

# Analyse lexicométrique d'un genre : la *déclaration de politique générale*

Pascal Marchand

LERASS (U. Toulouse 3) – F-31077 Toulouse – France

## Abstract

The policy statement of the Prime Minister, which gives rhythm to political life and French media, is a subject for expectations, forecasting, analysis and commentary. The lexicometrical analysis of policy statements allows us both to identify their type and to monitor major developments as they happen in the Fifth Republic. After the last presidential elections, the French Presidency and commentators have called a speech that the new President has delivered to his parliamentary majority a “policy statement”. Analysis of both lexical correspondences and indices of lexical richness as well as banality revealed that the Presidential “roadmap” was not a “policy statement”. In contrast, F. Fillon’s statement complies with the rules of the genre and presents itself as the richest of the Fifth Republic.

**Keywords:** policy statement, Prime minister, analysis of lexical correspondences, lexical richness and banality.

## Résumé

La *déclaration de politique générale* d'un Premier ministre rythme la vie politique et médiatique française et fait l'objet d'attentes, de prévisions, d'analyses et de commentaires. L'analyse lexicométrique des *déclarations de politique générale* permet à la fois d'en cerner le genre et d'en observer l'évolution au long de la Ve République. Après les dernières élections présidentielles, l'Élysée comme les commentateurs ont qualifié de *déclaration de politique générale* un discours que le nouveau Président a prononcé devant sa majorité parlementaire. L'analyse des correspondances lexicales, mais également des indices de richesse et de banalité, révèlent que la « feuille de route » présidentielle n'était pas une *déclaration de politique générale*. En revanche, celle de F. Fillon en respecte bien les lois du genre et se présente comme la plus riche de la Ve République.

**Mots-clés :** déclaration de politique générale, Premier ministre, analyse des correspondances lexicales, richesse et banalité lexicales.

On peut voir la déclaration de politique générale d'un Premier ministre comme un rituel un peu convenu, un marronnier médiatique un peu ennuyeux. On peut aussi y trouver la trace d'une histoire politique, de ses contextes, de ses événements et surtout des ressorts dont le pouvoir use pour s'adresser aux citoyens, les convaincre et les séduire. Si chacun de ces discours est tenu par un Premier ministre particulier, avec son histoire et sa personnalité, il l'est également pour un auditoire et dans un contexte général particulier. Plus qu'un simple exercice de style, il s'agit donc bien d'un témoignage riche en indicateurs historiques et linguistiques. L'analyse lexicométrique (Lebart & Salem, 1994) des *déclarations* de la Ve République permet d'identifier ces évolutions<sup>1</sup>.

---

<sup>1</sup> Les logiciels utilisés pour conduire ces analyses sont issus, soit de l'analyse statistique des données textuelles (*Lexico 3*, d'André Salem ; DTM de Ludovic Lebart), soit de l'analyse automatique du contenu (*Tropes*, de Pierre Molette). Voir : Marchand, 2007.

## 1. Caractéristiques du corpus

Après segmentation, le corpus présente 231 232 occurrences, qui représentent 12 741 formes dont 5 313 hapax. La fréquence maximale est « de » (12 528 occurrences). Après partition des trente-sept *déclarations*, on obtient le tableau 1.

## 2. Analyse des correspondances lexicales

Le tableau lexical croise les 2 097 formes apparaissant plus de 11 fois avec 36 *déclarations* (le discours qu'a prononcé A. Juppé à propos de la réforme de la protection sociale doit être placé en élément illustratif).

Partie	Nb occurrences	Nb formes	Nb hapax	Fréq. Max	Forme
1959 Debré	8810	2017	1188	442	de
1962 Pompidou	4395	1333	875	204	de
1968 Couve de Murville	4182	1176	706	181	de
1969 Chaban-Delmas	1804	742	517	95	de
1972 Messmer	7545	2061	1344	424	de
1974 Chirac	8519	1997	1224	509	de
1976 Barre	7013	1802	1135	388	de
1977 Barre	4121	1228	820	183	de
1978 Barre	6048	1636	1007	354	de
1981 Mauroy	11652	2648	1590	662	de
1981 Mauroy	8711	2168	1333	477	de
1983 Mauroy	9277	2112	1229	563	de
1984 Mauroy	5425	1507	916	287	de
1984 Fabius	5351	1477	953	224	de
1986 Chirac	9416	2415	1511	569	de
1987 Chirac	7856	1840	1109	415	de
1987 Chirac	6687	1837	1160	357	de
1988 Rocard	7610	2091	1366	407	de
1991 Rocard	1794	631	427	71	la
1991 Cresson	5265	1532	983	291	de
1992 Cresson	4266	1381	925	226	de
1992 Bérégoovoy	5484	1556	1006	227	de
1992 Bérégoovoy	2570	801	521	122	de
1993 Balladur	11531	2411	1406	658	de
1993 Balladur	7011	1736	1053	356	de
1995 Juppé	13063	2689	1495	787	de
1995 Juppé	4769	1324	840	322	de
1996 Juppé	5058	1303	787	243	de
1997 Jospin	6715	1794	1105	368	de
2002 Raffarin	9217	2087	1216	511	de
2003 Raffarin	3247	971	608	171	de
2004 Raffarin	3717	1040	640	212	de
2005 de Villepin	6595	1656	971	394	de
2006 de Villepin	4833	1366	871	247	de
2006 de Villepin	3527	1027	673	175	de
2007 Fillon	8148	2294	1546	409	de

Tableau 1 : Partition du corpus des déclarations de politique générale (1959-2007)



Les années 1970 sont celles de la « lutte contre l'inflation », mais également de termes comme « gouvernement », « progrès », « vie », « actions », « programme », « politique ». En revanche, l'utilisation des formes du « nous » est rare.

Les années 1980 sont les années du vocabulaire technique : des dates, des chiffres, des pourcentages, des « entreprises ». C'est aussi l'apparition du « nous » qui va avec la disparition de « gouvernement ».

Les années 1990 sont celles où apparaissent les termes « réforme », « emploi », « intégration », « immigration », « Europe », « exclusion », « pacte républicain ». Mais on trouve également « je », « il faut » et « l'État ».

Les années 2000 confirment cette tendance : on retrouve « l'emploi », mais aussi « nous ». On va trouver également « service public », « vérité », « égalité », « chances », « jeunes », « contrat », « accompagnement ». Mais on ne dit plus « gouvernement » et on n'emploie plus la négation (« ne/n'... pas »).

## 2.2. L'axe de la confiance (facteur 2)

L'analyse différencie, sur le deuxième facteur, deux cas bien distincts :

- les premières *déclarations*, prononcées lors de la nomination du nouveau Premier ministre (ou « discours d'investiture »), souvent, mais non exclusivement, consécutivement à une élection et dans une situation de bonne popularité (60,56 % de confiance moyenne)
- les déclarations suivantes, souvent provoquées par une motion de censure et dans une situation de relative impopularité. (40,72 % de confiance moyenne)

S'il est dans une situation de confiance, le Premier ministre aura tendance à prononcer un discours plus long et plutôt nominal. Les formes les plus significatives des « discours d'investiture » sont les suivantes : *doit, et, la, l', une, défense, coopération, monde, atlantique, républicain, qualité, vie, parlement, nation, publique, mer, sécurité, organisation, indépendance, publics, rapports...*

En revanche, si le Premier ministre est dans une situation de défiance, le discours sera plus court, plutôt verbal et fortement pronominalisé. Les formes les plus significatives sont : *vous, je, nous, Monsieur, m', me, votre*. On trouve également les verbes *avons, avez, suis, faisons, était, ai, dites, été, permis, pris, voudrais, avais, a, ont*. D'autres termes renvoient au débat ou à la polémique : *majorité, opposition, gauche, vrai, pas, alors, jamais, scrutin, oui, n...* On note enfin les dates (1981, 1982, 1983, 1984, 1986, 1987, novembre, mois) et les chiffres (% , mille, huit, milliards, sept, vingt, cent, deux, cinq, dix...).

Ces deux attitudes à l'égard du langage<sup>3</sup> discriminent avec une remarquable constance les discours d'investiture et les *déclarations* suivantes, du moins jusqu'aux *déclarations* de Jean-Pierre Raffarin. Avec ce dernier, les trois *déclarations* sont proches les unes des autres et dans un niveau « moyen » sur le deuxième facteur.

---

<sup>3</sup> Le terme d'*attitude à l'égard du langage* est issu de la théorie de Jean-Léon Beauvois et Rodolphe Ghiglione (1981), qui distinguaient une attitude paradigmatique et une attitude syntagmatique : le deuxième facteur de l'analyse semble bien renvoyer à une telle distinction.

### 2.3. 2007 : confirmation et innovation

Après la dernière élection présidentielle (6 mai 2007), le nouveau Président a convoqué les parlementaires de sa majorité avant la *déclaration de politique générale* de son Premier ministre, pour leur livrer sa « feuille de route ». Les commentateurs ont été unanimes pour dire qu'il avait prononcé un *discours de politique générale*. Or, la figure 2 montre que ce discours ne partage pas les propriétés d'une *déclaration de politique générale* : les procédures statistiques (AFC) isolent ce discours et l'opposent à l'ensemble des autres *déclarations* de la Ve République.

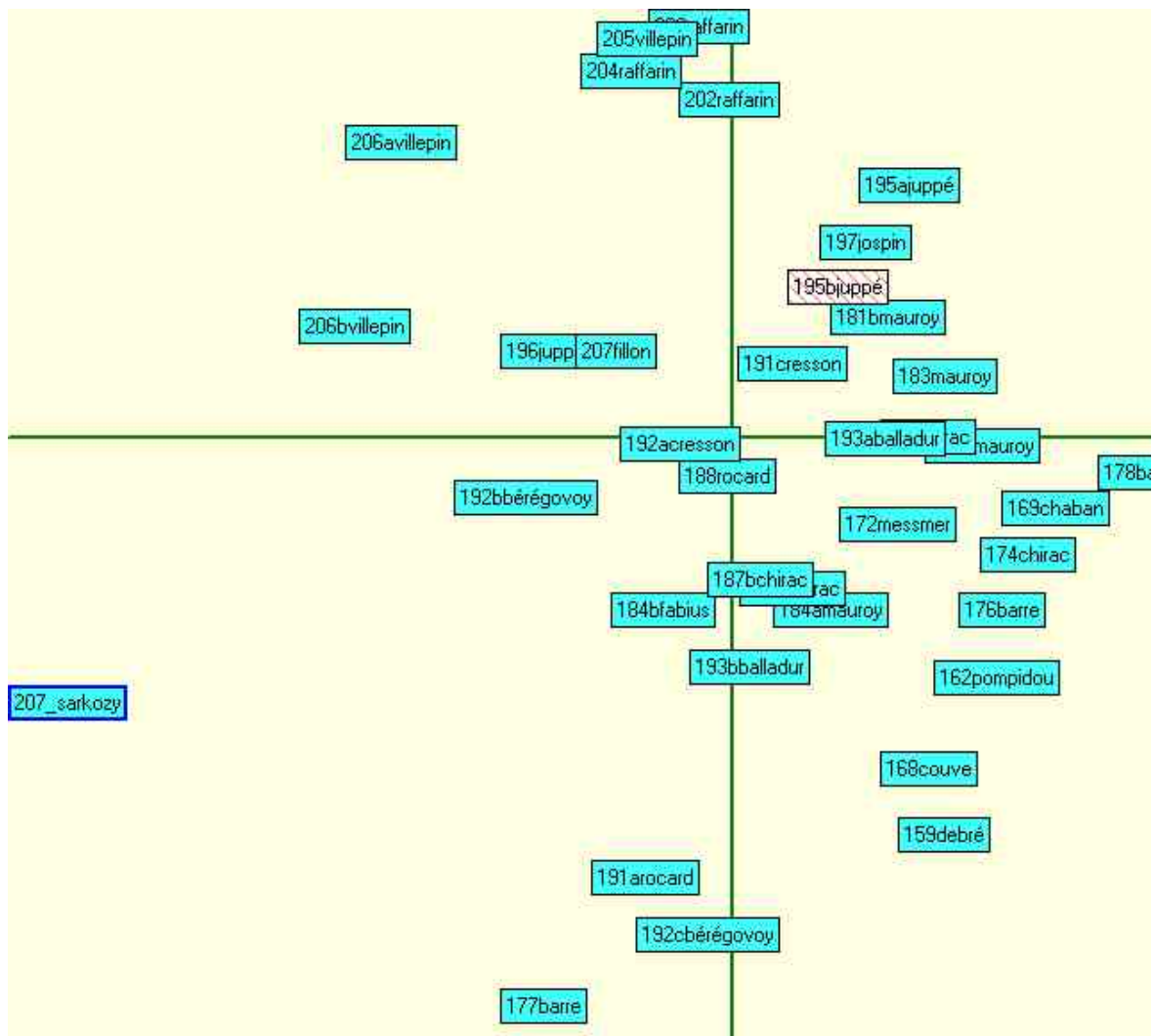


Figure 2 : AFC des déclarations de politique générale (1959-2007) et de la « feuille de route » de N. Sarkozy

Comparativement, on peut ramener les caractéristiques de ce discours présidentiel à quelques traits saillants :

- une forte utilisation des pronoms, notamment « on » (*on ne peut pas* : 13 fois ; *on va* : 7 fois), forme assez rare dans les *déclarations de politique générale*.
- une forte utilisation du « je » (*je veux prendre mes responsabilités* : 14 fois ; *je vais* : 6 fois), ce qui est moins rare, mais prend ici des proportions jamais atteintes par les Premiers ministres. La forme « je » s'accompagne souvent d'un évitement de la forme

« le gouvernement », et c'est également le cas dans le discours de Nicolas Sarkozy : le terme *gouvernement* est significativement évité, de même que *loi*, *projet*, *solidarité*, *république*, qui font partie des termes les plus fréquents (Fmax) des *déclarations de politique générale*.

- Un évitement des marques du « nous » (et notamment « notre pays »).
- une très grande proportion de négation (*ce n'est pas* : 23 fois), ce qui tendait à disparaître des dernières *déclarations de politique générale* depuis une dizaine d'années (particulièrement dans celles de Jean-Pierre Raffarin).

La *déclaration de politique générale* que François Fillon prononce, le mardi 3 juillet 2007, comprend 8 148 mots au total, et un lexique de 2 294 mots différents. Elle compte parmi les discours les plus longs (la moyenne est à 6 400 mots) mais se place en dixième position.

A la différence de la « feuille de route » de Nicolas Sarkozy, l'analyse statistique confirme qu'il s'agit bien d'une *déclaration de politique générale* et qui, de plus, se situe dans la mouvance descriptive des dernières années (sur *l'axe du temps*) et mobilise un style grammatical conforme aux discours « d'investiture » (sur *l'axe de la confiance*).

Les mots privilégiés par le nouveau Premier ministre sont : *identité*, *université*, *culture*, *fondements*, *enjeux*, *excellence*... Mais ce sont surtout des segments qui font la spécificité de cette déclaration : « *la France ne se* (dérobent, résignent, résigneront) », « *la France est grande lorsqu'elle* (défend, s'engage, milite, multiplie, montre) », « *(la ou cette) vieille croissance* (c'est) », « *une croissance nouvelle, forte, saine, solidaire* », « *il faut une* ».

En revanche, François Fillon a tendance à éviter des termes comme *aujourd'hui*, *économie*, *mois*, *majorité*, *avenir*, *esprit*, *développement*, *la communauté*, *la politique*, *mesures*, *nouvelles*, *la situation*, *je*, *action*, *développement*, *vie*... Il y a également un évitement systématique de certains « mots-outils » : *qu'* et *que*, *encore*, *ou*, *ainsi*, *mais*, *à*...

Ces évitements concernent des formes habituellement fréquentes dans les *déclarations*, ce qui attire notre attention sur l'importance de la richesse et de la banalité.

### 3. Richesse et banalité

Il y a plusieurs façons de calculer la richesse et la littérature en la matière est abondante. Nous avons utilisé, ailleurs, l'indice de Charles Muller, basé sur la loi binomiale (logiciel Hyperbase d'E. Brunet). Nous nous intéresserons ici à la distribution des hapax (les derniers rangs de l'index) et des formes de fortes fréquences (les premiers rangs de l'index).

L'opérationnalisation a donc recours aux « types généralisés » ou « Tgens » (voir Lamalle & Salem, 2002 ; Mange, Marchand & Salem, 2006). Ici, les termes les plus utilisés dans les *déclarations de politique générale* sont : France (862), gouvernement (790), politique (673), pays (589), français (478), État (440), entreprises (367), emploi (354), travail (303), économique (291), aujourd'hui (289), sociale (271), monde (268), République (259), action (258), loi (257), Europe (242), effort (233), avenir (232), développement (230), économie (230), nationale (228)... Certains Premiers ministres puisent largement dans ce vocabulaire banal (A. Juppé en 1996, J.-P. Raffarin, D. de Villepin dans sa dernière *déclaration*), d'autres beaucoup moins (R. Barre, M. Rocard). Si N. Sarkozy atteint, dans l'usage de la banalité, des seuils statistiques sans équivalent dans aucune autre *déclaration*, ce n'est pas le cas de F. Fillon.

Dans le graphique suivant, on montre la distribution de ces deux indicateurs : l'utilisation d'hapax (barres noires) et des fmax (barres blanches)<sup>4</sup>.

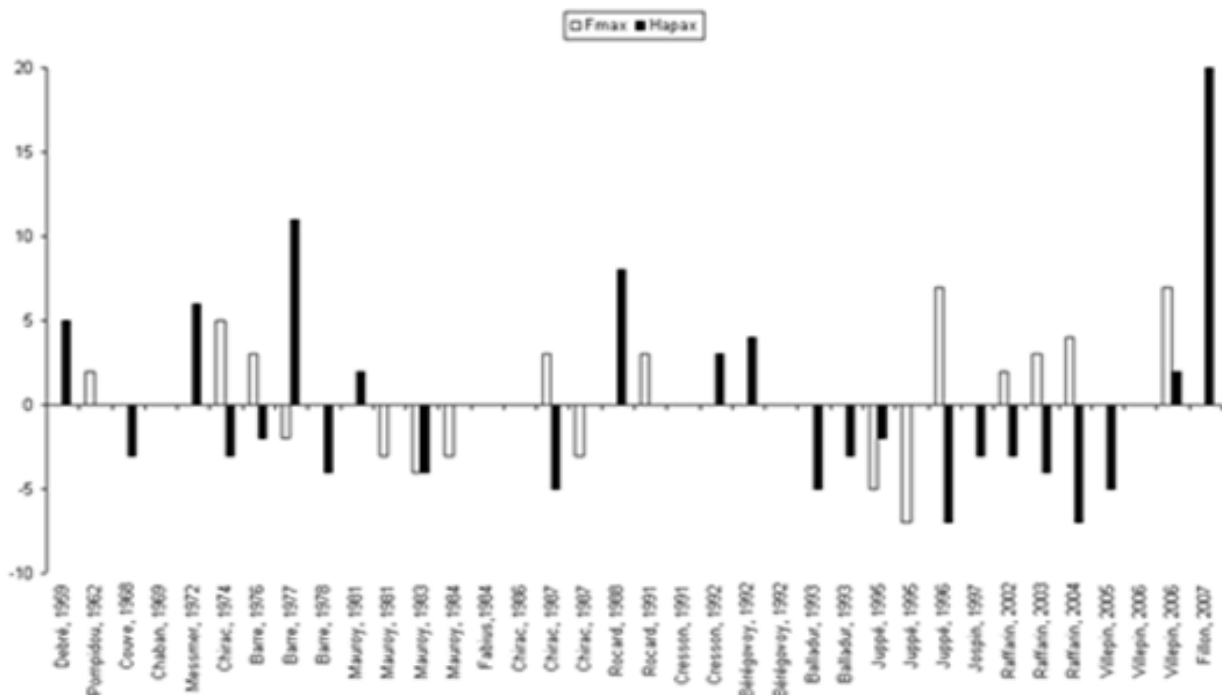


Figure 3 : Ventilation des fréquences max. et des hapax

On confirme ainsi que, contrairement à ses prédécesseurs et à son Président, F. Fillon ne puise pas particulièrement dans ce vocabulaire « banal ». Au contraire, le calcul d'accroissement du vocabulaire donne une confirmation, en indiquant qu'un afflux de mots nouveaux caractérise sa déclaration. Et la proportion d'hapax montre qu'il use de formes rares et révèle une richesse lexicale tout à fait particulière<sup>5</sup>. Il parvient même à produire un discours plus riche (lexicométriquement) que Raymond Barre en 1977 et Michel Rocard en 1988, qui détenaient le record.

On peut rechercher ces mots exclusifs de François Fillon dans des extraits particulièrement riches en hapax :

- Certains traduisent le souci d'un vocabulaire « recherché », comme dans l'extrait suivant :

(...) une *opportunité* de se *détacher* des *postures* idéologiques et des réflexes *claniques*.

L'immense *cohorte* de nos savants, *biologistes*, *mathématiciens*, philosophes, juristes, *historiens* qui *firent* notre rayonnement ne doit pas s'arrêter au seuil d'un siècle, où, précisément, le pouvoir de la matière *grise* dessinera notre avenir.

- D'autres hapax sont d'une toute autre nature, puisqu'ils renvoient à des événements d'actualité (voire à du « Storytelling ») :

La France est grande lorsqu'elle défend, à travers la libération d'*Ingrid Bétancourt* et des *infirmières bulgares* injustement *condamnées*, les droits *inaliénables* de tout être humain.

<sup>4</sup> L'ordonnée représente la probabilité de rejet de l'hypothèse nulle pour une spécificité positive (coordonnées positives) ou négative (coordonnées négatives).

<sup>5</sup> 67 % des formes graphiques et 19 % des occurrences de François Fillon sont des hapax. Par ailleurs, le corpus des trente-six *déclarations* comporte 5314 hapax et François Fillon à lui seul en représente 313, soit près de 6 %.

Et j'ai en mémoire ce *sous-officier* français, qui, il y a quelques années, m'expliquait que dans un *village* constamment *bombardé* de *Somalie*, la première tâche de sa compagnie *consista* à reconstruire la *maternité* détruite.

Une ultime confirmation de la richesse lexicale de la *déclaration* de François Fillon est encore donnée par la comparaison avec la base textuelle *Frantext*<sup>6</sup> : comme le montre le graphique suivant, le discours de François Fillon se positionne comme le texte qui puise le plus dans les formes rares de la langue française (groupes 1 et 2). En cela, il s'oppose radicalement à Jean-Pierre Raffarin, par exemple.

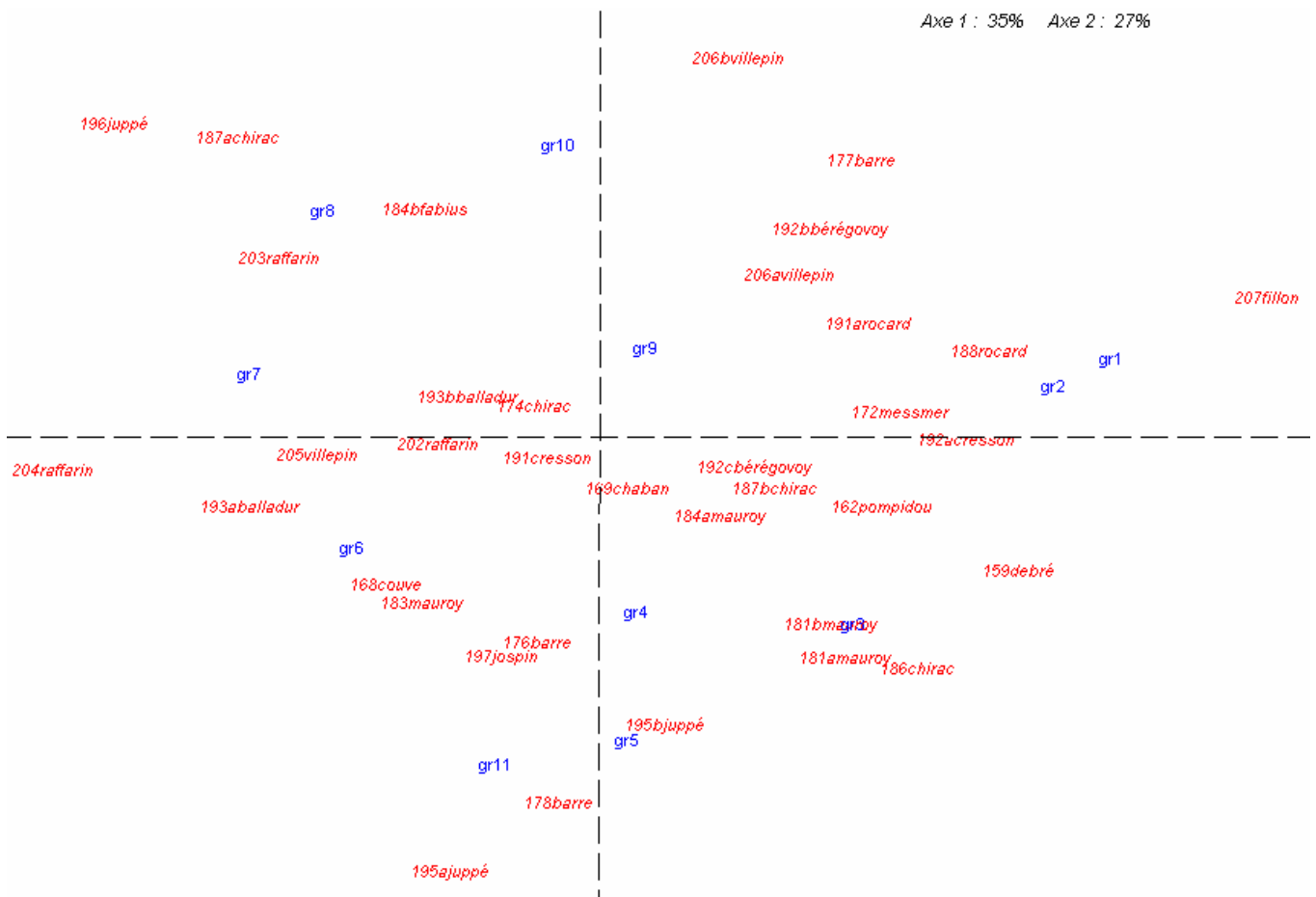


Figure 4: AFC croisant les déclarations de politique générale avec les groupes Frantext.

<sup>6</sup> Le corpus *Frantext* contient près de 4 000 textes appartenant aux domaines des sciences, des arts, de la littérature, des techniques, qui couvrent cinq siècles de littérature (du XVI<sup>e</sup> au XXI<sup>e</sup> siècle). Le logiciel *Hyperbase* (d'Étienne Brunet) range les mots du corpus des *déclarations de politique générale* en onze groupes selon leur fréquence dans *Frantext*. Dans le groupe 1 sont rangés les mots rares, qui apparaissent moins de 500 fois dans *Frantext* ; dans le groupe 2, les mots qui apparaissent entre 500 et 1 024 fois ; dans le groupe 3, entre 1 024 et 2 048 ; dans le groupe 4, entre 2 048 et 4 096 ; et ainsi de suite jusqu'au groupe 11, qui comprend les mots-outils, qui sont aussi les plus nombreux. On observe ainsi une opposition entre les *déclarations* des années 1960-1970 (Michel Debré, 1959 ; Georges Pompidou, 1962 ; Pierre Messmer, 1972), auxquelles il faut ajouter celles de Michel Rocard, et qui puisent dans un vocabulaire rare (groupes 1 à 3), et celles des années 1980-2000 (Laurent Fabius, 1984 ; Jacques Chirac, 1987 ; Édouard Balladur, 1993 ; Jean-Pierre Raffarin, 2002 et 2004 ; Alain Juppé, 1996 ; Dominique de Villepin, 2005 et 2006) qui puisent dans un vocabulaire beaucoup plus courant (groupes 6 à 10).

L'ensemble de ces indices statistiques ne permet donc pas de dire que le discours de N. Sarkozy était un *discours de politique générale*, tant il s'éloigne manifestement du genre<sup>7</sup>. En revanche, cette appréciation médiatique a sans doute eu un effet sur la *déclaration* de François Fillon : il était difficile pour le Premier ministre de tenir un discours véritablement engagé et engageant, qui aurait pu donner une impression de compétition d'ego.

Le discours de F. Fillon représente certainement une rupture avec le contexte de communication politique dans lequel il a été prononcé. Avec une *déclaration* moins marquée par le marketing politique et l'engagement personnel que les précédentes, mais aussi plus équilibrée et plus riche lexicalement, F. Fillon a-t-il traduit l'évolution annoncée vers l'effacement du Premier ministre et une modification de son statut constitutionnel ?

## Références

- Beauvois J.-L. & Ghiglione R. (1981). *L'homme et son langage - Attitudes et enjeux sociaux*. Paris, PUF.
- Lamalle C. & Salem A. (2002). Types généralisés et topographie textuelle dans l'analyse quantitative des corpus textuels. In A. Morin & P. Sébillot (Eds.), *Actes des 6<sup>es</sup> Journées d'analyse des données textuelles*. St Malo, 2002. Inria, 403-412.
- Lebart L. & Salem A. (1994). *Statistique textuelle*. Paris, Dunod.
- Mange J., Marchand P. & Salem A. (2006). Débats sur la toile. In *Actes des 8<sup>es</sup> Journées Internationales d'Analyse des Données Textuelles*. Besançon, 19-21 avril, Presses Universitaires de Franche-Comté, coll. Les cahiers de la MSH Ledoux, série « archive, bases, corpus », 667-676.
- Marchand P. (2007). *Le grand oral - Les discours de politique générale de la Ve République*. De Boeck / INA.
- Malrieu D. & Rastier F. (2001). Genres et variations morphosyntaxiques. *Traitement Automatique des Langues*, 42(2), 548-577.
- Mayaffre D. (2004). *Paroles de Président*. Honoré Champion.

---

<sup>7</sup> On emploie ici à dessein le terme de « genre ». Pour Denise Malrieu et François Rastier (2001), les variations morpho-syntaxiques selon les genres sont notables et l'étude des caractéristiques de genre revient à une étude des normes linguistiques. L'homogénéité générique des *déclarations de politique générale* n'était guère facile à démontrer tant qu'on ne disposait pas de texte qui s'en écarte significativement. On observait bien des évolutions chronologiques et des marquages contextuels, mais cela ne permettait pas d'invoquer rigoureusement un genre. La déclaration d'A. Juppé – *sur son plan de réforme de la protection sociale* - nous avait déjà orienté vers l'existence d'un genre, tant elle se démarquait significativement de toutes les autres. Mais c'est bien le discours de N. Sarkozy qui permet aujourd'hui d'être plus affirmatif.