



Analyse logométrique de la cohabitation Chirac/Jospin (1997-2002)

Damon Mayaffre

► **To cite this version:**

Damon Mayaffre. Analyse logométrique de la cohabitation Chirac/Jospin (1997-2002) : Explication de la défaite de Lionel Jospin à l'élection présidentielle de 2002. G. Purnelle, C. Fairon, A. Dister. JADT 2004, Mar 2004, Louvain-La-Neuve, Belgique. Presses universitaires de Louvain, Le poids des mots - Actes des 7èmes Journées Internationales d'Analyse statistique des Données Textuelles, vol. 2, pp.785-792, 2004, <<http://lexicometrica.univ-paris3.fr/jadt/jadt2004/tocJADT2004.htm>>. <hal-00554802>

HAL Id: hal-00554802

<https://hal.archives-ouvertes.fr/hal-00554802>

Submitted on 11 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

VOLUME

2

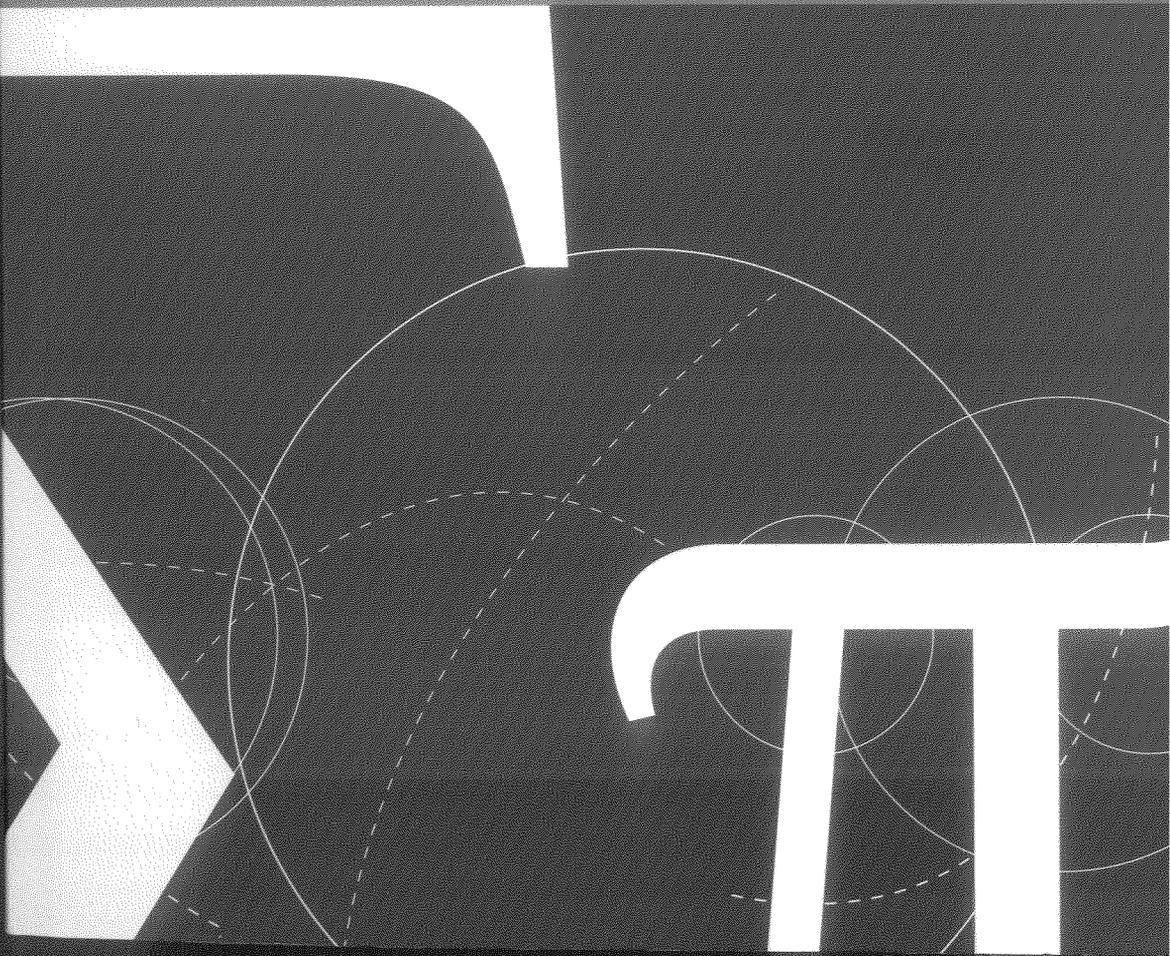
JADT 04

Le poids des mots

ACTES DES 7^{ES} JOURNÉES INTERNATIONALES D'ANALYSE STATISTIQUE
DES DONNÉES TEXTUELLES

Gérald Purnelle, Cédric Faron, Anne Dister (éds)

UCL PRESSES
UNIVERSITAIRES
DE LOUVAIN



Analyse logométrique de la cohabitation Chirac/Jospin (1997-2002).

Explication de la défaite de Lionel Jospin à l'élection présidentielle de 2002

Damon Mayaffre

CNRS-ILF-Université de Nice
UMR 6039, Bases, Corpus et Langage (Nice)
mayaffre@unice.fr

Abstract

We propose to use the logometric method to study the corpus of the Chirac-Jospin cohabitation (1997-2002). Among other results, the method enables one to account for the Prime Minister's defeat : in 2002 a slow evolution of his discourse led to a break-away or divergence from his discourse of the previous years. As will be shown both by the figures measuring intertextual distances and by tree and factor analyses of the vocabulary or the grammatical tags, the socialist locutor has blurred the identity of his speech to the extent of no longer making himself heard.

Résumé

L'on se propose d'analyser, grâce à la méthode logométrique, le corpus de la cohabitation Jospin/Chirac (1997-2002). Cette étude, entre autres résultats, permet d'expliquer la défaite du Premier ministre : à la suite d'un processus linguistico-politique, le discours qu'il tient en 2002 apparaît en rupture avec celui qu'il tenait les années précédentes. Comme le montreront le calcul des distances intertextuelles, les analyses factorielles et arborées réalisées sur le vocabulaire et les codes grammaticaux, le locuteur socialiste brouille l'identité de son discours jusqu'à ne plus se faire entendre.

Mots-clés : logométrie, distance intertextuelle, analyse arborée, individuation socio-linguistique, discours politique, Chirac/Jospin.

1. Introduction

Les périodes de cohabitation sont toujours des moments particuliers de la vie politique française. Mais la cohabitation Chirac/Jospin est exceptionnelle à double titre : d'abord par sa longueur (5 ans), ensuite par la situation originale du président, qui sans avoir souhaité explicitement la division de l'exécutif français l'a directement provoquée par la dissolution de 1997. Contrairement aux cohabitations Mitterrand/Chirac (1986-1988) et Mitterrand/Balladur (1993-1995), la cohabitation Chirac/Jospin (1997-2002) n'est pas une courte parenthèse politique imposée par le calendrier, mais un moment prolongé, « choisi » par les deux parties, durant lequel la France, non sans schizophrénie politique, apprend à parler à deux voix.

C'est cette polyphonie politique que nous voulons mettre à l'épreuve en étudiant les discours de Chirac et Jospin grâce aux traitements logométriques et au logiciel *Hyperbase*.

Corpus

Le corpus étudié est comparatif (conditions du traitement logométrique) : il est contrastif (Chirac vs. Jospin) et diachronique (1997-2002). Il est composé de tous les discours que le président et le Premier ministre ont prononcés ou écrits en France entre 1997 et 2002 soit quelque 2100 discours, 2.400.000 mots, 140.545 vocables (tableau 1).

Textes	Nombre de discours	Nombres d'occurrences	Nombre de vocables différents
Chirac1997	72	104.553	8902
Chirac1998	102	183.820	11.910
Chirac1999	110	176.352	10.989
Chirac2000	129	251.570	12.917
Chirac2001	122	219.236	12.110
Chirac2002	69	209.694	11.638
<i>Total partiel (1)</i>	<i>604</i>	<i>1.145.225</i>	
Jospin1997	68	56.229	5.935
Jospin1998	98	190.455	11.736
Jospin1999	107	225.158	12.915
Jospin2000	106	300.308	15.035
Jospin2001	113	276.747	14.077
Jospin2002	77	227.700	12.381
<i>Total partiel (2)</i>	<i>569</i>	<i>1.276.597</i>	
Total	1173	2.421.822	34.121

Tableau 1. Le corpus de la cohabitation Chirac/Jospin (1997-2002)

Méthode

Nous entendons par traitement logométrique d'une part la batterie d'outils documentaires et statistiques aujourd'hui disponibles sur des logiciels comme *Hyperbase* ou *Lexico* (spécificités, AFC, analyse arborée, calcul des distances, etc.), et d'autre part le fait que ces outils s'appliquent à tous les niveaux linguistiques (mots, lemmes, codes grammaticaux, enchaînements syntaxiques).

Par là cette contribution espère participer à l'amélioration de l'ADT à double titre, d'une part en comparant la performance des différents outils : l'analyse arborée notamment sera avantageusement mise en parallèle avec l'analyse factorielle. D'autre part, en comparant les niveaux de l'analyse linguistique : du point de vue de la distance entre les textes, les analyses sur les formes graphiques, les lemmes ou les fonctions grammaticales aboutissent à des résultats voisins, comme si un texte formait un tout linguistique aux dimensions multiples mais à l'unité remarquable.

Problématique socio-linguistique

Face à un corps de discours segmenté politiquement et chronologiquement, deux types de questionnements peuvent être formulés. Premièrement y a-t-il une identité linguistique chiracienne et une identité linguistique jospinienne ? Les deux hommes tiennent-ils le même langage ou bien des signes linguistiques forts et originaux marquent-ils leurs discours respectifs ? Cette interrogation prolonge notre réflexion sur l'individuation socio-linguistique des locuteurs politiques menée lors des dernières JADT à St. Malo (Mayaffre, 2002).

Deuxièmement, quel est le rythme et l'ampleur de l'évolution chronologique dans le corpus ? Les discours sont-ils identiques à eux-mêmes entre 1997 et 2002 ou changent-ils au fil des ans de manière continue et/ou chaotique ? Comment, finalement, évolue le parcours des deux hommes sous la cohabitation ?

Surtout, les deux types de questionnements peuvent être croisés. Quel est le facteur qui apparaît le plus prégnant dans le corpus ? Le facteur politique ou le facteur chronologique ? L'identité politico-linguistique de Chirac et Jospin résiste-t-elle aux contraintes imposées par le temps ? Les discours de Chirac — quelle que soit la conjoncture — s'opposent-ils toujours à ceux de Jospin ? Ou, au contraire, le temps, avec les thèmes qu'il charrie, commande-t-il les discours pour rapprocher les 2 hommes ?

Par là, le dénouement de 2002 et l'humiliante défaite de Jospin pourront être illustrés : à la suite d'un processus linguistico-politique, le discours que le Premier ministre tient en 2002 apparaît en rupture avec celui qu'il tenait les années précédentes, le locuteur socialiste brouille l'identité de son discours jusqu'à le rendre inaudible.

2. Identités discursives : deux hommes, deux statuts, deux discours

L'objet principal de cette contribution n'est pas là. Mais posons que Chirac et Jospin ont une identité linguistique différente¹. Plusieurs raisons — qu'il serait vain ici de vouloir démêler — expliquent cette individualisation linguistique : les deux principales étant le statut des locuteurs (président/Premier ministre) et la culture politique des deux hommes (droite/gauche).

Les spécificités lexicales des deux hommes sont en tout cas éloquentes (tableau 2).

Spécificités de Chirac (1997-2002)	Spécificités de Jospin (1997-2002)
Notre (+29)	Gouvernement (+32)
Tout (+26)	Ce (déterminant) (+26)
Vous (+24)	Député (+24)
Naturellement (+21)	Corse (nom) (+23)
Monde (+20)	Franc (nom) (+20)
France (+20)	Ministre (+19)
Votre (20)	Politique (nom) (+19)
Ce (pronom) (+19)	Loi (+19)
Et (+19)	1997 (+18)
Son (+17)	Emploi (+17)

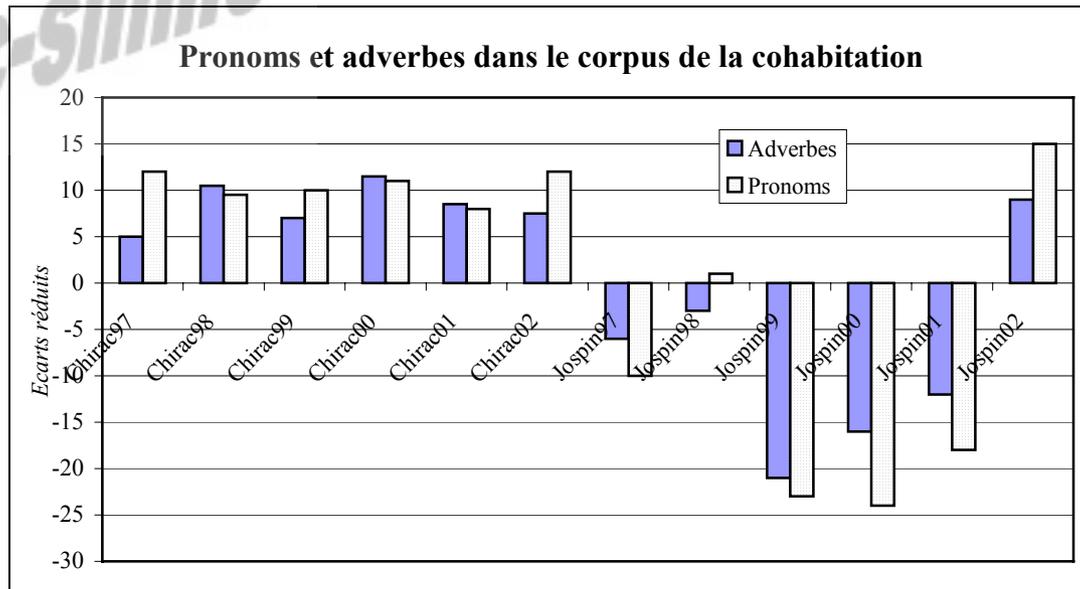
Tableau 2. Spécificités comparées de Chirac et Jospin (1997-2002)

Les spécificités de Chirac (ici les 10 plus importantes) relèvent — à l'exception de « France » et « monde » — de la mise en forme du discours. Il s'agit de mots-outils sans charge politique : « notre », « tout », « vous », « naturellement » etc. Les spécificités de Jospin sont, elles, de l'ordre du contenu politique du discours, qu'il s'agisse de termes institutionnels de la République ou de mots fortement estampillés de la langue politique : « gouvernement », « député », « loi », « emploi », « politique ». Du reste, la vacuité politico-linguistique est une

¹ Pour le développement qui suit cf. Mayaffre (2004 : chapitre 9).

des caractéristiques du discours de Chirac, tant au regard de Jospin que du discours d'un de Gaulle ou d'un Pompidou (Mayaffre, 2004 : Section I et II).

Grammaticalement, cela a une traduction nette et tout aussi explicite. Chirac (par rapport à Jospin) sur-utilise avant tout les adverbes ou les pronoms (graphique 1) et sous-utilise les noms porteurs de sens ou les adjectifs.



Graphique 1. Les pronoms et Les adverbes dans le corpus Cohabitation (1997-2002)²

En réalité, le discours présidentiel chiraquien ne dit pas grand chose, mais est fortement modalisé. Le discours de Jospin essaye de véhiculer un message quitte à paraître technique ou impersonnel (Mayaffre, 2004).

	C.97	C.98	C.99	C.00	C.01	C.02	J.97	J.98	J.99	J.00	J.01	J.02
C.97	0	0,110	0,116	0,113	0,124	0,145	0,197	0,167	0,183	0,179	0,173	0,168
C.98	0,110	0	0,089	0,094	0,098	0,119	0,182	0,148	0,168	0,167	0,162	0,151
C.99	0,116	0,089	0	0,088	0,092	0,123	0,180	0,151	0,160	0,163	0,160	0,155
C.00	0,113	0,094	0,088	0	0,083	0,118	0,171	0,142	0,155	0,150	0,147	0,146
C.01	0,124	0,098	0,092	0,083	0	0,107	0,166	0,143	0,150	0,150	0,139	0,148
C.02	0,145	0,119	0,123	0,118	0,107	0	0,178	0,150	0,168	0,168	0,161	0,139
J.97	0,197	0,182	0,180	0,171	0,166	0,178	0	0,133	0,123	0,124	0,126	0,175
J.98	0,167	0,148	0,151	0,142	0,143	0,150	0,133	0	0,114	0,118	0,113	0,131
J.99	0,183	0,168	0,160	0,155	0,150	0,168	0,123	0,114	0	0,094	0,096	0,158
J.00	0,179	0,167	0,163	0,150	0,150	0,168	0,124	0,118	0,094	0	0,092	0,152
J.01	0,173	0,162	0,160	0,147	0,139	0,161	0,126	0,113	0,096	0,092	0	0,144
J.02	0,168	0,151	0,155	0,146	0,148	0,139	0,175	0,131	0,158	0,152	0,144	0

Tableau 3.

Indices des distances entre textes selon le vocabulaire dans le corpus Cohabitation (1997-2002)

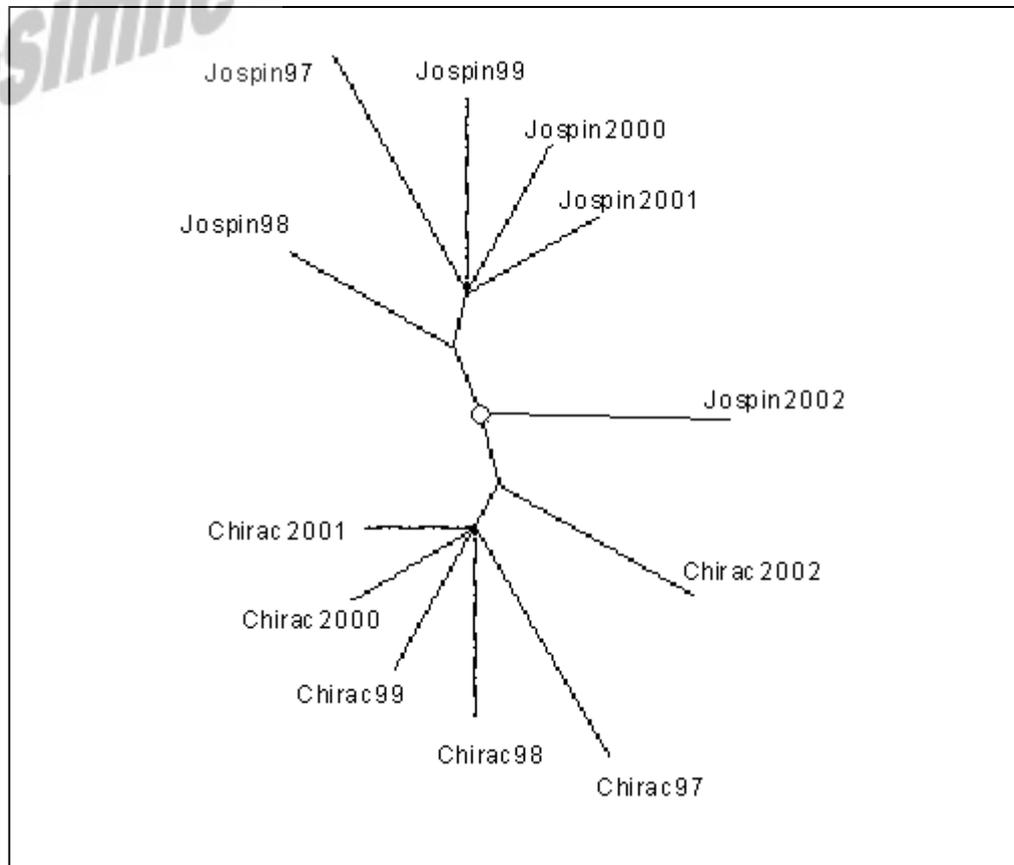
² La particularité de Jospin2002, évidente sur le graphique, fera l'objet de notre attention plus bas.

3. Le rythme de la cohabitation

Les identités respectives de Chirac et Jospin ainsi établies à grands traits, il convient de les mettre à l'épreuve du temps à l'image du dernier graphique. Le traitement diachronique sera ici global grâce à l'outil de synthèse que constitue l'analyse arborée.

Un indice de distance intertextuelle, calculé selon la méthode Labbé (2003), sur l'ensemble du vocabulaire employé dans les discours, a été établi pour aboutir au tableau matriciel suivant (tableau 3).

Une analyse arborée, selon la méthode Luong (1988 ; Barthélemy et Luong, 1998), qui allie techniques classificatoires et topologie mathématique³, donne une représentation fine et précise de cette grande matrice (graphique 2).



Graphique 2. analyse arborée selon le vocabulaire du corpus Cohabitation (1997-2002)

À partir d'un tel arbre, trois conclusions socio-linguistiques peuvent être développées.

D'abord, il existe une individuation linguistique réelle entre Chirac et Jospin. L'algorithme dessine nettement deux branches opposées. Sur l'une se distinguent toutes les ramifications de Chirac, sur l'autre toutes les ramifications de Jospin. (On enregistre une seule exception, nous y reviendrons, le discours de Jospin 2002 qui forme à lui tout seul une branche indépendante.) Le temps ne perturbe que très peu cette dichotomie politique. Après beaucoup d'autres expériences, avec des outils variés, sur des corpus politiques divers, nous pouvons à nouveau affirmer que le facteur individuel et politique semble toujours prépondérant dans le discours

³ L'analyse arborée a déjà fait l'objet de présentation aux JADT : succincte et pragmatique dans Luong et Mellet (1995) ; détaillée et technique dans Barthélemy et Luong (1998).

politique. Lorsque l'on croise plusieurs facteurs — l'identité politique du locuteur (sa personnalité et son statut), la chronologie, etc. — c'est toujours l'identité du locuteur — son appartenance politique et son statut — qui apparaît primordiale (Mayaffre, 2002).

Ensuite, le clivage entre Chirac et Jospin est réel mais *tend à se réduire* au fil de la cohabitation. Sur chacune des branches, c'est-à-dire localement sur l'arbre, le graphe se montre très précis notamment dans la description de la chronologie et de l'évolution du discours durant la période.

Il est très intéressant de remarquer comment le discours de Chirac 2002, par une sous-branche particulière, change en partie de nature et cherche à gommer son identité chiraquienne originelle. Il ne fait aucun doute que la campagne électorale se fait sur des thèmes jusqu'alors inexploités, avec d'autres mots que ceux des discours antérieurs. Pour des raisons électorales, Chirac semble inventer, en 2002, un autre discours. Néanmoins — et la précision prendra tout son sens plus loin —, ce discours de campagne en grande partie original reste clairement affilié au *parler* de Jacques Chirac : il reste, sur l'arbre, sur la même branche maîtresse chiraquienne avec les autres discours du président.

D'ailleurs, chez Chirac, cette évolution très marquée et un peu exceptionnelle du discours de 2002 ne doit pas nous masquer l'évolution chronologique progressive et cohérente de l'ensemble du discours. Sur l'arbre, dans le bouquet des textes de Chirac entre 1997 et 2001, on remarque que la distance (c'est-à-dire l'originalité) tend à se réduire au fil du temps : les ramifications sont en effet de plus en plus courtes, au fur et à mesure que s'égrainent les années. En d'autres termes — et l'observation attentive des indices du tableau 2 peut le confirmer — le discours de Chirac en 1997 est le plus éloigné du discours moyen (c'est-à-dire en grande partie, par effet de contraste dans notre corpus, du discours de Jospin), puis en 1998 le président perd en partie de son originalité comme l'atteste une ramification plus courte, en 1999, il en perd encore davantage, etc.

Du côté Jospin la même tendance est observée. À mesure que le temps s'écoule, Jospin se rapproche d'un discours moyen (c'est-à-dire dans notre corpus du discours chiraquien). Une seule exception perturbe légèrement ce bel ordonnancement chronologique : le discours de 1998 est très précocement édulcoré se démarquant de ses frères pour se rapprocher du centre topologique de l'arbre⁴. Mais au-delà de cette exception, c'est le discours de 1997 qui est le plus typé ou le plus marginal (ramification la plus longue), puis celui de 1999, puis celui de 2000, etc.

C'est pourquoi la conclusion est évidente : les deux hommes déteignent l'un sur l'autre, et leurs discours tendent à se rapprocher. L'effet émoulinant de la cohabitation où le clivage gauche/droite est estompé est ainsi illustré. Sans jamais tenir le même discours, Jospin et Chirac perdent petit à petit, entre 1997 et 2002, leur identité propre pour employer de plus en plus un vocabulaire commun.

4. La défaite discursive de Jospin en 2002

Mais l'intérêt de l'arbre présenté se trouve ailleurs, dans la situation particulière du texte Jospin2002. Le discours de Jospin n'évolue pas seulement mais transmue durant l'année 2002 pour devenir méconnaissable. Que les discours des deux hommes tendent à se rapprocher a déjà été dit. Que, symétriquement à Jospin, le discours de campagne de Chirac soit particulier

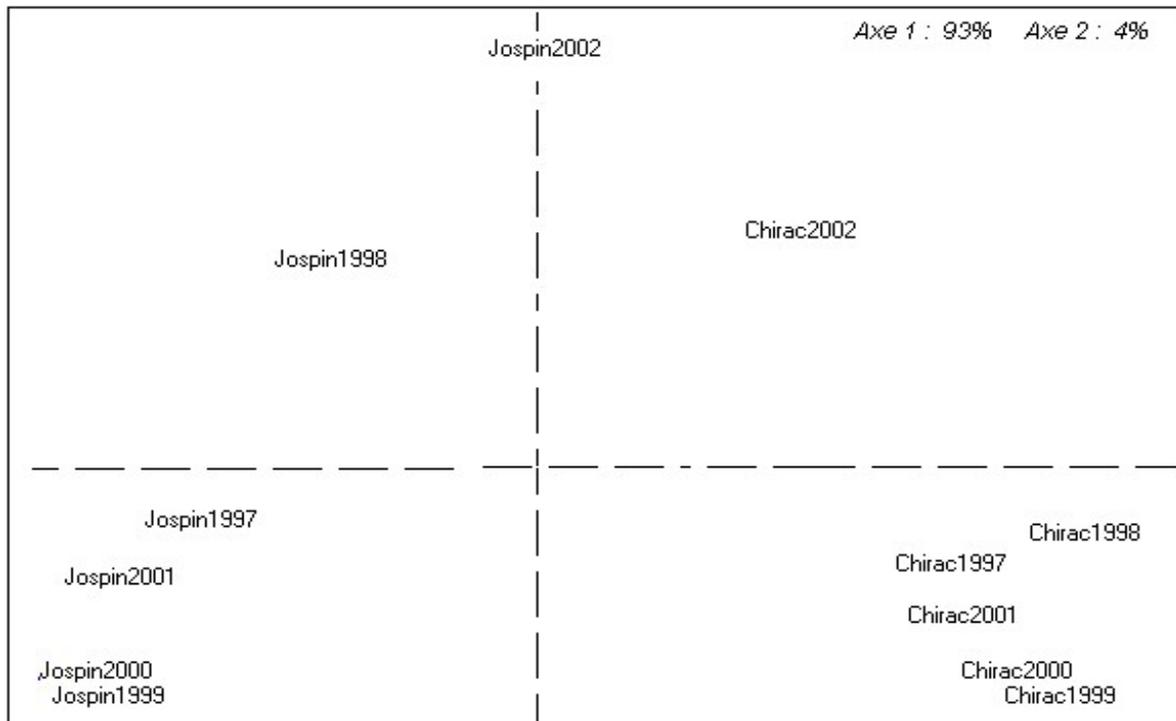
⁴ Une explication de la particularité de l'année 1998 durant la gouvernance de Lionel Jospin est donnée dans Mayaffre (2004).

et opportuniste a déjà été souligné. Mais la localisation de Jospin2002 dépasse tous ces constats pour expliquer la défaite du Premier ministre aux élections présidentielles.

Il ne s'agit plus d'une évolution du discours mais d'une rupture linguistique étonnante. Dans le clivage linguistico-politique Jospin / Chirac dont elle donne une belle représentation, l'analyse arborée n'arrive plus à reconnaître et à classer Jospin2002. Le discours de campagne du Premier ministre se trouve aussi éloigné du discours de Chirac que des siens propres durant les années précédentes. En 2002 Jospin tient un discours à part, original, mutant, qui ne s'apparente pas à celui qu'il a tenu précédemment⁵. La personnalité linguistique (sinon politique) de l'émetteur, que les Français avaient appris à reconnaître durant 5 ans, est totalement brouillée. En termes rapides, Jospin se met subitement, comme Chirac, à utiliser des adverbes et des pronoms et à négliger les noms surtout lorsqu'ils sont politiques (Mayaffre, 2004). Le graphique 1, produit plus haut, n'est qu'une illustration générale du phénomène.

Pour cette raison, le discours du candidat Jospin devient inaudible particulièrement pour l'électorat de gauche, habitué à un autre discours. Est-ce la faute des conseillers en communication ou des plumes du candidat ? Est-ce la faute des médias et de ses adversaires qui ont su entraîner Jospin sur des thématiques qui lui étaient étrangères ? En tout cas les mots que le Premier ministre candidat se met alors à utiliser ne lui appartiennent pas. En 2002, Jospin vocalise un discours étranger, « inidentifiable ». Dans ces conditions, il avait peu de chances de se faire entendre.

Pour conclure, l'analyse est confirmée par une AFC issue du traitement des données lexicales (graphique 5). Ici l'axe 1 (horizontal) et l'axe 2 (verticale) ont été croisés.



Graphique 3. AFC du corpus Cohabitation (1997-2002) selon le vocabulaire

⁵ On remarquera de manière symptomatique que le texte le plus éloigné (0.175) du texte Jospin2002 n'est pas un texte de Chirac mais le texte Jospin1997 : dans le corpus, les discours qui ressemblent le moins aux discours de Jospin 2002 sont les discours de Jospin tenus 5 ans auparavant.

Le schème général de l'AFC est identique à celui de l'arbre. Les conclusions linguistico-politiques, en trois points, sont confirmées : il existe bien une identité lexicale chiraquienne *versus* jospinienne, apparente sur l'axe 1 (axe horizontal recueillant 93% de l'information) ; les discours de Chirac2002 ou Jospin1998 apparaissent comme particuliers dans leur famille respective (particulièrement sur l'axe 2) ; enfin Jospin2002 n'est plus seulement particulier mais totalement a-typique, exactement central sur l'axe 1 et excentré sur l'axe 2.

Nous noterons enfin, à titre méthodologique, que localement sur l'AFC, les concentrés de points (Chirac1997, Chirac1998, Chirac1999, Chirac2000, Chirac2001 et Jospin1997, Jospin1999, Jospin2000, Jospin2001) sont peu précis et difficiles à interpréter sans la consultation de l'axe 3 voire de l'axe 4. L'analyse arborée, elle, nous avait pourtant permis de constater, du premier coup d'œil, l'évolution qui avait cours durant la cohabitation, en soulignant — par la dimension de plus en plus courte au fil des années des ramifications menant aux textes — comment les discours du président et du Premier ministre perdaient, dans la durée, de leur originalité respective pour tendre vers leur confusion.

Références

- Barthélemy J.-P. et Luong X. (1998). Représenter les données textuelles par les arbres. In *Actes des JADT 1998* : 49-73.
- Brunet Ét. (1988). Une mesure de la distance intertextuelle : la connexion lexicale. *Revue informatique et statistique dans les sciences humaines, Le nombre et le texte, hommage à Etienne Evrard* : 81-116.
- Brunet Ét. (2002). Le lemme comme on l'aime. In *Actes des JADT 2002* : 221-233.
- Brunet Ét. (2003). Peut-on mesurer la distance entre deux textes ? *Corpus*, vol. (2) : 47-70.
- Corpus* (2003). Vol. (2), *La distance intertextuelle*.
- Labbé C. et Labbé D. (2003). La distance intertextuelle. *Corpus*, vol. (2) : 95-118.
- Labbé D. et Hubbert P. (1998). La connexion des vocabulaires. In *Actes des JADT 1998* : 361-370.
- Lebart L. et Salem A. (1994). *Statistique textuelle*. Dunod.
- Xuan L. (1988). *Méthodes d'analyse arborée. Algorithmes, applications*, Thèse d'Etat, Paris V.
- Luong X. (1994). « L'analyse arborée des données textuelles : mode d'emploi. *Travaux du cercle linguistique de Nice*, vol. (16) : 25-43.
- Luong X. et Mellet S. (1995). Les calculs multidimensionnels au service de l'analyse syntaxique diachronique. In *Actes des JADT 1995* : 281-289.
- Luong X. et Mellet S. (2003). Mesures de distance grammaticale entre les textes. *Corpus*, vol. (2) : 141-166.
- Mayaffre D. (2002). Discours politique, genres et individualisation socio-linguistique. In *Actes des JADT 2002* : 517-529.
- Mayaffre D. et Luong X. (2003). Arbres et généalogie politique. Représentation arborée du discours de Jacques Chirac (1995-2002). *Histoire et Mesure*, vol. (XVIII-3/4).
- Mayaffre D. (2004), *Jacques Chirac (1995-2003) : le Président de la parole*. Champion, Lettres numériques. (A paraître).
- Pincemin B. (2002). Similarités texte-textes. Expérience d'une application de diffusion ciblée et propositions. In *Matemáticas y Tratamiento de Corpus* (s. n.). Logroño : Fundacion San Millán de la Cogolla : 35-52.