à ceux d ' entre vou se substituent aux droits de l ' homme : **je** pense à l ' Afghanistan , au Liban , à la solitude , qui souffrent du chômage . **je** pense , bien entendu , aussi à celles uisque j ' ai prononcé le mot sécurité , **je** précise qu ' il s ' applique aussi bie ive la France ! mes chers compatriotes , **je** remercie la tradition qui me vaut , po à moi , garant de l ' unité nationale , **je** serai là pour assurer la continuité de n montrent pas beaucoup ces temps - ci : **ie** souhaite ardemment qu ' ils se reprenn

2.2.2. Concordances du lemme croire

Retour										CONCORDANCE										Irier Notes Imprimer Sommaire									
Forme	Lemme	Expr.	Initial	Final	Chain	Liste	Tout	Nb	29																				
M1 114e		oin qu' on l' aide et que l' on y croie , elle a besoin que les peuples																											
DG 41c		avec sérénité , parce qu' on peut croire que la suppression prochaine de																											
DG 43c		p saisi de vertige . On a même pu croire , un moment , qu' il s' abandon																											
GP 50d		prochain printemps . Ils veulent croire que vieillir est un moyen de ma																											
M2 128c		pas l' aumône mais des raisons de croire dans un régime de liberté et de																											
JC 155b		comme certains ont voulu le faire croire , mais pour garder et léguer à																											
JC 170a		être fiers de notre patrie et de croire en notre avenir . Cet avenir ,																											
DG1b		ide , laborieuse , cohérente , je crois que jamais , en dépit des diffic																											
DG 15f		bres au tableau . Tout de même je crois qu' en 1962 nous n' avons pas pe																											
DG 40c		as un grand peuple en marche . Je crois , cependant , qu' au total , et																											
VG 70a		ne lui rend pas service . Je ne crois pas que cela le rende plus heure																											
VG 70e		ne et heureuse année 1976 , et je crois qu' Anne Aymone veut aussi vous																											
M1 110b		gagne les victoires et , moi , je crois de toutes mes forces à la France																											
M1 111b		s , ne se sont pas produites ; je crois pouvoir dire : au contraire .																											
M1 117a		C' est mon rôle , je crois , que d' exprimer au nom de tous																											
M2 131a		Mais je crois encore aux chances de la paix et																											
M2 152c		n et plein d' espoir en vous . Je crois aux forces de l' esprit et je ne																											
JC 167e		t en jeu pour vous dire ce que je crois être bon pour les Français ou ,																											
GP 54b		tout va mal . En vérité , qui le croit réellement ? Personne ne peut ni																											
M2 120e		laissés pour compte qu' on ne le croit généralement , que ce soit pour																											
M2 139c		parfois de vieilles haines qu' on croyait oubliées . Tel est le cas de l																											
GP 53a		problèmes , mon Gouvernement a , croyez - le , une claire conscience .																											
VG 71a		toire a rendu sceptique , vous ne croyez pas beaucoup à la sincérité des																											
M1 93c		distribution . Ce sera aussi , croyez - moi , la meilleure façon de v																											
M1 102c		ce sens l' a été grâce à vous qui croyez en la France . C' est encore																											
M1 112a		Croyez - moi , un peuple , pour être g																											
M2 123c		es fruits de l' effort commun . Croyez - le , mes chers Compatriotes ,																											
M2 145b		r la lutte contre le sida . Et ne croyez - vous pas insupportable d' ass																											
DG 14d		, celui des crises politiques , a cru trouver l' occasion de revenir nou																											

2.2.3. Concordances d'une liste de quelques formes avec Lexico 3

Groupe de formes

Nom du groupe : ➔

Le motif :

☒ Ajouter

Rechercher Enregistrer ☐

Supprimer Charger

Forme	Fréquence
je	337
i	85
me	22
mon	28
ma	19
mes	100

Concordances

Forme : Tri : Regroupement : Largeur :

☐ Expression rationnelle Type de documentation : Délimiteurs de séquence :

Partie : 3vg, Nombre de contextes : 131

reviens d ' Afrique . dans trois jours , j ' accueillerai en terre française ' Guadeloupe
et ceux que nous rencontrerons en 1977 . j ' adresse aussi mes vœux à ceux qui sont
ier vœu que j ' exprime ce soir , et que j ' adresse en votre nom à la France . et
vivent à la réflexion . c ' est pourquoi j ' ai choisi de vous dire 2 choses toutes
n haut que ce sont les plus françaises . j ' ai confiance dans le progrès de la France
ogres de la France . je vous le répète , j ' ai confiance dans le progrès de la France
ent et affectueusement à la France , que j ' ai pour mission de conduire et de servir
me mais de la confiance dans un pays que j ' aime et que je respecte , une confiance
es et les français , mes compatriotes et mes amis . a la France . je souhaite que l '
aises et français , mes contemporains et mes amis , je vous demande de prendre en charge
ons que je vous ai déjà décrites et dont j ' aurai l ' occasion de vous parler à nouveau
- mêmes , si désirée par les français . mes chers amis , je reviens à mes mots du début
dans le monde . françaises et français , mes chers amis , je vous exprime les vœux très
e la république . mes chers françaises , mes chers français , ce soir est une fête pour
les vœux du président de la république . mes chers françaises , mes chers français ,
vous , les françaises et les français , mes compatriotes et mes amis . a la France .
te pour elle et pour le monde . a vous , mes compatriotes , mes vœux s ' adressent à
' unité des français . quant à moi , je me consacrerai à cette recherche de l ' unité
t courageuses . françaises et français , mes contemporains et mes amis , je vous demande
ce une bonne et heureuse année 1976 , et je crois qu ' Anémone veut aussi vous adresser
déclaration , d ' une interview , et on me demande : " pour vous , la France , qu '
' affaire des éducateurs et des jeunes - mon dernier souhait sera pour l ' unité de la
iens et de services produits en France , je dis bien la quantité , a augmenté de 15
tes simples , mais qu ' il me semble que je dois rappeler dans un moment difficile pour
ité de sang - froid de quelques hommes . j ' espère que la sagesse des dirigeants l '
grand pays . c ' est le premier vœu que j ' exprime ce soir , et que j ' adresse en
urprimer a vous toutes et à vous tous . j ' exprime les vœux du président de la république

2.2.4. Les concordances de deux lemmes éloignés avec Weblex. Utilisation de CQP = Corpus Query Processor

Chercher dans la page :

Il y a 7 occurrences de "France"[*]"enfant" dans le corpus voeux2-author-Degaulle (propr

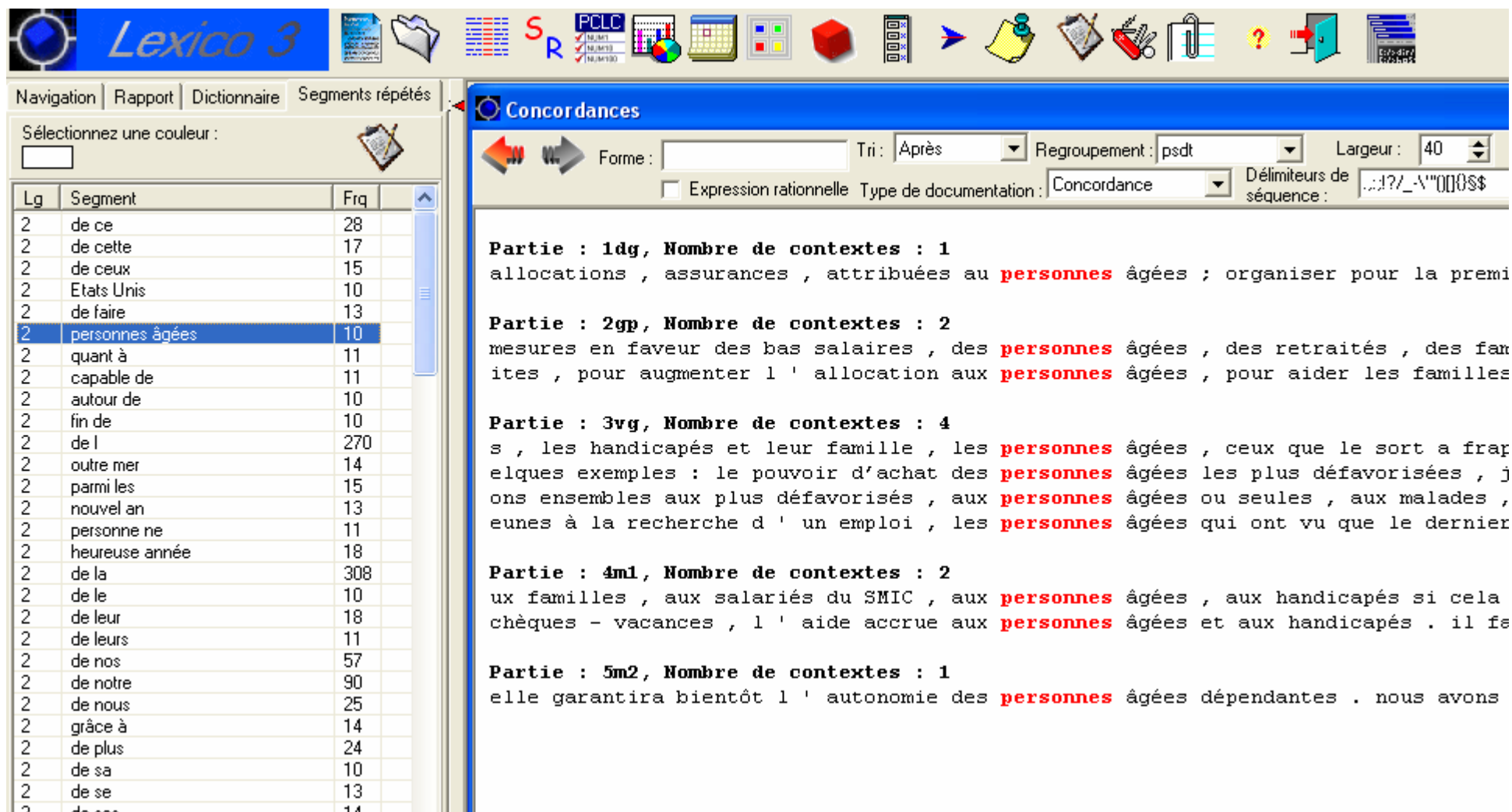
1	Degaulle, 1960	ur et à son énergie . Car l'Algérie a besoin de la communauté française , et la	France , pour son oeuvre , a besoin d' elle en Algérie . Bien entendu , et , quoi qu' il arrive , la : enfants
2	Degaulle, 1963	ustes envers nous-mêmes . Car le bilan est positif . Pendant ces douze mois , la	France a continué de monter . En 1963 , notre population s' est augmentée de presque six cen est né environ neuf cent mille bébés . La proportion croissante des jeunes devant progressiv processus , on peut penser que , parmi les enfants
3	Degaulle, 1963	ns , j' ai l' honneur et la charge de parler en notre nom à tous , j' offre à la	France , cette fois encore , les voeux très ardents et très confiants de ses enf
4	Degaulle, 1964	mmes sur la terre n' ont tant attendu de nous ni éprouvé tant d' attrait pour la	France . Certes , la vie est la vie , autrement dit un combat , pour une nation comme pour un ho , toujours et partout , des difficultés à vaincre , des efforts à déployer , des peines à supporter , de dignité , de justice , de fraternité . Mais , tous ensemble , nous sommes un peuple , dont l' év les enfants
5	Degaulle, 1965	ns nos têtes , dans nos coeurs et dans nos mains tout ce qu' il faut pour que la	France parcourt une étape décisive de son progrès , pour qu' elle apporte plus de justice enco entre tous ses enfants
6	Degaulle, 1965	eurs voeux pour 1966 et que , tous ensemble nous souhaitons une bonne année à la	France . Vive la République ! Vive la France . Françaises , Français , Je vous souhaite une bon nom de nous tous , mes voeux sont l' expression de ceux que la France forme pour chacu

2.3. Les segments répétés

Les successions de formes contiguës du texte

- Avec Lexico 3
- Avec Weblex

2.3.1. Les segments répétés avec Lexico 3



The screenshot shows the Lexico 3 software interface. The top menu bar includes 'Navigation', 'Rapport', 'Dictionnaire', and 'Segments répétés'. The 'Segments répétés' window on the left displays a table of segments and their frequencies. The 'Concordances' window on the right shows search results for the segment 'personnes âgées'.

Segments répétés Table:

Lg	Segment	Frq
2	de ce	28
2	de cette	17
2	de ceux	15
2	Etats Unis	10
2	de faire	13
2	personnes âgées	10
2	quant à	11
2	capable de	11
2	autour de	10
2	fin de	10
2	de l	270
2	outre mer	14
2	parmi les	15
2	nouvel an	13
2	personne ne	11
2	heureuse année	18
2	de la	308
2	de le	10
2	de leur	18
2	de leurs	11
2	de nos	57
2	de notre	90
2	de nous	25
2	grâce à	14
2	de plus	24
2	de sa	10
2	de se	13
2	de ses	14

Concordances Window:

Forme : [] Tri : Après Regroupement : psdt Largeur : 40

☐ Expression rationnelle Type de documentation : Concordance Délimiteurs de séquence : []

Partie : 1dg, Nombre de contextes : 1
allocations , assurances , attribuées au **personnes** âgées ; organiser pour la premi

Partie : 2gp, Nombre de contextes : 2
mesures en faveur des bas salaires , des **personnes** âgées , des retraités , des fam
ites , pour augmenter l ' allocation aux **personnes** âgées , pour aider les familles

Partie : 3vg, Nombre de contextes : 4
s , les handicapés et leur famille , les **personnes** âgées , ceux que le sort a frag
elques exemples : le pouvoir d'achat des **personnes** âgées les plus défavorisées , j
ons ensemble aux plus défavorisés , aux **personnes** âgées ou seules , aux malades ,
eunes à la recherche d ' un emploi , les **personnes** âgées qui ont vu que le dernier

Partie : 4ml, Nombre de contextes : 2
ux familles , aux salariés du SMIC , aux **personnes** âgées , aux handicapés si cela
chèques - vacances , l ' aide accrue aux **personnes** âgées et aux handicapés . il fe

Partie : 5m2, Nombre de contextes : 1
elle garantira bientôt l ' autonomie des **personnes** âgées dépendantes . nous avons

2.3.2. Les segments répétés du corpus vœux. Weblex

f	segment	1
8	<u>chacune et à chacun d' entre vous</u>	7
5	<u>à chacune et à chacun d' entre</u>	7
5	<u>à chacune et à chacun de vous</u>	7
3	<u>C' est au nom de l' emploi</u>	7
3	<u>au nom de l' emploi que nous</u>	7
3	<u>est au nom de l' emploi que</u>	7
3	<u>À chacune et à chacun d' entre</u>	7
8	<u>chacune et à chacun d' entre</u>	6
8	<u>et à chacun d' entre vous</u>	6
6	<u>à chacune et à chacun de</u>	6
5	<u>chacune et à chacun de vous</u>	6
5	<u>à chacune et à chacun d'</u>	6
4	<u>Nous ne sommes pas les plus</u>	6
4	<u>À chacune et à chacun d'</u>	6
3	<u>C' est au nom de l'</u>	6
3	<u>Il n' y a pas d'</u>	6
3	<u>Personne ne peut nier que la</u>	6
3	<u>au nom de l' emploi que</u>	6
3	<u>ceux qui ont la chance d'</u>	6
2	<u>à</u>	2

2.4. Les cooccurrences

Présence de deux formes non contiguës

- **Les cooccurrences de deux formes dans hyperbase**
- **Utilisation de la fonction thème d'hyperbase pour les découvrir. Images : 1 et 2.**
- **Le calcul des cooccurrences dans weblex**
- **Cooccurrences gauches et droites**
- **Graphique**
- **Résumé du graphique**

2.4.1. Les cooccurrences : Unité France

Retour	Notes	Cherche	Forme	Lemme	initial	final	LONG	parag	CONTEXTES page 1	Cliquez un extrait pour voir la page		
			expression	cooccurr.	liste	14			autres résultats			

Le bonheur pour la FRANCE , c' est la confiance en elle - même et c' est l' UNITÉ . La confiance en elle - même , parce que nous traversons une époque difficile , qui est une époque d' évolution et d' adaptation dans le monde , à la recherche d' un nouvel équilibre . La FRANCE doit avoir confiance en elle - même , parce qu' elle est capable de surmonter ces difficultés . Quand nous pensons à ce qui s' est passé en FRANCE depuis le début de ce siècle , et dont beaucoup d' entre vous se souviennent , quand nous pensons à la première guerre qui a décimé la plupart des familles françaises , quand nous pensons au deuxième conflit dans lequel notre pays a été totalement occupé et son économie désorganisée et lorsque nous constatons que nous avons été capables de surmonter ces difficultés , la FRANCE peut avoir confiance en elle - même pour surmonter les difficultés du présent .

VG . Page: 75 a (occ. 2)

mon dernier souhait sera pour l' UNITÉ de la FRANCE . La FRANCE est au total , malgré d' inutiles querelles , plus unie qu' il y a un an . On pouvait craindre qu' elle ne sorte déchirée du grave débat politique de mars dernier ` élections législatives ` . La haine et la rancune pouvaient nouveau diviser les français . Mais le sens de l' UNITÉ ce sens venu du fond des âges et qui est le certificat de naissance des nations le sens de l' UNITÉ l' a emporté grâce au _concours , je tiens à le dire , des uns et des autres .

VG . Page: 78 a (occ. 3)

En 1980 , la FRANCE a besoin du sérieux , du courage , des facultés d' adaptation , de la générosité et de l' UNITÉ des Français .

VG . Page: 79 c (occ. 4)

2.4.2. Les contextes courts de je

 C:\HYPERBASWOEUX.EXE

Retour  Notes  Cherche

Forme	Lemme	initial	final	LONG	50
expression	cooccure.	liste	346		

CONTEXTES page 1

autres résultats   Sommaire

Cliquer un extrait pour voir la page

, c' est à vous que JE les adresse et , en
DG Page: 30 c (15ème occ.)

, c' est à vous que JE les adresse et , en
DG Page: 30 d (16ème occ.)

JE vous souhaite une bonne année ! Parlant au nom de
DG Page: 35 b (17ème occ.)

JE vous souhaite une bonne année ! Parlant au nom de
DG Page: 35 b (18ème occ.)

De tout mon coeur , JE souhaite une bonne année à la
DG Page: 39 c (19ème occ.)

pas aux hommes et JE ne le prédis pas .
DG Page: 40 a (20ème occ.)

pas aux hommes et JE ne le prédis pas .
DG Page: 40 b (21ème occ.)

peuple en marche . JE crois , cependant , qu'
DG Page: 40 b (22ème occ.)

peuple en marche . JE crois , cependant , qu'
DG Page: 40 c (23ème occ.)

2.4.3. L'environnement de la forme je

Retour

Environnement d'un mot (ou groupe de mots)

Cliquer sur un mot pour voir les contextes

seuil

Sommaire

écart	corpus	texte	mot	HIERARCHIQUE
12.68	66	49	SOUHAITE	
9.66	326	98	VOUS	
7.70	12	13	DIS	
7.16	25	18	SUIS	
6.64	20	15	SAIS	
5.34	19	12	VOUDRAIS	
5.28	77	26	VOEUX	
5.28	11	9	CROIS	
5.16	27	14	PENSE	
4.82	16	10	FORME	
4.35	7	6	DEMANDE	
4.04	11	7	VEUX	
3.98	665	107	QUE	
3.88	6	5	AVAIS	
3.75	4	4	VAIS	
3.75	4	4	TIENS	
3.75	4	4	DISAIS	
3.66	17	8	ADRESSE	
3.65	30	11	NOM	
3.45	2	3	GARDIEN	
3.38	5	4	FIER	
3.18	3	3	SALUE	
3.18	3	3	ASSURE	
3.11	48	13	DIRE	
3.11	22	8	ME	
3.10	27	9	COEUR	
3.10	6	4	VOIS	
3.10	6	4	PROMIS	
2.88	7	4	BONSOIR	
2.85	11	5	SÛR	
2.82	4	3	RÉPÈTE	

écart	corpus	texte	mot	ALPHABETIQUE
2.07	42	9	AI	
2.11	12	4	AMIS	
2.51	2	2	ARDEMENT	
3.18	3	3	ASSURE	
2.51	2	2	ATOMIQUE	
3.88	6	5	AVAIS	
2.51	2	2	AVIEZ	
2.71	74	16	BONNE	
2.88	7	4	BONSOIR	
2.51	2	2	CHANTIERS	
3.10	27	9	COEUR	
2.15	3	2	CONTRIBUE	
5.28	11	9	CROIS	
4.35	7	6	DEMANDE	
2.15	3	2	DIRAI	
3.11	48	13	DIRE	
7.70	12	13	DIS	
3.75	4	4	DISAIS	
2.35	15	5	DIT	
2.51	2	2	DOIS	
2.15	3	2	ÈRE	
2.33	6	3	ESPÈRE	
2.11	12	4	EURO	
3.38	5	4	FIER	
4.82	16	10	FORME	
3.45	2	3	GARDIEN	
2.51	2	2	HOMMAGE	
33.69	346	265	JE	
2.50	41	10	LUI	
3.11	22	8	ME	
2.33	6	3	MOT	

2.4.4. Extrait de la liste de toutes les cooccurrences significatives

Cooccurrences du corpus "voeux2" (forme)

Seuils : f 3, cf 3, p 5.0E-2, d_m 1000.0

A B	f _A	f _B	cf	p	d _m
Chers Compatriotes	34	34	34	2e-75	0.0
Françaises Français	38	124	34	8e-41	0.7
chers compatriotes	21	28	18	1e-34	0.0
Vive République	55	60	29	3e-34	1.0
bonne année	51	199	37	4e-30	0.7
Nouvel An	12	16	12	1e-28	0.0
Je souhaite	136	65	34	2e-26	0.2
personnes âgées	12	10	10	1e-25	0.0
pouvoir achat	18	9	9	2e-20	1.0
intérêts particuliers	15	10	9	2e-20	0.0
Bonne année	24	199	21	2e-20	0.4
bonne heureuse	51	22	14	6e-19	1.6
heureuse année	22	199	19	3e-18	1.6
souhaite bonne	65	51	18	3e-16	7.0
je souhaite	208	65	30	5e-16	0.2
peut nier	41	10	9	1e-15	0.0
cent mille	12	10	7	3e-15	0.4
départements territoires	6	6	5	1e-13	1.4

2.4.5. Cooccurrences droites et gauches de France

Lexicogramme du pôle "France" dans le corpus voeux2-author-Degaulle (p3)

Seuils : f 3, cf 3, p 5.0E-2, d_m 1000.0

France

(70)

cooccurents gauches					cooccurents droits				
	f	cf	p	d _m		f	cf	p	d _m
vive	18	9	1e-04	1.0	enfant	8	5	1e-03	8.6
souhaiter	12	6	2e-03	4.5					
nom	10	5	5e-03	13.2					
adresser	4	3	8e-03	14.7					
année	47	11	2e-02	6.5					
voeu	10	4	3e-02	13.0					
maintenant	6	3	3e-02	16.3					
terre	6	3	3e-02	16.7					

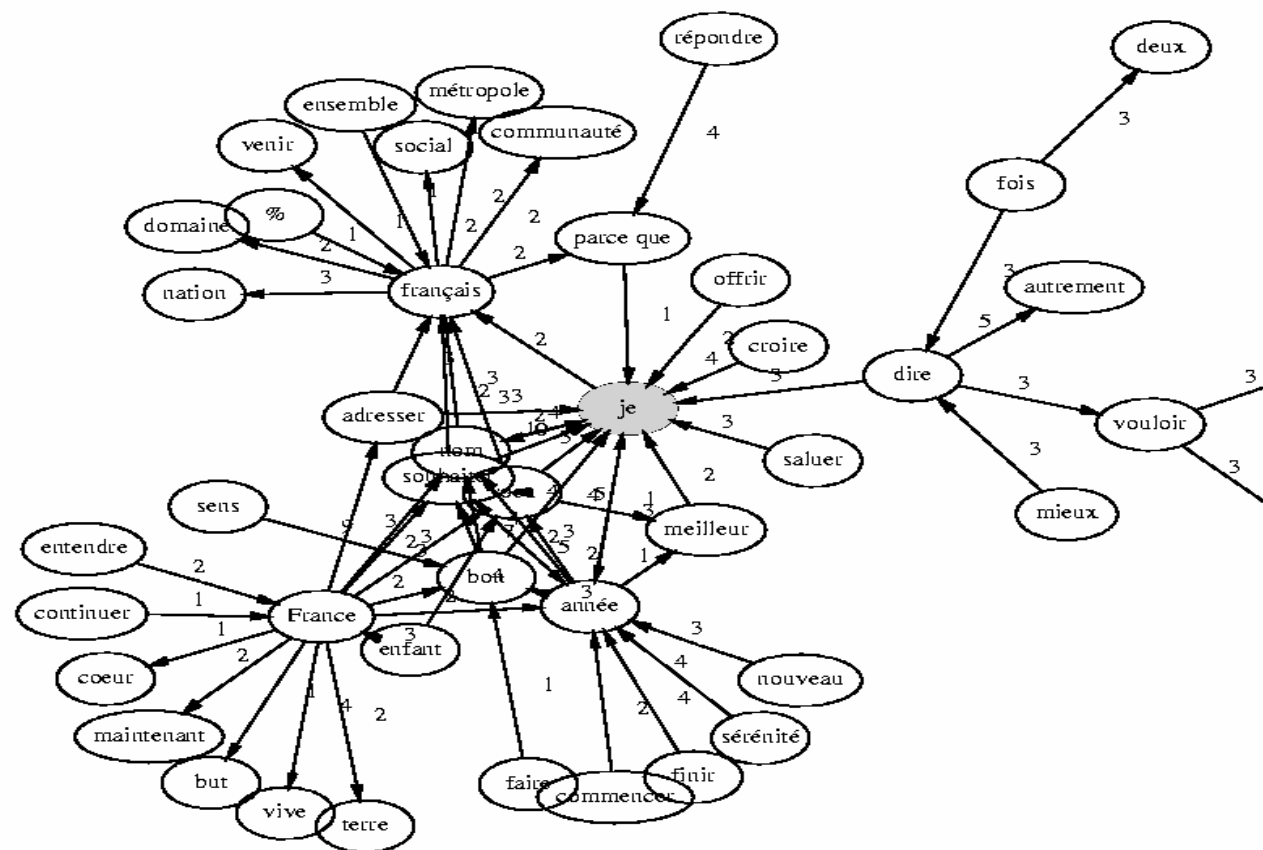
2.4.6. Graphique des cooccurrences de je

Lexicogramme récursif autour du pôle je dans le corpus vœux2-author-Degaulle

Seuils : **p** 2e-01, **r** 3, **f** 3, **d_m** 1000.0, **pl** 3

Synthèse

0 noeuds (0 arcs)



2.4.7. Résumé du graphique des cooccurrences de je

Calcul en cours ([abandon](#))...

Synthèse des lexicogrammes récursifs du corpus `voeux2-author-Degaulle`

Seuils : **p** 5e-02, **r** 3, **f** 3, **d_n** 1000.0, **pl** 3

Tri décroissant par le nombre de noeuds par lexicogramme.

- ♦ 28 (nous avancer prospérité voie développement Amérique latin disposer lui-même part prendre propre évolution aide état Europe détente union coopération puis grand nombre peuple Afrique)
 - ♦ 12 (communauté France finir nouveau sérénité bon je nom souhaiter année commencer Algérie)
 - ♦ 11 (enfant meilleur voeu métropole nation catégories domaine économique social français adresser)
 - ♦ 5 (d abord expansion niveau vie agir)
 - ♦ 3 (commerce industrie agriculture)
-

3. L'approche statistique ou lexicométrique

De l'étude des fréquences des formes au positionnement des textes

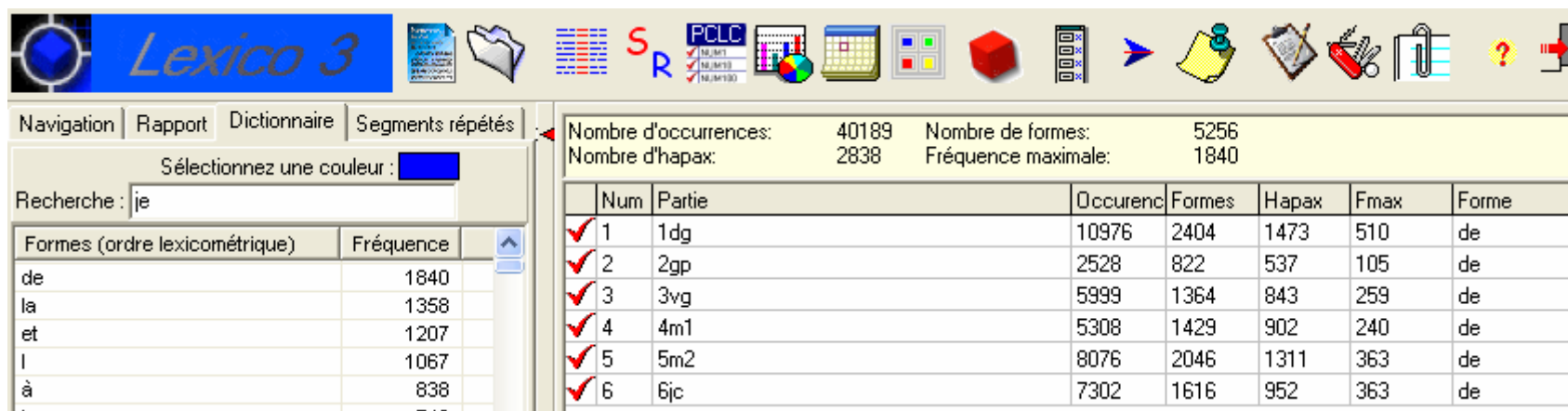
- Les caractéristiques d'un corpus
- Les spécificités d'une partie de celui-ci
- Rapprochement et éloignement des textes le composant

3.1. Les caractéristiques d'un corpus

3.1.1. L'étude des formes en fonction de leur fréquence

- Le nombre d'occurrences
- Les fréquences absolues et relatives
- Les mots outils
- Les hapax

3.1.2. Le nombre d'occurrences du corpus et de ses composantes



The screenshot shows the Lexico 3 software interface. The top menu bar includes Navigation, Rapport, Dictionnaire, and Segments répétés. Below the menu, there's a search bar with the text "je" and a color selection button. The main display area is divided into two sections. The top section shows summary statistics: "Nombre d'occurrences: 40189", "Nombre de formes: 5256", "Nombre d'hapax: 2838", and "Fréquence maximale: 1840". The bottom section is a table with columns: Num, Partie, Occurenc, Formes, Hapax, Fmax, and Forme. The table lists six entries, each with a checkmark in the first column.

Num	Partie	Occurenc	Formes	Hapax	Fmax	Forme
✓ 1	1dg	10976	2404	1473	510	de
✓ 2	2gp	2528	822	537	105	de
✓ 3	3vg	5999	1364	843	259	de
✓ 4	4m1	5308	1429	902	240	de
✓ 5	5m2	8076	2046	1311	363	de
✓ 6	6jc	7302	1616	952	363	de

3.1.3. Fréquence de nous

Nb. d'occurrences du corpus : 40189
 Fréquence de nous : 645
 Fréquence relative de nous : $\frac{645}{40189} = 0.016$
 Cadence : $\frac{40189}{645} = 62$

Lexico 3

Navigation | Rapport | Dictionnaire

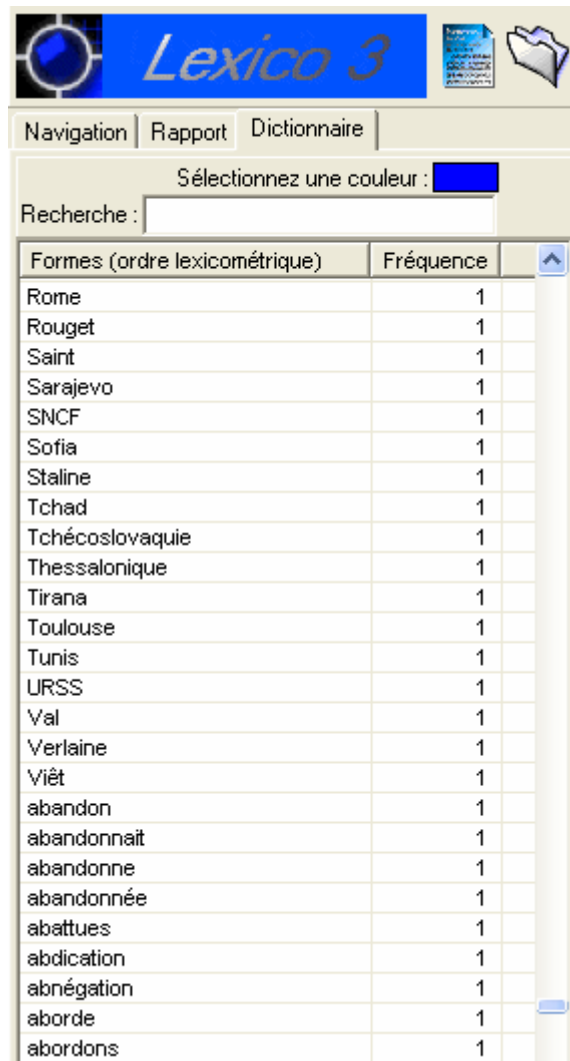
Sélectionnez une couleur :

Recherche :

Formes (ordre lexicométrique)	Fréquence	Formes/SR	1dg	2gp	3vg	4m1	5m2	6jc
nous	645	nous	199	+3 52	+2 77	-3 75	111	-2 131
d	597	nous	0	0	0	0	1	0
en	582	nouveau	2	1	6	+2 2	3	4
		nouveaux	3	1	0	2	0	4
		nouvel	1	-3 1	1	6	+3 5	4
		nouvelle	11	1	3	1	-2 10	10

3.1.4. Les mots outils dans le tableau lexical entier

3.1.5. Les vrais et les faux hapax



Formes (ordre lexicométrique)	Fréquence
Rome	1
Rouget	1
Saint	1
Sarajevo	1
SNCF	1
Sofia	1
Staline	1
Tchad	1
Tchécoslovaquie	1
Thessalonique	1
Tirana	1
Toulouse	1
Tunis	1
URSS	1
Val	1
Verlaine	1
Viêt	1
abandon	1
abandonnait	1
abandonne	1
abandonnée	1
abattues	1
abdication	1
abnégation	1
aborde	1
abordons	1

3.2. Les spécificités d'une partie

Les formes caractérisant une partie d'un corpus

- La loi hypergéométrique
- J.C. Spécificités positives et négatives
- J.C. Spécificités phrastiques

3.2.1. J.C. Spécificités positives



Navigation | Rapport
Dictionnaire | Spécifs - Part : psdt

Corpus de référence : 2gp, 3vg, 4m1, 5m2, 6jc,
Parties sélectionnées : 6jc,
Spécificités ☒ positives ☐ négatives

Terme	Frq Tot.	Frq P...	Spécif
solidarité	33	22	10
euro	12	12	10
doit	61	28	7
une	354	102	7
avenir	46	23	7
valeurs	12	10	7
chaque	27	15	6
être	109	37	5
créer	12	8	5
ambition	16	10	5
état	46	20	5
cohésion	13	9	5
faire	68	24	4
choix	15	9	4
aussi	91	29	4
publics	13	8	4
construire	13	8	4
nation	36	14	4
compatriotes	62	21	4
nouvelles	18	10	4
conscience	15	9	4
siècle	21	10	4
libertés	14	7	3
outré	17	7	3
trop	24	9	3
surtout	14	7	3

3.2.2. J.C. Spécificités négatives



Lexico 3

Navigation | Rapport

Dictionnaire | Spécifs - Part : psdt

Corpus de référence : 2gp, 3vg, 4m1, 5m2, 6jc
 Parties sélectionnées : 6jc
 Spécificités ☒ positives ☐ négatives

Terme	Frq Tot.	Frq P...	Spécif
lui	41	3	-2
mois	34	2	-2
ans	42	3	-2
leurs	42	3	-2
on	112	10	-3
pays	118	11	-3
ils	66	5	-3
sans	85	8	-3
par	122	13	-3
que	656	94	-3
à	838	125	-3
a	261	34	-3
du	281	34	-3
depuis	38	2	-3
peuples	39	2	-3
celle	32	1	-3
ni	34	1	-3
françaises	39	2	-3
paix	62	5	-3
votre	58	4	-3
économique	45	3	-3
politique	52	3	-3
en	582	80	-4
qu	306	25	-7

3.2.3. J.C. Spécificités phrastiques

C:HYPERBASWOEUX.EXE

Refaire résumé 5M2 Voir les mots spécifiques et cacher les passages spécifiques Trier Cherche CLIC+MAJ: Recherche du mot dans les textes Retour Sommaire 6JC

Cliquer sur un extrait pour voir le contexte 5M2 6JC

2.634 Oui , notre pays est porteur d'un message . Un message de fraternité , de tolérance et de justice . C'est ainsi que je vois la France . Une France qui regarde en FACE son passé , les HEURES sombres comme les HEURES glorieuses . Une France FIDÈLE à son histoire , à ses VALEURS , et décidée à les défendre . Une NATION qui entre dans l'AVENIR avec CONFIANCE , parce qu'elle a choisi le progrès et la SOLIDARITÉ .

2.398 Cette France VIVANTE et FORTE , cette France ACCUEILLANTE , apaisée , elle sera la nôtre grâce aux CHOIX que nous avons faits : CHOIX de l'Europe , CHOIX de la MODERNITÉ , RESPECT de l'identité française .

2.356 Choisissons résolument d'AVANCER . C'est ainsi que nous ouvrirons notre vie politique , que nous FERONS mieux vivre notre DÉMOCRATIE . C'est ainsi que nous inventerons une nouvelle SOLIDARITÉ , une SOLIDARITÉ RESPONSABLE pour ramener vers l'EMPLOI ceux qui en sont exclus , PARFOIS depuis longtemps . C'est ainsi que nous donnerons à nos enfants une bonne formation pour l'EMPLOI . C'est ainsi que nous garantirons l'AVENIR de nos RETRAITES .

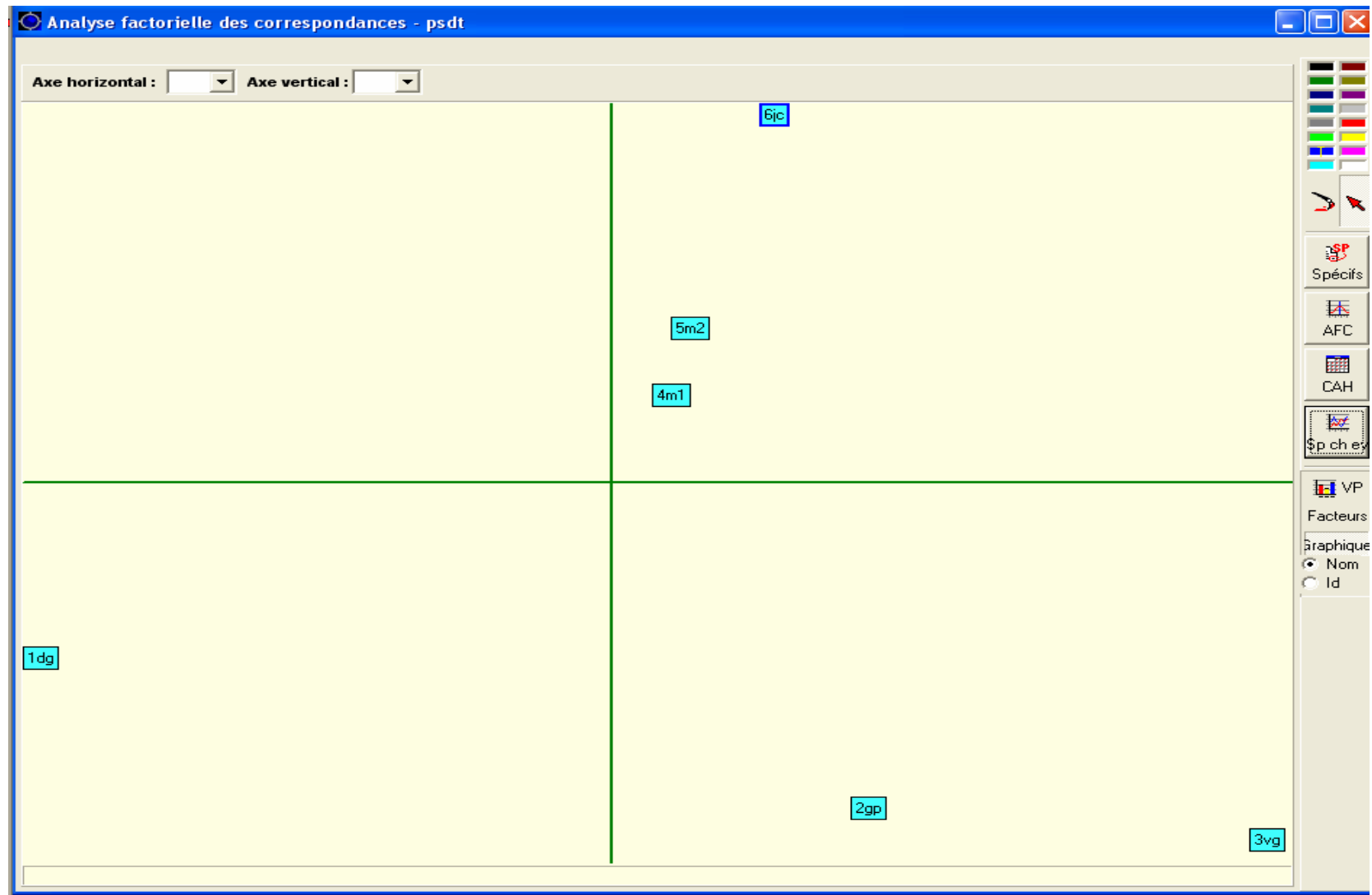
écart	corpus	texte	mot
6.0	34	24	solidarité
5.7	358	124	une
5.7	12	12	Euro
5.1	46	26	avenir
5.0	12	11	valeurs
4.4	62	29	doit
4.4	16	12	ambition
4.0	42	21	Etat
4.0	18	12	nouvelles
3.8	27	15	chaque
3.8	12	9	créer
3.8	6	6	patrimoine
3.6	109	40	être
3.6	69	28	faire
3.6	13	9	construire
3.5	16	10	face
3.4	14	9	publics
3.4	14	9	cohésion
3.4	7	6	énergies
3.4	5	5	mondialisation
3.4	5	5	modernité
3.4	5	5	2001
3.2	163	52	nos
3.2	15	9	devons
3.2	15	9	conscience
3.2	15	9	choix

3.3. Rapprochement et éloignement des textes d'un corpus

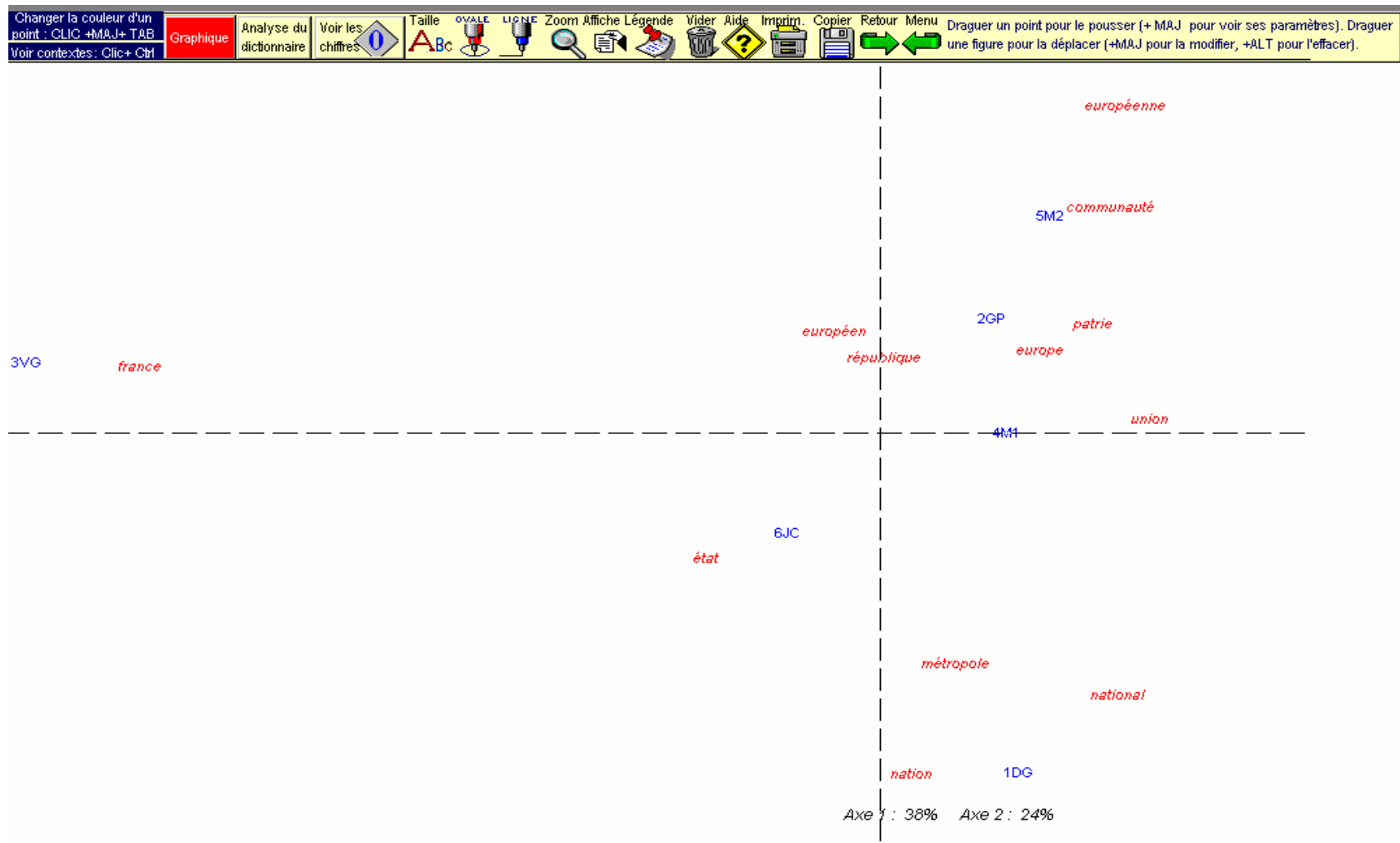
Oppositions, rapprochements et distances entre les textes d'un corpus

- **L'analyse factorielle du dictionnaire**
- **AFC sur une liste de mots**
- **Le classement automatique hiérarchisé**

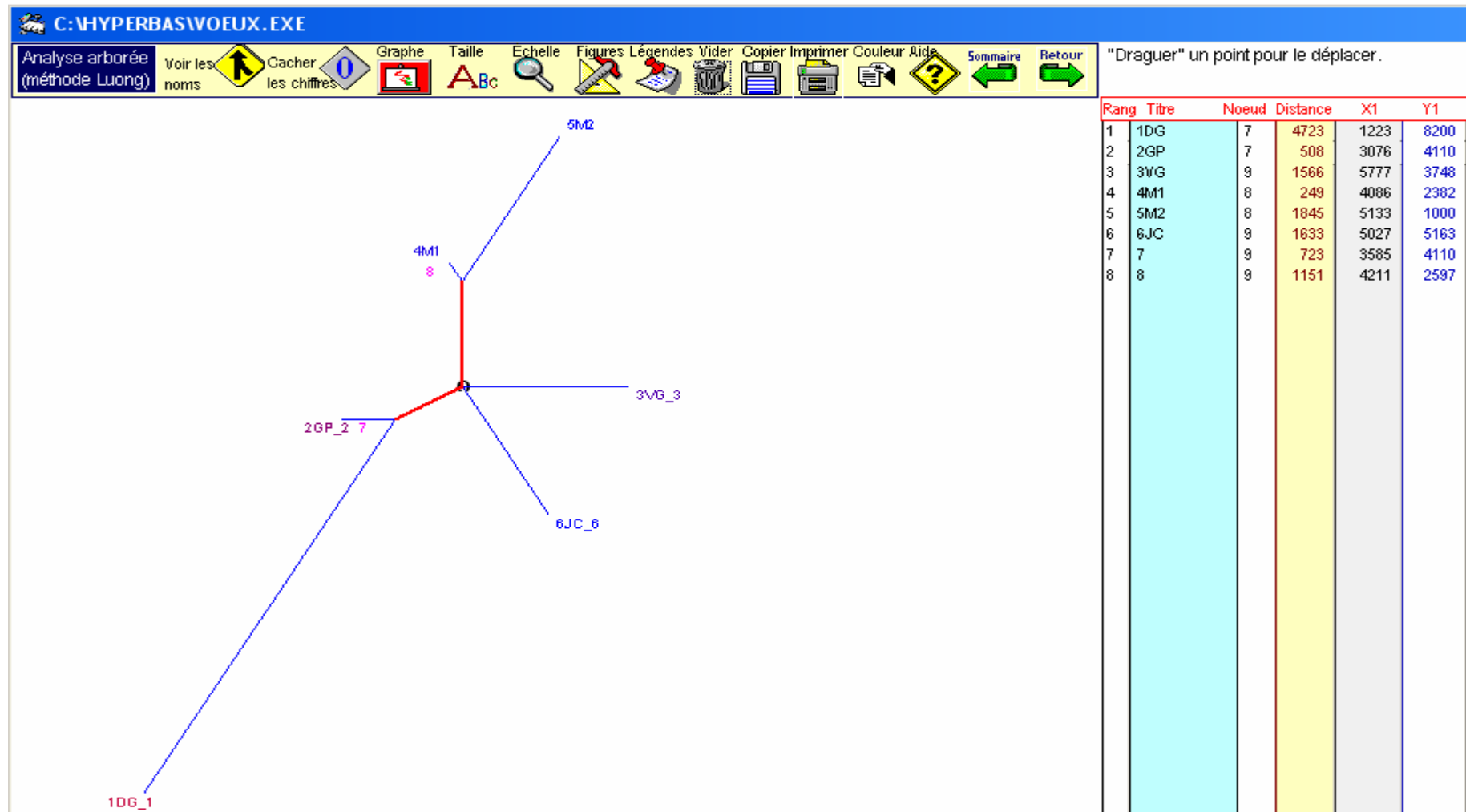
3.3.1. L'analyse factorielle du dictionnaire



3.3.2. AFC sur une liste de mots



3.3.3. Le classement automatique hiérarchisé



4. Une autre approche en gestation

- L'intégration de corpus et de textes formatés différemment dont ceux en xml (TEI)
- Les traitements préalables des textes : treetagger, cordial
- Un travail plus tourné sur des formes lemmatisées
- Des analyses plus fines des formes et des segments
- Un ou des logiciels exploitant un fond commun de petits outils

4.1. Lemmes et catégories grammaticales

- Hy : Formes non lemmatisées et lemmes
- Hy : Formes et catégories grammaticales
- Hy : concordance D N A V (det. nom. adj. ver.) nom = France
- Nooj : concordance (N, V, A) nom = France, verbe = être,

4.1.1. Formes et lemmes

C:\HYPERBASVELYSEE.EXE

Sommaire Retour N° Mots 173 Lettres 948 Page 1 DG 3

Ecarts Textes Recherche Notes Code/Lemme page CLIC sur un mot : autres contextes

L' unité française se brisait .
 La guerre civile allait commencer .
 Aux yeux du monde , la France paraissait sur le point de se dissoudre .
 C' est alors que j' ai assumé la charge de gouverner notre pays .
 Le drame de l' Algérie , bouleversant les populations , mettant l' armée à dure épreuve , soulevant sur place une vague d' indignation pour ce qui était du présent , de changement et de fraternité pour ce qui était de l' avenir , a déclenché cette crise nationale .
 Mais , de toutes les manières , celle - ci devait éclater .
 Car , depuis 12 ans , le régime des partis , flottant sur un peuple profondément divisé au milieu d' un univers terriblement dangereux , se montrait hors d' état d' assurer la conduite des affaires .
 Non point par incapacité ni par indignité des hommes .
 Ceux qui ont participé au pouvoir sous la quatrième République étaient des gens de valeur , d' honnêteté , de patriotisme .

le 7 unité 2 français 3 se 5 briser 1 .
 le 7 guerre 2 civil 3 aller 1 commencer 1 .
 au 7 oeil 2 du 7 monde 2 , le 7 France 2 paraître 1
 sur 9 le 7 point 2 de 9 se 5 dissoudre 1 .
 ce 5 être 1 alors 6 que 8 je 5 avoir 1 assumer 1 le 7 charge 2
 de 9 gouverner 1 notre 7 pays 2 .
 le 7 drame 2 de 9 le 7 Algérie 2 , bouleverser 1
 le 7 population 2 , mettre 1 le 7 armée 2 à 9
 dur 3 épreuve 2 , soulever 1 sur 9 place 2 un 7
 vague 2 de 9 indignation 2 pour 9 ce 5 qui 5 être 1
 du 7 présent 2 , de 9 changement 2 et 8 de 9 fraternité 2
 pour 9 ce 5 qui 5 être 1 de 9 le 7 avenir 2 , avoir 1 déclencher 1
 ce 7 crise 2 national 3 .
 mais 8 , de 9 tout 3 le 7 manière 2 , celui 5 2
 ci 6 devoir 1 éclater 1 .
 car 8 , depuis 9 12 an 2 , le 7 régime 2 de 7 parti 2 ,
 flotter 1 sur 9 un 7 peuple 2 profondément 6 diviser 1
 au 7 milieu 2 de 9 un 7 univers 2 terriblement 6
 dangereux 3 , se 5 montrer 1 hors 9 de 9 état 2
 de 9 assurer 1 le 7 conduite 2 de 7 affaire 2 .
 non 6 point 2 par 9 incapacité 2 ni 8 par 9 indignité 2
 de 7 homme 2 .
 celui 5 qui 5 avoir 1 participer 1 au 7 pouvoir 2 sous 9
 le 7 quatrième 3 République 2 être 1 un 7 gens 2
 de 9 valeur 2 , de 9 honnêteté 2 , de 9 patriotisme 2 .

verbe 1, substantif 2, adjectif 3, numéral 4, pronom 5, adverbe 6, déterminant 7, conjonction 8, préposition 9, interjection 0

4.1.2. Formes et catégories

C:\HYPERBASVELYSEE.EXE

Sommaire Retour N° Mots 173 Lettres 948 Page 1 DG 3

Ecartis Textes Recherche Notes Code/Lemme page CLIC sur un mot : autres contextes

L' unité française se brisait .
 La guerre civile allait commencer .
 Aux yeux du monde , la France paraissait sur le point de se dissoudre .
 C' est alors que j' ai assumé la charge de gouverner notre pays .
 Le drame de l' Algérie , bouleversant les populations , mettant l' armée à dure épreuve , soulevant sur place une vague d' indignation pour ce qui était du présent , de changement et de fraternité pour ce qui était de l' avenir , a déclenché cette crise nationale .
 Mais , de toutes les manières , celle - ci devait éclater .
 Car , depuis 12 ans , le régime des partis , flottant sur un peuple profondément divisé au milieu d' un univers terriblement dangereux , se montrait hors d' état d' assurer la conduite des affaires .
 Non point par incapacité ni par indignité des hommes .
 Ceux qui ont participé au pouvoir sous la quatrième République étaient des gens de valeur , d' honnêteté , de patriotisme .

_Da_ms_T _Nc_fs_T _Afpfs_T _Px3__V _Vmü3sV
 _Yps
 _Da_fs_T _Nc_fs_T _Afpfs_T _Vmü3sV
 _Vmn__V _Yps
 _Da_p_H _Nc_mp_H _Da_ms_H _Nc_ms_H
 _Ypw _Da_fs_T _Np_fs_T _Vmü3sV
 _Sp__H _Da_ms_H _Nc_ms_H _Sp__H
 _Px3_dH _Vmn__H _Yps
 _Pd__nS _Vmip3sV _Rgp__8 _Cs _Pp1__sS
 _Vaip1sV _VmpasmV _Da_fs_D _Nc_fs_D
 _Sp__H _Vmn__H _Ds1_s_D _Nc_m_D _Yps
 _Da_ms_P _Nc_ms_P _Sp__P _Da_ms_P
 _Np_fs_P _Ypw _Vmpp
 _Da_p_D _Nc_fp_D _Ypw _Vmpp _Da_ms_P
 _Nc_fs_P _Sp__F
 _Afpfs_F _Nc_fs_F _Ypw _Vmpp _Sp__L
 _Nc_fs_L _Da_fs_P
 _Nc_fs_P _Sp__P _Nc_fs_P _Sp__H _Pd__dH
 _Pr__nS _Vmü3sV
 _Da_ms_H _Nc_ms_H _Ypw _Sp__H _Nc_ms_H
 _Cc__H _Sp__H _Nc_fs_H
 _Sp__H _Pd__dH _Pr__nS _Vmü3sV _Sp__B
 _Da_ms_B _Nc_ms_B _Ypw _Vaip3sV _VmpasmV
 _Dd_fs_T _Nc_fs_T _Afpfs_T _Yps
 _Cc _Ypw _Sp__H _Afpfp_H _Da_p_H
 _Nc_fp_H _Ypw _Pd_fs_P _Nc__T
 _Rgp__L _Vmü3sV _Vmn__V _Yps
 _Cc _Ypw _Sp__H _Mc__H _Nc_mp_H
 _Ypw _Da_ms_T _Nc_ms_T _Da_p_T _Nc_mp_T
 _Ypw
 _Vmpp _Sp__F _Da_ms_F _Nc_ms_F _Rgp__F

dnapv
 dnavv
 dndn
 dnvsdnspv
 pvrpvpvdsnvdn
 dnstdn , vdn
 vdnstan
 vsndnsnspivdn
 snscnspivsdn
 vvdna
 c , sadn , pnrvv
 c , smn , dndn ,
 vsdnrvdnstdnra
 pvssnsvdn
 msnscn
 plvvdnsdanvdsn
 , sn , sn

verbe 1, substantif 2, adjectif 3, numéral 4, pronom 5, adverbe 6, déterminant 7, conjonction 8, préposition 9, interjection 0

Informations sur le codage grammatical

4.1.3. Concordance dnav n=France

C:\HYPERBAS\ELYSEE.EXE

FormeLemmeCodeSyntaxeExpr.InitialFinalChainListeToutNb11CONCORDANCETrierNotesRetourSommaire

Clic sur une ligne pour voir la page

DG 364a|comme l' action propre à la France nouvelle est à la fois inélucta

DG 459a|Progrès démographique . La France moderne pourrait compter 100000

DG 908a|vaillant bataillon somali que la France libre avait recruté et qui s' e

GP 1288a|où la Corse et avec elle la France entière célèbrent le deuxième c

GP 1569a|sommes Européens . Seule une France forte permettra de trouver son

GP 2359a|nos intérêts essentiels . Une France forte est seule capable d' avoi

GP 2359a|alliés et pas des maîtres . Une France forte est seule capable de coop

VG 2893a|pas très connu c' est que la France actuelle est plus sure que la F

VG 3146a|ologique et non sociologique . La France sociologique est largement

M2 7136a|chance pour la France , que la France européenne sera plus forte et p

JC 9544a|ponsables les uns des autres . La France blessée veut se retrouver rasse

4.1.4. Concordance (N, V, A) nom = France, verbe = être,

The screenshot displays the NooJ software interface with three main windows:

- Locate a pattern in AI_vo_No**: This window shows the search pattern configuration. The "Pattern is:" section has "a NooJ regular expression:" selected, with the pattern "<France><être, V><A>" entered. The "Limitation" section shows "Only: 100 matches" selected.
- Concordance for Text AI_vo_No.not**: This window displays the results of the search in a table format.

Text	Before	Seq.	After
e de son avenir, la	France est résolue/	à poursuivre l'oeuvre immense de sa rénovation i	
vous. Car, quand la	France est malheureuse/	, il n'y a pas de bonheur pour les Français dignes	
la situation de la	France est exceptionnelle/	et puisque, tout cela, nous l'avons réalisé tous en	
Je souhaite que la	France soit intelligente/	en 1977 pour sortir des idées étroites, des conce	
e rayonnement de la	France est grand/	dans le monde, dans cette Europe qu'il faut cons	
ientôt 7 ans que la	France fut défendue/	, écoutée, respectée : elle l'est. □ □ C'était mon d	
r la position de la	France est forte/	à l'étranger, ceux qui voyagent le savent, ils en se	
ue. Grâce à elle la	France sera prête/	avant ses concurrents, à en tirer le meilleur profit	
rer, alors, oui, la	France sera bien/	partie pour le siècle qui vient ! □ □ Partout, je re	
elles le seront. La	France est dynamique/	, courageuse, ambitieuse. Elle a besoin d'abord d	
- AI_vo_No.not**: This window shows the text being analyzed, "De tout mon coeur, je souhaite une bonne année à la France. Par là même, mes voeux vont à chacune et à chacun de vous. Car, quand la France est malheureuse, il n'y a pas de bonheur pour les Français dignes de ce nom. Mais, quand la France réussit, tous ses enfants voient grandir leurs chances. Oui, tous ! c'est-à-dire, ceux de notre métropole, ceux de nos départements et territoires d'outre-mer, ceux qui vivent à l'étranger, enfin, cas très émouvant et qui nous est d'autant plus cher, ceux de la nation française au Canada." Below the text, a list of annotations is shown, including "France,N+z1", "est,A+z1+m+s", "malheureux,A+z1+f+s", and "il,PRO+z1+3".

5. Quelques liens

- Logiciels
- Corpus

5.1. Logiciels

- Hyperbase : Etienne Brunet, UFR Lettres, 98 bd Herriot, 06204 Nice Fax : (33) 04 93 37 54 45 Courriel : brunet@unice.fr
- [Lexico 3](http://www.cavi.univ-paris3.fr/Ilpga/ilpga/tal/lexicoWWW/) <http://www.cavi.univ-paris3.fr/Ilpga/ilpga/tal/lexicoWWW/>
- [Weblex](http://weblex.ens-lsh.fr/doc/weblex/index.html#sommaire) <http://weblex.ens-lsh.fr/doc/weblex/index.html#sommaire>
- [Nooj](http://perso.orange.fr/rosavram/pages/noojpag.html) <http://perso.orange.fr/rosavram/pages/noojpag.html>

5.2. Corpus

- [Damon Mayaffre : Politext](http://www.unice.fr/ILF-CNRS/politext/) <http://www.unice.fr/ILF-CNRS/politext/>
- [La documentation Française](http://discours-publics.vie-publique.fr/rechlogos/servlet/RechServlet?_page=ACCUEIL&_type=NEW) http://discours-publics.vie-publique.fr/rechlogos/servlet/RechServlet?_page=ACCUEIL&_type=NEW
- [Textopol](http://textopol.free.fr/) <http://textopol.free.fr/>
- [Jean-Marc Leblanc](http://www.univ-paris12.fr/www/labos/ceditec/leblanc.html/) <http://www.univ-paris12.fr/www/labos/ceditec/leblanc.html/>