



HAL
open science

Des mots pour des voix : 132 discours pour devenir président de la République française

Dominique Labbé, Denis Monière

► **To cite this version:**

Dominique Labbé, Denis Monière. Des mots pour des voix : 132 discours pour devenir président de la République française. *Revue Française de Science Politique*, 2008, 58 (3), pp.433-455. halshs-00349256

HAL Id: halshs-00349256

<https://shs.hal.science/halshs-00349256>

Submitted on 6 Mar 2013

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.

Des mots pour des voix

132 discours pour devenir président de la République française

Dominique Labbé (PACTE – CNRS - Institut d'Études Politiques de Grenoble)

Denis Monière (Université de Montréal)

Résumé :

Analyse lexicométrique des discours de campagne des 4 principaux candidats à l'élection présidentielle de 2007 (F. Bayrou, J.-M. Le Pen, S. Royal et N. Sarkozy). Le vocabulaire, les thèmes, les phrases permettent de caractériser les styles et les stratégies de communication. Les deux finalistes sont particulièrement étudiés. La campagne de S. Royal a été dominée par trois thèmes - son "pacte présidentiel", la "valeur travail" et les "services publics" - et par ses attaques contre N. Sarkozy alors que ce dernier a donné la priorité au "pouvoir d'achat", à l'histoire de France et à l'identité nationale et qu'il a choisi de nommer le moins possible ses adversaires jusqu'à la campagne pour le second tour. Pour S. Royal, la politique est une obligation morale et de compétence alors que N. Sarkozy met l'accent sur la conviction et la volonté. Enfin, S. Royal construit des phrases plus longues et plus complexes mais avec un vocabulaire peu spécialisé contrairement à N. Sarkozy dont les phrases sont très courtes, répétées mais avec un vocabulaire adapté au thème traité.

Abstract :

Lexicon analysis of the speeches of the four major candidates who were running for the French presidency in 2007 (F. Bayrou, J.-M. Le Pen, S. Royal et N. Sarkozy). Their vocabularies, themes and sentences show the characteristics of their personal style and communication strategies. The two finalists are more specifically studied. The three most important themes in the S. Royal campaign were: her "presidential program", the "work status", the preservation of the "public services" and she attacked a lot N. Sarkozy whereas this one gave priority to the "purchasing power", history of France, "national identity" and he chose not to pronounce his opponents' names until the campaign for the second ballot. In S. Royal's mind, politics means "moral responsibility" and "ability to rule" whereas in N. Sarkozy's mind it means "belief" and "willingness". At least, S. Royal sentences are much more longer and more complicated but their vocabulary is not specially dedicated to the theme of these sentences. On the contrary, N. Sarkozy sentences are very short and the vocabulary used in them is specific to the subject he is talking about.

Mots clefs :

France – Campagne électorale – Bayrou – Le Pen – Royal – Sarkozy – Lexicométrie - Statistique lexicale.

Manuscrit de l'article paru dans la *Revue française de science politique*, 2008-3, vol 58, p. 433-455.

Toute citation doit être faite à partir du texte de la revue disponible sur CAIRN.

Jusqu'à récemment, l'analyse des discours de campagne était délaissée par les politologues. Les études électorales - dominées par un mélange des paradigmes dits de "Columbia" et de "Chicago" - s'intéressaient peu à ces discours électoraux parce qu'ils étaient réputés avoir peu d'incidence sur le choix des électeurs¹ ; celui-ci étant considéré comme relativement stable d'une élection à l'autre et comme conditionné par des variables démographiques, sociales, culturelles (sexe, âge, scolarité, pratique religieuse etc.) et surtout politiques (les alignements partisans)².

À ces raisons théoriques s'ajoutaient des obstacles pratiques. L'accès aux discours était difficile. Le chercheur ne pouvait pas suivre matériellement toutes les équipes de campagne afin de recueillir les messages qui avaient été émis. L'enregistrement et la transcription des verbatim étaient onéreux. Enfin ces contraintes étaient amplifiées par des préjugés négatifs à l'endroit des politiques. La thèse la plus répandue soutenait qu'il disaient tous la même chose : c'était bonnet blanc, blanc bonnet. Pourquoi s'intéresser à un tissu de banalités, de propos vagues?

Mais avec le déclin des identités partisanses et l'apparition de l'électeur versatile qui change son vote d'une élection à l'autre et qui attend la fin de la campagne pour prendre une décision, l'étude des discours devient incontournable puisque ce qui se passe et se dit durant la campagne peut avoir un effet significatif sur le résultat du scrutin. Dans certaines campagnes électorales au Canada et au Québec, les intentions de vote ont beaucoup varié entre le début et la fin de la campagne électorale (1984, 1988, 2003, 2007). De plus, les électeurs retardent de plus en plus leur décision de vote attendant d'avoir en main tous les programmes avant de faire un choix.³ Dans ce contexte, les discours électoraux sont considérés comme essentiels au processus démocratique parce qu'ils fournissent aux électeurs le moyen de comparer les offres de politiques et

¹ Pour une synthèse en français de la littérature concernant les analyses électorales : N. Mayer et P. Perrineau, *Les comportements politiques*, Paris, A. Colin, 1992 et P. Bréchon, *Comportements et attitudes politiques*, Grenoble, Presses universitaires de Grenoble, 2006.

² Pour une synthèse en français de la littérature concernant les alignements partisans : P. Martin, *Comprendre les évolutions électorales*, Paris, Presses de la fondation des sciences politiques, 2000.

³ Voir F. Fletcher, « The Media and the 1984 Landslide » dans H. Penniman, *Canada at the Polls*, Duke University Press, 1987, p. 161; A. Blais et al, « L'élection québécoise : un bilan des sondages », *Revue canadienne de science politique*, juin 1986, p. 336 ; D. Monière et A. Blais, « la campagne électorale de 1989 », dans *L'année politique au Québec, 1988-1989*, Montréal, Québec Amérique, 1989, p. 132. Sur le phénomène des électeurs tardifs, voir Ian McAllister, « The New Politics of Late Deciding Voters », dans D.M. Farrell et R. Smitt-Beck *Do Political Campaigns Matter?*, New York, Routledge, 2002, p. 22-41.

d'évaluer les performances gouvernementales⁴. Inspirées par ce nouveau paradigme, les recherches se sont concentrées principalement sur les effets de la communication et ont tenté d'expliquer le rôle des médias, tout particulièrement celui de la télévision, en négligeant largement le discours⁵.

Les travaux portant directement sur les discours électoraux sont encore très peu fréquents et ont surtout privilégié certains types de discours plus facilement accessibles - comme les débats télévisés ou les messages publicitaires - mais ces corpus ne sont pas exhaustifs et ne couvrent pas l'ensemble des déclarations des candidats faites jour après jour, celles-ci n'étant presque jamais transcrites et répertoriées.

Aujourd'hui, l'accès à des corpus significatifs n'est plus un obstacle, car les campagnes se déroulent aussi sur Internet que les partis et les candidats utilisent pour diffuser les interventions qu'ils jugent importantes. On peut donc maintenant rassembler chronologiquement tous les discours importants prononcés par les principaux candidats et suivre ainsi le déroulement de la compétition électorale.

C'est ce que nous avons fait avec les 4 candidats arrivés en tête de l'élection présidentielle française de 2007. Nous avons appliqué à leurs discours les méthodes déjà utilisées sur d'autres corpus politiques⁶.

Après avoir décrit ce corpus et évoqué les méthodes d'analyse, un certain nombre de résultats sont présentés qui suggèrent certaines pistes particulièrement intéressantes pour l'étude de la communication politique permettant de circonscrire les thématiques de la campagne, les stratégies déployées par les candidats et les styles de communication privilégiés.

⁴ Voir Darrell West, *Air Wars*, Washington, Congressional Quarterly, 1997, p.185 ; Denis Monière, *Le discours électoral : les politiciens sont-ils fiables?*, Montréal, Québec-Amérique, 1988.

⁵ En France, la science politique s'est intéressée avec beaucoup de retard à cette question. C'est en 1999 que, pour la première fois, la question figure à l'ordre du jour du congrès de l'Association française de science politique. A cette occasion, aucune des interventions ne portait sur l'analyse du discours ni sur les campagnes électorales (J. Gerstlé (éd.), *Les effets d'information en politique*, Paris, L'Harmattan, 2001).

⁶ Par exemple : D. Labbé, *Le vocabulaire de F. Mitterrand*, Paris, Presses de la FNSP, 1990 ; D. Labbé et D. Monière, *Le discours gouvernemental*, Paris, H. Champion, 2003 et *Les mots qui nous gouvernent*, à paraître en 2008.

1. Corpus et traitements préalables

En analyse du discours, tout repose sur la qualité du corpus de textes rassemblés et sur celle des traitements préalables de ces textes.

Traitements préalables

Ces textes font l'objet de certains traitements préalables⁷.

En premier lieu, l'orthographe a été soigneusement corrigée et les graphies multiples ont été standardisées. Par exemple, "événement" et "évènement" ne sont qu'un même mot écrit de deux manières différents de même que "puis" et "peux", etc.). Cela concerne particulièrement les sigles, les abréviations, les noms propres, les chiffres et les dates dont la transcription est d'une infinie variété... Ces tâches sont partiellement effectuées par des automates, mais les interventions manuelles sont nombreuses et suivent des règles précises.

En second lieu, des balises indiquent les sources du texte puis délimitent les séquences (début et fin des propos des orateurs, interruptions, questions et réponses pour les interviews). Ce travail est indispensable pour ne pas confondre les propos du candidat – qui font seuls l'objet des analyses présentées ici - avec ceux des autres participants aux réunions publiques qu'il est pourtant essentiel de conserver pour la bonne compréhension du texte.

Enfin, la lemmatisation attache à chacun des mots du texte une étiquette contenant la graphie standard et l'entrée sous laquelle le mot peut être retrouvé dans un dictionnaire. Par exemple, "puis" peut recevoir deux étiquettes différentes (selon le contexte) : "puis, adverbe" ou "pouvoir, verbe". Ce travail est indispensable pour trois raisons essentielles. En premier lieu, dans tout texte en français, plus du tiers des mots sont "homographes" : une seule graphie mais plusieurs entrées possible dans le dictionnaire (comme par exemple, la forme "pouvoir" qui intéresse tant les politologues : substantif masculin ou verbe ?). La lemmatisation met de l'ordre dans ce chaos et permet de consulter une base de données textuelles comme on le fait avec un dictionnaire. Deuxièmement, la langue française étant fortement flexionnelle, la lemmatisation réduit

⁷ Pour le détail de ces traitements : D. Labbé, *Normes de saisie et de dépouillement des textes politiques*, Grenoble, Cahiers du CERAT, 1990.

considérablement le nombre des individus rares – sur lesquels le statisticien n'a rien à dire sinon qu'ils sont rares. Enfin, la lemmatisation débouche sur des analyses intéressantes, comme on le verra plus loin.

Le corpus

Les discours ont été téléchargés sur les sites des candidats, soit : *Désir d'avenir* pour S. Royal; *Ensemble tout devient possible* et le site de l'UMP pour N. Sarkozy; le site du Front national pour J.-M. Le Pen et enfin le site *La France de toutes nos forces* pour F. Bayrou. Comme les sites de campagnes sont devenus un mode de communication électorale valorisé par les candidats, ces corpus peuvent être considérés comme représentatifs de leur communication électorale même s'il n'existe pas de preuve de leur exhaustivité (tableau 1). La cueillette a débuté avec les déclarations de candidature : S. Royal, le 17 novembre 2006, J.-M. Le Pen, le 12 novembre, F. Bayrou le 2 décembre et N. Sarkozy, le 18 décembre. Pour analyser un corpus homogène et ne pas mélanger les genres, nous n'avons retenu que les discours prononcés devant des auditoires et exclu les conférences de presse, les entretiens et les lettres car certains candidats ne mettaient pas systématiquement en ligne ce genre de messages.

Tableau 1 : Corpus des discours des candidats à l'élection présidentielle

	Nombre de discours	Longueur en mots	Formes standardisées différentes	Vocables différents
Bayrou	28	227 446	11 200	6 396
Le Pen	21	87 114	10 961	7 298
Royal (1er tour)	31	146 687	10 206	6 103
Royal (2e tour)	8	48 269	5 093	3 255
Royal total	39	194 956	11 632	6 738
Sarkozy (1er tour)	36	231 024	12 289	7 215
Sarkozy (2 ^e tour)	8	68 843	6 414	4 059
Sarkozy total	44	299 867	13 817	7 879
Total Premier tour	116	692271	24 082	13 156
Total Deuxième tour	16	117112	8 793	5 251
Total présidentielle	132	809383	25 419	13 658

Le tableau 1 montre que, tant par le nombre de discours que par le nombre de mots, les principaux candidats ne se sont pas engagés avec une égale intensité dans l'effort de persuasion. N. Sarkozy a été le plus bavard et J.-M. Le Pen le moins intervenant. Ce

dernier n'a prononcé en moyenne qu'un discours à tous les 8 jours – en s'en tenant à ce qui est indiqué par son site Internet - alors que F. Bayrou en faisait un tous les 5 jours, S. Royal tous les 4,4 jours et N. Sarkozy tous les 3,7 jours. Cette relative inertie du candidat de l'extrême-droite n'est peut-être pas étrangère à sa contre-performance en vertu du principe qui veut que, moins on est actif, moins on est visible et moins on est en mesure d'influencer le choix des électeurs et plus on laisse ses adversaires accaparer l'attention et imposer leurs thématiques.

La fréquence des discours n'est pas le seul facteur explicatif, il faut aussi prendre en considération les contenus valorisés par chaque candidat, et pour cela s'intéresser aux contenus des messages à l'aide d'instruments qui ne fasse pas intervenir la subjectivité de l'observateur.

2. Le choix des mots

La statistique appliquée au discours repose sur le postulat selon lequel il est possible d'inférer des significations à partir des fréquences de vocabulaire, et d'en extraire des informations sur celui qui parle et sur le contenu de son message. Si on reprend la métaphore de l'urne pour décrire l'ensemble du vocabulaire d'une langue, on peut supposer que le fait qu'un groupe de mots particuliers sortent plus souvent que les autres peut ne pas être fortuit et que certains écarts – par rapport aux autres – peuvent être réellement significatifs.

En premier lieu, les vocables sont classés par ordre décroissant de fréquence (index "hiérarchiques") afin de retrouver les mots "favoris" des locuteurs, par exemple, les noms et les adjectifs les plus utilisés (tableau 2 et 3).

Les mots favoris des candidats

Tableau 2. Les dix premiers substantifs classés par ordre décroissant de fréquence

Bayrou	Le Pen	Royal 1 ^{er} tour	Sarkozy 1 ^{er} tour	Royal 2 ^e tour	Sarkozy 2 ^e tour
pays	France	France	France	France	France
France	an	pays	politique	entreprise	Français
entreprise	pays	Français	Français	Français	république
monde	Français	travail	travail	pays	politique
Français	million	entreprise	enfant	femme	pays
chose	Etat	emploi	homme	travail	travail
société	peuple	jeune	république	salarié	enfant
Europe	politique	valeur	vie	emploi	droit
république	monsieur	Europe	pays	valeur	peuple
an	travail	droit	valeur	Europe	madame

Ces listes sont toujours très instructives. Par exemple, *travail* est un mot privilégié par S. Royal comme par N. Sarkozy. Seul F. Bayrou parle du *monde* et il partage avec S. Royal un intérêt fort pour l'*Europe* alors que les deux autres ont manifestement été plus discrets sur ce point. Autre surprise : les *entreprises* ne figurent pas parmi les mots préférés de Sarkozy, contrairement à F. Bayrou et S. Royal. Seul J.-M. Le Pen place le *peuple* en haut de son vocabulaire. N. Sarkozy et F. Bayrou font un emploi beaucoup plus fort l'adjectif *grand* par rapport à Royal. Bayrou et J.-M. Le Pen donnent à l'adjectif *français* une place sans commune mesure avec celle que leur accordent S. Royal et N. Sarkozy.

D'autres informations sont moins évidentes. Ainsi l'apparition de *madame* dans les propos de N. Sarkozy entre les 2 tours : il s'agit essentiellement de *Madame Royal*, ce qui indique une nette personnalisation de sa campagne, alors qu'avant le premier tour, il avait pris soin de ne pas nommer ses adversaires (comme on le verra plus loin). Ou encore l'utilisation abondante de *an* par J.-M. Le Pen qui révèle une dimension caractéristique du personnage qui sera analysée dans la troisième section à propos du tableau 7.

Beaucoup d'autres informations peuvent être tirés des simples index tant ils sont riches. Mais ces commentaires laissent en suspens deux questions : quels écarts sont réellement significatifs du point de vue statistique et quel sens donnent les locuteurs à ces mots très fréquents mais aussi assez vagues.

La réponse à la première question est donné par un calcul qui recherche les suremplois (et les sous-emplois) statistiquement significatifs de chacun des locuteurs comparés à tous les autres⁸.

Les mots caractéristiques des candidats

Le tableau ci-dessous indique quels ont été les substantifs les plus nettement suremployés par chacun des 4 candidats lors du 1^{er} tour (les suremplois sont caractéristiques du locuteur avec moins de 1 chance sur 10 000 de se tromper). Faute de place, les sous-emplois ne sont pas présentés ici.

Tableau 3. Les substantifs les plus caractéristiques (classement par indice décroissant)

<p>Bayrou : pays, entreprise, monde, chose, société, peuple, question, famille, besoin, citoyen, manière, situation, idée, président, éducation, nombre, élection, retraite, raison, problème, chance, gouvernement, moment, charge, décision, sujet, université, ami, mois, effet, parti, réforme, dette, engagement, exemple, réalité, planète, texte, fond, guerre</p> <p>Le Pen : an, million, peuple, monsieur, nation, immigration, année, madame, liberté, insécurité, effet, loi, euro, compatriote, système, condition, candidat, frontière, milliard, siècle, mesure, paysan, front, ministre, niveau, patrie, immigré, classe, population, terre, impôt, résultat, agriculture, chiffre, produit, nature, environnement, armée, préférence, bien, étranger, équilibre</p> <p>Royal : entreprise, emploi, jeune, valeur, femme, pacte, besoin, salarié, avenir, recherche, place, énergie, présidente, heure, logement, sécurité, formation, région, droite, effort, façon, combat, territoire, confiance, égalité, défense, solidarité, parole, travail, puissance, action, mouvement, aide, développement, débat, salaire, compte, inégalité, valeur, choix, précarité</p> <p>Sarkozy : politique, travail, enfant, homme, valeur, république, vie, droit, école, culture, identité, liberté, gauche, histoire, président, moyen, idée, force, autorité, crise, coeur, condition, respect, devoir, sentiment, principe, volonté, travailleur, pensée, rêve, parent, salaire, cause, morale, civilisation, jeunesse, création, courage, terre, justice, sport, peur</p>
--

Ces brèves listes indiquent assez clairement les axes de campagne privilégiés par chacun. Pour résumer, F. Bayrou a fait essentiellement campagne sur l'*entreprise* et sur les questions de *société* et du *monde* ; J.-M. Le Pen a repris ses thèmes classiques

⁸ Ce calcul est une adaptation des "spécificités du vocabulaire" (Lafon P., *Dépouillements et statistiques en lexicométrie*, Genève-Paris: Slatkine-Champion, 1984). Pour la discussion statistique : Labbé C. et Labbé D., *Que mesure la spécificité du vocabulaire ?* Grenoble, CERAT, 1994. Repris dans : *Lexicometrica*, 3, 2001. Pour la prise en compte de différences de densités entre catégories grammaticales : Monière D., Labbé C; et Labbé D., "Les particularités d'un discours politique : les gouvernements minoritaires de Pierre Trudeau et de Paul Martin au Canada", *Corpus*, 4, 2005, p 79-104.

(*nation, immigration, insécurité...*) avec un intérêt nouveau pour l'*agriculture* et pour les chiffres ; S. Royal partage avec F. Bayrou la même insistance sur les *entreprises* mais il s'agit ici d'*emploi des jeunes, des femmes, des salaires, du logement*, etc. Quant à N. Sarkozy il a parlé *politique, valeur, république, histoire*, mais aussi *identité, culture, civilisation* ; il a aussi beaucoup traité des *enfants* et des *parents* (*famille* est un peu plus loin dans la liste). Il est manifestement celui qui a été le plus divers.

N. Sarkozy et S. Royal privilégient tous deux les mots *travail* et *valeur* mais leur donnent-ils le même sens ? C'est la deuxième question posée à propos des mots les plus fréquents du discours politique.

Le sens des mots

Les corpus "lemmatisés" permettent de répondre à cette question⁹. L'algorithme relève les mots qui sont les plus étroitement associés avec le vocable dont on recherche le sens – comme les définitions du dictionnaire – et il repère les phrases les plus caractéristiques du mot étudié, comme les dictionnaires illustrent leurs définitions par quelques citations. Le tableau 4 présente les mots les plus fortement associés à *travail* chez S. Royal et N. Sarkozy (faute de place, la liste complète n'est pas reproduite).

Le tableau suggère que N. Sarkozy et S. Royal partagent pas mal de choses quand ils parlent du *travail*. Par exemple, cinq des dix verbes les plus étroitement associés à *travail* leur sont communs (*faire, payer, taxer, travailler, vivre*) et *valeur* est le substantif le plus significatif chez les deux.

⁹ Sur le calcul du sens des mots : Hubert P. & Labbé D., "La structure du vocabulaire du général de Gaulle", In Bolasco S. et Al. *IIIe Giornate internazionali di analisi statistica dei dati testuali*. Rome, CISU, II, p 165-176. ; Labbé C. & Labbé D., "How to measure the meanings of words ? Amour in Corneille's work", *Langage Resources Evaluation*, 2005, 39, p 335-351. Pour la comparaison du sens d'un même mot chez deux auteurs différents : "La France chez de Gaulle et Mitterrand", in Fiala P., Lafon P. (dir), *Des mots en liberté. Mélanges Maurice Tournier*, Fontenay-aux-Roses, ENS-Editions, 1998, p 183-193. Le calcul porte sur les associations positives (les mots s'attirent mutuellement) et négatives (les mots se repoussent ou s'excluent mutuellement). Faute de place, on se limite aux premières (les attirances mutuelles).

Tableau 4. Vocables les plus fortement associés à *travail* chez S. Royal et N. Sarkozy

<p>Royal : Noms propres : CPE Verbes : travailler, vivre, retrouver, payer, arriver, taxer, saper, revaloriser, généraliser, tuer, faire Substantifs : valeur, salarié, salaire, heure, famille, sécurité, an, précarité, million, condition, contrat Adjectifs : social, professionnel, supplémentaire, bas, payé, précaire, décent, garanti, petit, scolaire</p> <p>Sarkozy : Noms propres : Jaurès Verbes : travailler, faire, vivre, taxer, souhaiter, permettre, créer, payer, transmettre, cesser Substantifs : valeur, crise, revenu, famille, emploi, travailleur, mérite, fruit, effort, temps, condition, chômage Adjectifs : social, pauvre, particulier, fiscal, ouvrier, remarquable, régulier, désespérant, moyen, indéterminé, rémunéré</p>
--

Cependant, la courte liste des principaux substantifs et adjectifs, associés à *travail*, montre aussi des oppositions très nettes. N. Sarkozy privilégie les *travailleurs*, le *mérite*, les *fruits du travail* et de *l'effort*, le *travail ouvrier*... Il s'agit de *moins taxer* les revenus du travail. En revanche, S. Royal parle des *salariés*, des *heures supplémentaires*, des *conditions* et du *contrat de travail*, et promet de *revaloriser* les *bas* ou les *petits salaires*. Les phrases les plus caractéristiques illustrent clairement cette opposition entre les deux candidats (tableau ci-dessous).

Tableau 5. Phrases les plus caractéristiques du vocable *travail* chez S. Royal et N. Sarkozy

<p>Royal : "Ce sont les bas salaires qui tuent la valeur travail" (1^{er} février 2007 : tous les mots de cette phrase sont caractéristiques de l'univers lexical de <i>travail</i> chez Royal). "Ce sont les salariés "jetés" qui tuent la valeur travail" (1^{er} février 2007). "C'est le chômage qui tue la valeur travail, c'est la précarité qui brise la valeur travail, c'est le CPE qui a contrecarré la valeur travail" (23 mars 2007). " Ce sont les exonérations fiscales aux plus riches qui tuent la valeur travail" (1^{er} février 2007). "La France présidente, c'est celle qui réhabilite la valeur du travail" (20 avril 2007). "C'est parce que les salariés seront bien payés que la valeur travail sera respectée, c'est parce que les salariés seront respectés que la valeur travail sera garantie, c'est parce que l'égalité salariale entre les hommes et les femmes sera réalisée que les salariés seront motivés" (4 mai 2007). "Ce sont les inégalités économiques et sociales qui sapent et qui gangrènent la valeur travail" (1^{er} février 2007). "La valeur travail, c'est d'abord payer le travail à sa valeur" (2007).</p>

Sarkozy :

"La crise de la valeur travail est au coeur de la crise morale" (formule répétée 5 fois : 25 et 36 janvier, 1^{er} et 20 mars, 11 avril 2007).

"Je souhaite une école qui place au coeur de ses valeurs le travail, l'effort, le mérite." (formule répétée 4 fois : 25 et 26 janvier, 2 février, 13 mars 2007).

"Avec la crise de la valeur travail, c'est l'espérance qui disparaît." (formule répétée 7 fois : 14, 25 et 26 janvier, 7 février, 1^{er} et 20 mars, 11 avril 2007).

Cela met en lumière une caractéristique particulière du discours électoral : répéter sans se lasser les mêmes slogans. De ce point de vue, N. Sarkozy surpasse de loin S. Royal dont les collaborateurs n'ont pas semblé avoir le même talent pour trouver les petites phrases "coup de poing" que ses auditoires auraient pu retenir et répéter autour d'eux.

Cet échantillon des phrases les plus représentatives montre également comment la campagne électorale est une sorte de dialogue à distance. Elle permet surtout de relever l'un des échanges les plus significatifs de la campagne électorale. En janvier, N. Sarkozy, au nom de J. Jaurès et de L. Blum, accuse la gauche d'avoir oublié la *valeur travail*. S. Royal répond "captation d'héritage" et accuse la droite d'avoir *tué* la *valeur travail* :

"Ce sont ceux qui ont érigé la précarité en norme de référence qui ont détruit la *valeur travail* et qui n'ont plus aujourd'hui aucune légitimité politique pour la défendre. C'est nous qui la défendrons, parce que c'est le respect du salarié et du travailleur, la sécurité dans le contrat de travail, le salaire décent et durable qui constituent les fondamentaux de la *valeur travail* et non pas le contraire" (1^{er} février).

N. Sarkozy réplique et finit par trouver le thème qu'il reprendra dans une demi-douzaine de ses derniers discours de la campagne du 1^{er} tour :

"Voyez-la, écoutez-la cette gauche qui depuis mai 68 a cessé de parler aux travailleurs, de se sentir concernée par le sort des travailleurs, d'aimer les travailleurs parce qu'elle rejette la *valeur travail*, parce que la valeur travail ne fait plus partie de ses valeurs, parce que son idéologie à elle ce n'est pas l'idéologie de Jaurès, ce n'est pas l'idéologie de Blum qui respectaient le travail, qui aimaient les travailleurs, son idéologie à elle c'est l'idéologie du partage du travail, des 35 heures, de l'assistanat" (29 avril 2007).

C'est donc en étudiant la manière dont un orateur associe les mots entre eux que l'on peut découvrir les thèmes qu'il développe.

Des mots aux thèmes

Pour décrypter les principaux thèmes des candidats, nous comparons les syntagmes les plus fréquemment répétés par les deux finalistes¹⁰. Cette dernière analyse est limitée aux deux finalistes et à leurs syntagmes composés de deux vocables dont le premier est un substantif. Comme précédemment, une fois les principaux syntagmes identifiés, l'algorithme recherche les phrases les plus caractéristiques du thème considéré.

Tableau 6. Principaux syntagmes répétés chez S. Royal et N. Sarkozy (fréquence absolue)

Royal :

pacte présidentiel 98, France président 92, élection présidentiel 56, valeur travail 50, service public 49, puissance public 43, formation professionnel 42, homme (et) femme 38, peuple français 34, pouvoir (d') achat 32, désir (d') avenir 27, sécurité social 27, bas salaire 26, logement social 26, dialogue social 25, nouveau règle 25, enseignement supérieur 24, énergie renouvelable 22, progrès social 22, règle (du) jeu 22, France nouveau 21, matière grise 21, ordre juste 21, république nouveau 20, efficacité économique 19, pacte social 19, pays pauvre 18, volonté politique 19, petit retraite 18, changement profond 17, valeur ajouté 17.

Sarkozy :

pouvoir (d') achat 71, président république 71, pensée unique 61, président (pour la) France 56, histoire (de) France 47, identité national 42, égalité (des) chance 39, homme politique 37, promotion social 37, droit (de l') homme 36, crise moral 34, idée (de l') homme 33, plein emploi 33, politique (et la) vie 33, intérêt général 30, valeur travail 30, développement durable 29, homme (et) femme 29, service public 27, élection présidentiel 26, droite républicain 25, fonction présidentiel 24, fruit (du) travail 24, démocratie irréprochable 22, culture commun 21, parti socialiste 21, capitalisme financier 20, crise (de la) valeur 20, génération futur 20, heure complémentaire 20, revenu (du) travail 20

La lecture de cette liste ne surprendra sans doute pas, mais elle montre que chaque candidat avait choisi ses principaux thèmes en fonction de sa famille idéologique. Outre *l'élection présidentielle*, il n'y a que 3 recoupements entre les deux : (les) *hommes* et (les) *femmes* (S. Royal elle-même place les hommes avant les femmes), *valeur travail* et *service public*) et quelques chassés-croisés caractéristiques : c'est N. Sarkozy qui parle beaucoup du *parti socialiste* - alors que S. Royal évite plutôt ce thème (elle l'utilise 5 fois avant le premier tour et 2 fois dans l'entre-deux-tours) – et qui dénonce le *capitalisme financier*.

¹⁰ Sur cette notion : A. Pibarot & D. Labbé, "Les syntagmes répétés dans l'analyse des commentaires libres", in Mellet S. (ed), *4e Journées d'analyse des données textuelles*, Nice, 1998, p 507-516. Le tableau 6 comporte les vocables (ou entrées de dictionnaire) et non les mots. Ainsi par exemple, "service public" et "services publics" forment un même syntagme, le pluriel n'étant qu'une flexion d'un même vocable.

Les thèmes sociaux ont donc dominé la campagne présidentielle de S. Royal. Plutôt que de parler de "programme" ou de "projet", elle préfère *pacte*. Ce mot apparaît principalement dans deux syntagmes : *pacte présidentiel* et *pacte social*. Voici les deux phrases, les plus caractéristiques de *pacte présidentiel*, qui résument bien les principaux thèmes de la campagne de S. Royal, tels qu'ils apparaissent grâce aux mots caractéristiques et aux syntagmes répétés :

"le *pacte présidentiel* est un projet cohérent qui s'appuie non seulement sur des valeurs, je vous l'ai dit, sur des règles du jeu neuves et claires où chacun saura ce qu'il a à faire dans cette France qui veut à nouveau bouger et se redresser, mais aussi des décisions rapides, immédiates, qui répondront aux urgences, et je l'ai dit et je le répète, aux urgences de la lutte contre le chômage, aux urgences de la défense du pouvoir d'achat, aux urgences des moyens à rendre à l'école et à la formation, à l'urgence écologique aussi ; à la nécessité aussi de faire revenir la France à la table de l'Europe" (30 mars 2007).
"Lorsque l'on voit les sept piliers de mon *pacte présidentiel*, on se rend compte que la questions des femmes y est majeure, puisque lorsque je parle de la lutte contre le chômage, je sais que les premières victimes du chômage sont les femmes ; lorsque je parle de la lutte contre la précarité, je sais que 80% des travailleurs précaires sont des femmes ; lorsque je parle de la lutte contre toutes les formes de violence, je sais que les femmes sont les premières victimes de toutes les formes de violence ; lorsque je parle du droit d'accès à la formation professionnelle, je sais que ce sont les femmes qui sont le plus écartées de la formation professionnelle ; lorsque je parle de droit à la santé, de revalorisation des bas salaires ou des petites retraites, je sais que ce sont d'abord les femmes qui sont frappées par les bas salaires et par les petites retraites" (7 mars 2007).

Le premier thème de campagne de N. Sarkozy a donc été le *pouvoir d'achat*. Cela est confirmé par certains syntagmes verbaux fréquents (non reproduits ici) : *vouloir travailler* (plus), *falloir travailler* (plus), *travailler (plus pour) gagner (plus)*... Comme on l'a montré plus haut, à propos du *travail*, la dénonciation de la *pensée unique* – en l'occurrence, il s'agit de celle qu'il prête à S. Royal et au Parti socialiste - a été l'autre thème principal de la campagne de N. Sarkozy, suivie par la revendication de *l'histoire de (la) France* et la défense de *l'identité nationale*. Voici, pour deux de ces thèmes, la phrase la plus caractéristique :

Je veux en finir avec la *pensée unique* qui nous a laissé dans la situation où nous sommes, qui s'est trompée sur tout, qui a échoué sur tout, qui a creusé les déficits, qui a augmenté le chômage, qui a fait exploser la pauvreté, l'exclusion, la précarité, qui a raté l'intégration, qui a laissé se développer les discriminations, qui installé le communautarisme, qui a abaissé l'autorité de l'Etat, qui a affaibli l'identité nationale, qui a fabriqué l'échec à l'école, qui a dévalorisé les diplômes et le travail, qui a favorisé l'assistanat" (20 mars 2007).

"Je continuerai donc de parler de *l'identité nationale* parce que c'est un sujet qui concerne tous les Français, parce que je ne veux pas laisser le monopole de la nation à l'extrême-droite, parce que je me fais une haute idée de la France, parce que je n'ai pas honte de mon pays, parce que je n'ai pas honte de ses valeurs, de sa culture, parce que je n'ai pas à m'excuser d'être Français, parce qu'au contraire j'en suis fier, parce que j'aime la France pour ce qu'elle m'a donné en partage, pour ce qu'elle m'a permis de devenir" (13 mars 2007)

Non seulement les deux thèmes se recoupent étroitement, mais ils sont tous les deux clairement dictés par le souci de reconquérir une partie de l'électorat de la droite républicaine tentée par le vote d'extrême droite.

3. Tactiques et stratégies de communication

Ce qui précède aura fait sentir combien certains indicateurs sont révélateurs des stratégies poursuivies. Quelques indices simples permettent de caractériser ces stratégies.

L'ancrage spatial, temporel, économique

Lorsqu'il s'adresse à un auditoire, tout orateur doit choisir. Il peut situer son propos, le rattacher à des lieux, des dates, des phénomènes connus de son auditoire. Ses paroles y gagneront un ancrage dans le temps ou dans l'espace mais leur portée, ou leur généralité, en seront réduites. Par conséquent, la densité des toponymes, des patronymes, des dates et des chiffres permet d'évaluer l'ancrage spatial, temporel, social, économique des discours candidats (tableau 7).

J.-M. Le Pen cite un très grand nombre de lieux, de noms de personnes, de dates, secondairement de chiffres. Ainsi s'explique la très forte densité du substantif "an" dans le tableau 2 Du coup, son discours donne le sentiment d'entendre un cours d'histoire. Cet ancrage temporel et spatial est bien illustré par sa phrase la plus caractéristique :

"Sur 20 ans de Parlement européen, j'ai d'ailleurs tellement fait, avec mes amis députés, pour la défense des paysans de France, mais aussi d'Europe et du monde, face aux idéologues du libre-échange planétaire qui sèment la jachère en Beauce, l'arrachage des vignes en Languedoc-Roussillon ou l'élimination des éleveurs de l'Allier, tout en éliminant aussi les campesinos du Mexique, des producteurs de riz du Sénégal ou les petits céréaliers du Rio grande do Sul, que j'ai quelque légitimité à m'adresser à nos 540000 exploitants, au 1.1 million d'actifs permanents et surtout aux 2.3 millions de nos retraités agricoles à qui l'on ose verser, pour seul paiement de leur vie au service de

l'alimentation des Français, que 350 euros de retraite par mois, et même encore moins pour la veuve du paysan" (6 mars 2007).

Tableau 7 Poids des toponymes, patronymes et nombres (en ‰ mots)

	Mots à majuscule	Nombres	Somme d'argent
Sarkozy (1er tour)	19,5	9,1	2,9
Royal (1er tour)	18,7	11,5	3,0
Royal (2e tour)	19,2	11,6	3,1
Bayrou	16,8	15,4	3,2
Sarkozy (2 ^e tour)	22,3	14,8	3,7
Le Pen	25,7	43,5	6,9

À l'opposé, N. Sarkozy est celui qui, au premier tour, chiffre le moins ses propos. Il avait annoncé qu'il ferait campagne sur les "idées". Il produit effectivement un discours valable pour tous, en tout temps et en tout lieu. En revanche, pour le 2e tour, il cite plus de noms propres (+14%) - en premier lieu : *Madame Royal* – et de chiffres (+64%). Cette avalanche de chiffres vise à montrer, par effet de contraste, que son adversaire ne connaît pas ses dossiers et n'a pas les compétences pour présider la République française.

Pour le premier tour, S. Royal semble avoir adopté une stratégie assez proche de celle de N. Sarkozy. Elle n'en change pas entre les deux tours et, par contraste avec le discours de son adversaire, elle a donc semblé plus "abstraite" et moins en prise sur l'actualité...

Le rapport à l'adversaire.

Pour évaluer et comparer les comportements polémiques des candidats, nous utilisons comme indicateur les références aux patronymes des adversaires classées par rapport à l'ensemble des références à des personnes physiques (tableau 8). Les 4 corpus n'ayant pas la même longueur, les fréquences absolues sont converties en "pour 10 000 mots", ce qui permet de constater que J.-M. Le Pen cite le nom de ses concurrents 4,4 fois plus que ne le fait F. Bayrou (dernière ligne du tableau 8) et 8,4 fois plus que N. Sarkozy.

Tableau 8. Principaux patronymes cités par les 4 candidats
(Classement par ordre décroissant de fréquences exprimées pour 10 000 mots)

Bayrou		Le Pen		Royal		Sarkozy	
Sarkozy	1,5	Sarkozy	6,8	Sarkozy	1,7	Jaurès	2,7
Royal	1,3	Bayrou	3,9	Jaurès	1,5	de Gaulle	1,9
Chirac	1,2	Royal	3,9	Mitterrand	1,3	Ferry	1,5
Mendès- France	0,8	Chirac	3,8	Bayrou	0,9	Royal	1,4
Schuman	0,7	Le Pen	2,9	Blum	0,4	Blum	1,1
Bayrou	0,6	Mitterrand	1,4	Césaire	0,4	Fillon	0,6
de Gaulle	0,6	Giscard	1,3	Delors	0,4	Pompidou	0,6
Jospin	0,6	Hulot	0,8	Zapatero	0,4	Hollande	0,5
Hulot	0,5	Rocard	0,7	Ayrault	0,3	Chirac	0,5
Le Pen	0,5	de Gaulle	0,6	Fillon	0,3	Voltaire	0,5
				(...)		(...)	
				Royal	0,1	Bayrou	0,2
				Le Pen	0,0	Le Pen	0,2
						Sarkozy	0,1
Références aux 3 adversaires	3,3		14,6		2,6		1,7

Plus un candidat mentionne ses adversaires plus son propos est critique; à l'inverse moins un chef est nommé par les autres, plus il est marginalisé dans la dynamique électorale. La principale caractéristique de la dernière campagne électorale saute aux yeux : N. Sarkozy a été au cœur des propos tenus par les trois autres candidats alors que lui-même a préféré les ignorer assez largement et parler d'abord de J. Jaurès, de J. Ferry ou de L. Blum. « Je me sens l'héritier de Jaurès, de Blum etc. » a-t-il lancé le 25 janvier 2007 (thème repris ensuite une trentaine de fois). Il s'est beaucoup référé à C. De Gaulle et à G. Pompidou mais aussi à F. Fillon. Seule "Madame Royal" émerge à la quatrième place mais l'essentiel de ces mentions viennent de l'entre-deux-tours. On notera également, chez N. Sarkozy, l'importance relative des références à F. Hollande qui sont à mettre en relation avec ses attaques contre le Parti socialiste (tableau 9).

S. Royal concentre ses attaques sur N. Sarkozy. Elle se réclame de l'héritage de F. Mitterrand - associé à l'occasion à J. Jaurès et à L. Blum - pour tenter de contrer, sans grand succès, la stratégie de récupération de son adversaire.

F. Bayrou se réclame de l'héritage de P. Mendès-France « ce grand défenseur de la démocratie » qu'il mentionne 19 fois qu'il dit vénérer avec C. de Gaulle « parce qu'ils

n'étaient pas des hommes d'intérêts partisans » (12 mars). Toutefois, si lui et J.-M. Le Pen évoquent souvent Chirac, c'est pour s'en dissocier et critiquer sa présidence.

Les candidats qui sont le plus à l'offensive sont donc ceux qui ont peu de chance de l'emporter alors que les deux présidentiables se montrent plus circonspects quant à la désignation de leurs adversaires. La même remarque est valable pour les partis (tableau 9).

Tableau 9. Références aux partis (sigle ou nom, pour 10 000 mots)

	UDF	FN	PS	UMP	Total (autres que le parti du candidat)
Bayrou	(0,7)	0,0	2,5	2,2	4,4
Le Pen	1,1	(4,8)	1,4	1,5	4,0
Royal	0,1	0,1	(0,3)	0,8	1,0
Sarkozy	0,2*	0,1	0,8	(0,2)	1,0
Total	0,2	0,0	1,0	1,0	

* Chez N. Sarkozy il s'agit 4 fois sur 7 de ses « amis de l'UDF », c'est-à-dire les dissidents anti-Bayrou.

F. Bayrou est celui qui fait le plus référence aux partis. Il s'attaque exclusivement au Parti socialiste et à l'UMP, qui sont ses concurrents les plus immédiats. J.-M. Le Pen est pratiquement au même niveau mais répartit équitablement ses attaques contre les 3 autres. Là encore, c'est un thème de campagne ancien qu'il avait résumé autrefois par la *bande des quatre larrons* (devenus trois du fait de l'effondrement du parti communiste). S. Royal cherche à rester au dessus de la mêlée en parlant peu de son propre parti et en ignorant aussi les partis de ses adversaires. Enfin N. Sarkozy ne ménage pas le Parti socialiste mais ignore, pour ainsi dire, les autres partis (y compris le sien).

Dans l'ensemble on peut dire que les deux présidentiables privilégient une stratégie d'évitement de leurs adversaires, alors que les deux autres candidats, en retard dans les intentions de vote, préfèrent une stratégie de la dénonciation qui vise à débaucher les électeurs qui leur sont les plus proches.

Pouvoir ou vouloir ?

Quelle relation l'orateur entretient-il avec ce qu'il dit, avec son auditoire, le monde qui l'entoure ? Il existe plusieurs indices commodes pour répondre à cette question.

En premier lieu, il s'agit des combinaisons verbe + verbe, dans lesquelles le second verbe est à l'infinitif. On appelle parfois le premier de ces verbes un "pseudo-auxiliaire" parce qu'il joue un rôle semblable à "avoir" ou "être" avec le participe passé. Ces combinaisons signalent une tension et indiquent l'orientation de cette tension. En effet, s'il existe une infinité de verbes modalisateurs ou pseudo-auxiliaires, ceux-ci peuvent être regroupés en quelques catégories (en fonction de l'idée introduite par le premier verbe) : la foi dans une idée (*croire*), l'obligation (*devoir*), la nécessité (*falloir*), la possibilité (*pouvoir*), la connaissance (*savoir*), la volonté (*vouloir*). Cette dernière est la plus présente dans le discours politique français contemporain – quel que soit le locuteur - et trois combinaisons figurent toujours dans le groupe de tête : *vouloir dire*, *vouloir être*, *vouloir faire* (ou tout autre verbe d'action à la place de *faire*). Toutes ces combinaisons ont été relevées chez les deux principaux candidats avant le premier tour et le calcul présenté ci-dessus a été appliqué pour discriminer les pseudo-auxiliaires significatifs. Le tableau 11 présente les résultats de ce calcul pour les deux seuls finalistes.

Tableau 10. Principaux verbes pseudo-auxiliaires caractéristiques de Royal et de Sarkozy (comparés à tous les autres)

	Royal	Sarkozy
Croire	≈*	+
Devoir	+	+
Falloir	-	-
Pouvoir	-	+
Savoir	+	-
Vouloir	≈	+
Total des pseudo-auxiliaires	-	+

* ≈ sans différence significative par rapport à la moyenne des trois autres.

Le tableau se lit ainsi : par rapport à la moyenne des trois autres candidats, S. Royal ne se différencie pas pour l'emploi de la modalité du "croire" contrairement à N. Sarkozy qui la suremploie. Cette information ne surprend pas puisqu'il a affirmé à plusieurs reprises faire campagne "sur ses idées". Lors de son investiture, il annonce : "Je ferai campagne avec humilité et abnégation pour proposer mes idées et les faire partager par le plus grand nombre" (8 janvier 2007)¹¹. Il affirme aussi reprendre les

¹¹ Affirmation reprise les 14 et 26 janvier ; les 2 et 15 février ; les 9, 13, 22, 30 et 31 mars ; les 10, 11, 13, 18, 22, 24, 26, 27, 29 avril.

idées de Jaurès que la gauche a abandonné (11 février). Pour lui la politique ce n'est pas les rivalités de personnes mais "la confrontation des idées et des convictions" (13 avril).

Les deux finalistes s'accordent pour privilégier la modalité du *devoir* et éprouvent une certaine répugnance envers "il faut". En revanche, S. Royal fait campagne sur l'obligation morale (*devoir*) et sur la raison (*savoir*) alors que N. Sarkozy privilégie le possible (*pouvoir*), l'obligation morale et la volonté (*vouloir*), spécialement la volonté de parler (*vouloir dire*) et d'agir (*vouloir faire*). La dernière ligne indique que, par rapport à ses trois concurrents principaux, S. Royal a moins utilisé les pseudo-auxiliaires contrairement à N. Sarkozy qui était dans la situation inverse. La tension qui se dégage des propos de la première est donc significativement moins forte, au risque de paraître manquer de "battant" et de volonté par rapport au second.

Les pronoms

Après les combinaisons verbe+verbe, les pronoms *personnels* constituent le second indice d'une certaine conception du monde, de la place qu'y tient le locuteur et de ses relations avec les autres¹². L'emploi du *je* est révélateur d'une conception personnalisée de la gouverne alors que le recours au *nous* indique que le locuteur refuse le *je* individualiste et préfère se référer à un émetteur qui l'inclue avec d'autres (dont font normalement partie ceux qui l'écoutent). L'usage du *vous* sert surtout à désigner l'interlocuteur dans le sens du « je vous parle... » Dans une élection présidentielle où il s'agit de voter pour un candidat, la forte personnalisation des discours est la norme. Cependant, les pronoms ont d'autres usages *non-personnels* (Benveniste) : désigner des personnes absentes de l'interlocution (il, ils), une idée ou une chose (ce), établir une liaison entre les membres de la phrase (il, qui, que...), etc. Le tableau 12 récapitule la densité de ces deux catégories dans les discours des 4 candidats.

¹² E. Benveniste. "La nature des pronoms". *Problèmes de linguistique générale*. Paris, Gallimard, 1970, I, p. 251-257. C. Kerbrat-Orrecchioni. *L'énonciation de la subjectivité dans le langage*. Paris, A. Colin, 1981.

Tableau 11. Poids des pronoms personnels et des autres pronoms (en %)

	Pronoms personnels	Non-personnels
Bayrou	75,2	70,6
Le Pen	36,2	38,1
Royal (1er tour)	62,5	66,1
Sarkozy (1er tour)	61,4	68,0
Royal (2e tour)	72,4	69,4
Sarkozy (2e tour)	71,5	70,2

J.-M. Le Pen se distingue des autres candidats en tenant des propos plus impersonnels. À l'opposé, F. Bayrou est celui qui introduit les plus fortes relations avec son auditoire soit sous forme interlocutoire (je-vous) soit sous forme inclusive (je-nous). S. Royal et N. Sarkozy semblent avoir adopté sensiblement les mêmes stratégies au premier et au second tour. Au premier, le discours est assumé mais l'impersonnel domine ; au second tour, il est plus fortement personnalisé, la tension "je-vous" domine.

Nom ou verbe ?

Rappelons que les textes ne sont pas "enfournés" directement dans l'ordinateur. Au préalable, les graphies des mots sont corrigées et standardisées, puis la lemmatisation associe à chacun de ces mots un lemme comprenant un mot « vedette » (l'entrée de dictionnaire) et un code renvoyant à sa catégorie grammaticale. Cette dernière information permet de dresser une sorte de carte d'identité grammaticale de chaque locuteur (tableau 12).

J.-M. Le Pen est celui qui utilise le moins de pronoms personnels (et le moins de verbes) à l'opposé de N. Sarkozy et de S. Royal qui, au second tour, ont été ceux qui en ont employé le plus. Les pronoms ont des densités d'emploi relativement proches de celles des verbes et leur poids relatifs semble inverse à celui des substantifs et des déterminants.

Tableau 12. Densité d'utilisation des principales catégories grammaticales (% mots)

Catégories	Bayrou	Le Pen	Royal (1er tour)	Sarkozy (1er tour)	Royal (2e tour)	Sarkozy (2e tour)
Verbes	166,0	113,9	149,8	163,6	165,8	174,0
<i>Formes fléchies</i>	107,6	67,4	97,0	101,9	107,9	108,3
<i>Participes passés</i>	19,8	15,4	15,1	16,1	20,6	22,5
<i>Participes présents</i>	2,0	5,1	2,1	2,5	2,0	2,0
<i>Infinitifs</i>	36,6	26,0	35,6	43,1	35,2	41,2
Noms propres	16,8	25,7	18,7	19,5	19,2	22,3
Noms communs	173,1	211,4	184,7	181,1	173,5	169,0
Adjectifs	48,3	76,2	57,9	48,3	51,7	42,6
<i>Adj. participe passé</i>	4,3	9,4	6,8	5,1	4,7	4,3
Pronoms	145,8	74,3	128,6	129,4	141,8	141,7
<i>Personnels</i>	75,3	36,3	62,5	61,7	72,4	76,1
<i>Démonstratifs</i>	21,1	10,6	21,5	22,1	22,3	22,1
<i>Possessifs</i>	0,5	0,2	0,4	0,3	0,2	0,4
<i>Indéfinis</i>	4,4	2,9	4,3	6,2	4,6	5,7
<i>Relatifs</i>	34,5	19,3	32,0	31,3	31,6	32,5
Déterminants	172,8	217,0	180,2	181,1	168,8	176,1
<i>Articles</i>	124,1	143,3	133,7	140,6	125,8	129,5
<i>Nombres</i>	15,4	43,5	11,5	9,1	11,6	14,8
<i>Possessifs</i>	11,9	16,9	14,6	16,3	10,4	14,1
<i>Démonstratifs</i>	9,9	6,5	10,7	5,7	10,5	7,1
<i>Indéfinis</i>	11,6	6,7	9,6	9,5	10,4	10,5
Adverbes	64,0	54,5	63,7	68,2	70,3	75,1
Prépositions	145,1	173,5	152,1	148,7	144,5	135,6
Conjonctions	64,0	51,7	62,5	57,3	62,1	57,9
<i>Coordination</i>	30,5	34,7	37,4	26,1	34,0	23,7
<i>Subordination</i>	33,5	17,0	25,1	31,2	28,1	34,2

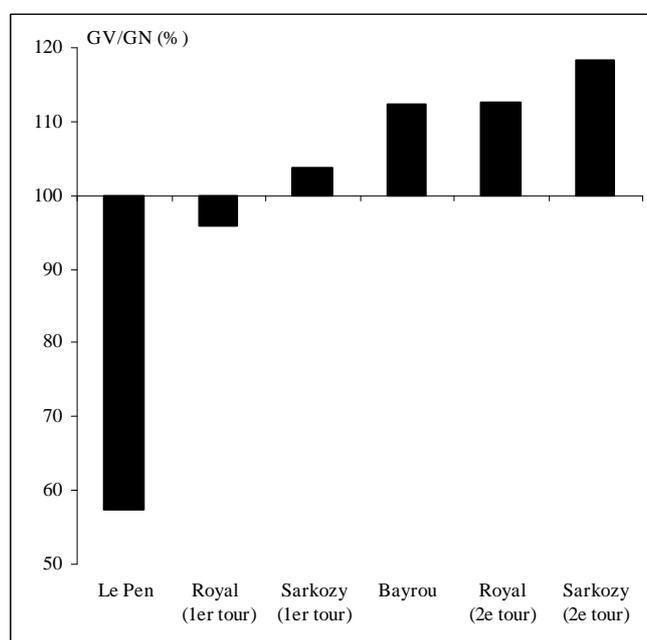
Les notions de groupes du nom et du verbe peuvent résumer ces relations. Le premier est constitué des pronoms, des verbes et des adverbes ; le second des substantifs, adjectifs, déterminants et prépositions. Naturellement, le partage n'est pas exclusif : certains adverbes peuvent se glisser dans le groupe du nom, certaines prépositions s'utilisent dans le groupe du verbe, etc. Mais au total, on observe une démarcation nette entre les candidats selon qu'ils privilégient l'un ou l'autre de ces deux groupes (tableau 14) En moyenne, le groupe du verbe couvre 37% des textes contre 62% pour le groupe du nom. En rapportant la performance de chaque candidat à ces moyennes, on obtient le graphique reproduit sous le tableau 13.

Tableau 13. Densité relative des groupes du nom et du verbe (% mots)

	GV	GN
Le Pen	260	739
Royal (1er tour)	367	631
Sarkozy (1er tour)	392	605
Royal (2e tour)	406	592
Bayrou	409	587
Sarkozy (2e tour)	425	569

* le total n'est pas égal à 100 du fait de l'absence des interjections, locutions et mots étrangers

Graphique 1. Densité relative des groupes du nom et du verbe (rapportée à la moyenne)



Ces chiffres soulignent la singularité de J.-M. Le Pen et, par contraste, la relative proximité des trois autres. Entre les deux tours de l'élection, le poids du groupe verbal augmente nettement chez N. Sarkozy comme chez S. Royal. Cela permet de penser que les densités relatives du nom et du verbe sont plutôt à rattacher à des choix de communication – ou aux circonstances – plutôt qu'à des dimensions stylistiques stables propres à chaque locuteur.

Selon l'interprétation dominante chez les théoriciens du style, un excédent du groupe du verbe (pronom+verbe+adverbe) indique une tension et une orientation vers l'agir, car la plupart des verbes impliquent un sujet et une action, donc une confrontation avec la réalité. À l'inverse, le suremploi du groupe du nom (substantif+adjectif+déterminant)

traduirait une pensée plus abstraite et moins engagée dans l'action. Dans la bouche d'un homme politique, cela trahirait une conception plus passive ou plus conservatrice du gouvernement et de l'Etat. Au Canada, cette tendance s'observe chez les gouvernements minoritaires dont la survie dépend du bon vouloir de l'opposition¹³.

4. Le style de la communication

Le concept de style appliqué aux discours désigne l'ensemble des choix de vocabulaire et d'expression qui caractérisent un locuteur et le distinguent des autres¹⁴. Tout locuteur a une façon particulière de dire les choses et son style est un écart par rapport à la pratique des autres. Dès lors, la statistique permet de mesurer objectivement les différences entre plusieurs locuteurs et donc de caractériser le style de chacun. Pour caractériser le style discursif d'un locuteur, il existe un grand nombre d'indices possibles. A titre d'exemple, les indicateurs suivants sont présentés : richesse et diversité du vocabulaire, longueur et complexité de la phrase.

Richesse ou simplicité ?

La richesse du vocabulaire est fonction de plusieurs facteurs dont deux ont une influence prépondérante : le nombre de thèmes traités – plus ce nombre est grand, plus la richesse sera élevée - et l'étendue du lexique que l'orateur choisit de mobiliser. Pour estimer cette richesse, on peut utiliser le nombre de vocables différents (V') qu'auraient contenu les autres corpus s'ils avaient eu une longueur égale à celle du plus petit (Royal 2^e tour : 48 269 mots)¹⁵.

¹³ Monière D., Labbé C; et Labbé D., "Les particularités d'un discours politique : les gouvernements minoritaires de Pierre Trudeau et de Paul Martin au Canada", art. cit.

¹⁴ Les études comparatives du style des hommes politiques sont peu nombreuses, déjà relativement anciennes et elles font peu de place à la statistique (en dehors de décomptes assez rudimentaires). Par exemple : J. Roche. *Le style des candidats à la présidence de la République*. Toulouse : Privat, 1969.

¹⁵ Calculée à l'aide du modèle de partition du vocabulaire. Cf. Hubert P. et Labbé D., "Un modèle de partition du vocabulaire", in Labbé D., Thoiron P. et Serant D., *Etudes sur la richesse et la structure lexicale*, Paris-Genève, Slatkine-Champion, 1998, p. 93-114.

Tableau 13 Richesse du vocabulaire des candidats
(classement en fonction décroissante de V')

	V' (48 269)	Royal (tour 2) = 100
Le Pen	5 381	165
Royal (tour 1)	3 719	114
Sarkozy (tour 1)	3 560	109
Sarkozy (tour 2)	3 455	106
Bayrou	3 383	104
Royal (tour 2)	3 255	100

J.-M. Le Pen l'emporte très largement sur les autres. Quand S. Royal (2^e tour) emploie 100 vocables différents, il en mobilise 165. Cela tient-il à la personnalité du leader du FN, politicien traditionnel, adepte du beau langage à l'ancienne ? Cela pourrait aussi signaler une incapacité à centrer sa campagne sur quelques thèmes sans cesse répétés comme l'ont plutôt fait les autres. Il faudrait disposer d'autres corpus de J.-M. Le Pen pour trancher sur ce point.

L'autre élément remarquable tient à la rupture de ton intervenue chez S. Royal entre les deux tours (recul de 14% de V'). Pendant la campagne du 1^{er} tour, elle aborde un nombre de thèmes un peu plus grand que les autres avec un vocabulaire relativement étendu. En revanche, au second tour, elle restreint considérablement le contenu de ses interventions et choisit la simplicité. C'était manifestement le choix de Bayrou dès le début. Mais les 3 principaux candidats (Bayrou, Royal et Sarkozy) sont extrêmement proches les uns des autres, ce qui suggère un même choix de communication. En effet, les fluctuations observées chez S. Royal indiquent simplement que la "richesse du vocabulaire" ne mesure pas l'étendue du lexique de l'orateur mais qu'elle dépend de sa stratégie de communication.

Il est possible d'aller plus loin en scindant la "richesse du vocabulaire" en deux indices : la *diversité* et la *spécialisation* du vocabulaire¹⁶. Dans sa relation à l'auditoire et à l'objet qu'il traite, l'orateur peut choisir la diversité de l'expression, ou au contraire la simplicité, sans reculer devant la répétition. Dans le premier cas, il mobilisera un vocabulaire étendu pour préciser sa pensée au risque de brouiller son message. Dans le second cas, son message sera plus accessible mais il risque d'ennuyer son auditoire ou de laisser paraître une pensée "pauvre". *L'indice de diversité* rend compte de cette propension à diversifier l'expression l'aide du nombre de mots différents observés dans

¹⁶ On trouvera des éléments de comparaison (Premiers ministres français, canadiens et québécois depuis 1945) dans Labbé D. & Monière D., *Le vocabulaire... op. cit.*, p 131-133.

toutes les tranches — ici de 1000 mots contigus — que l'on peut extraire d'un texte, nombre estimé à l'aide du modèle de partition du vocabulaire¹⁷. L'indice neutralise ainsi les différences de tailles entre les textes et fournit une norme de comparaison entre corpus (tableau 14).

Tableau 14. Diversité et spécialisation des vocabulaires

	Diversité (1000 mots)	Ecart-type	Spécialisation (%)
Bayrou	323,5	8,4	5,0
Le Pen	399,1	9,8	7,2
Royal (1 ^{er} tour)	353,1	8,9	7,2
Royal (2 ^{er} tour)	317,2	8,3	2,9
Sarkozy (1 ^{er} tour)	310,9	8,2	15,1
Sarkozy (2 ^e tour)	282,2	7,6	23,4

A l'écrit, l'indice de diversité varie généralement entre 425‰ (texte très élaboré, voire précieux) à 310‰ (texte peu élaboré contenant beaucoup de mots répétés à courte distance), le point de basculement se situant vers 370‰. Au dessus de cette valeur moyenne, la recherche dans l'expression peut tourner au maniérisme (relation au destinataire à dominante poétique) ; en dessous, la volonté de faire simple domine (relation au destinataire à dominante pédagogique ou polémique).

L'indice de *spécialisation* mesure le poids relatif des vocables réservés à un thème et à un seul dans l'ensemble du vocabulaire utilisé par le locuteur. Entre 0 et 10%, le locuteur a choisi l'hyper-généralité : le même vocabulaire sert quel que soit le thème abordé (c'est le cas de la majorité des politiques contemporains). Entre 10 et 20%, la spécialisation est moyenne. Les valeurs au-dessus de 20% se rencontrent en général dans les documents techniques ou dans des ouvrages "à plusieurs mains".

Seul J.-M. Le Pen a choisi de diversifier son expression – ou de s'en tenir à un discours écrit assez littéraire. Au premier tour, S. Royal utilise une forme à la limite inférieure de l'écrit, mais au second tour, elle choisit un style qui s'apparente plus à l'oralité avec une plus faible diversité.

F. Bayrou et, surtout N. Sarkozy, ont choisi la simplicité et la répétition plus caractéristiques de l'oralité. Au second tour, N. Sarkozy renforce encore ce choix (sa diversité passe en dessous de 300 (oral familier, conversation quotidienne).

¹⁷ Hubert P. et Labbé D., "Un modèle...", *art cit* ; *Vocabulary Richness*, communication au congrès de l'ALLC-ACH. Paris, La Sorbonne, mars 1994. Reproduit dans : *Lexicometrica*, numero 0, hiver 1997-98.

L'indice de spécialisation révèle deux choix opposés :

- Nicolas Sarkozy a choisi de traiter un nombre important de thèmes en mobilisant, pour chacun de ces thèmes, un vocabulaire spécifique à ce thème. En quelque sorte, il privilégie le "mot juste" sur le "mot passe-partout". Il renforce encore cette attitude au second tour pour atteindre une valeur qui se rencontre rarement dans le discours politique contemporain.

- tous les autres ont choisi l'hyper-généralité avec un vocabulaire spécialisé inférieur à 10% du vocabulaire total. S. Royal entre les deux tours se distingue tout particulièrement en faisant un choix exactement inverse à celui de N. Sarkozy comme si elle se réfugiait dans les généralités.

Longueur et complexité de la phrase

Enfin la longueur de la phrase et sa structure sont deux éléments caractéristiques d'un style. La "phrase" est une suite de mots délimitée par le point ou, lorsqu'ils sont suivis d'un mot commun en majuscules, par les trois points de suspension, le point d'interrogation, le point d'exclamation.

Pour établir la longueur moyenne de la phrase de chaque candidat, la longueur du corpus (nombre total de mots) est divisé par le nombre de phrases. Pour calculer l'indice de complexité de la phrase, on divise le nombre de ponctuations faibles (, ; :) par ce même nombre total de phrases.

Tableau 15. Indices de longueur et de complexité des phrases

	Longueur	Complexité
Bayrou	28.71	2.47
Le Pen	28.06	2.71
Royal	27.48	2.31
Sarkozy	22.38	1.33

Ces deux indicateurs montrent que le style oratoire de Sarkozy se distingue nettement de celui de ses adversaires : ses phrases sont beaucoup plus courtes et moins complexes (elles contiennent deux fois moins de césures). Ce phrasé « coup de poing » ressemble aux messages publicitaires. Il dénote une recherche d'efficacité communicationnelle et révèle un style personnel radicalement différent de celui

dominant dans la classe politique française où l'on fait des phrases relativement compliquées, du moins en regard du goût contemporain.

M. Cressot écrivait à ce propos que "La phrase longue répond à une vision totale et complexe et à une volonté de rendu synthétique."¹⁸ Elle est l'outil de la démonstration. La phrase courte présente une vision fragmentaire et nette. Elle est plus incisive et appropriée à l'argumentation et à la polémique.

Naturellement, un même orateur utilise ces deux types de phrases mais en quantité variable et la proportion plus ou moins grande des différentes longueurs donne une coloration particulière au discours.

Conclusions

L'analyse des discours de campagne met bien en relief les différences de style et de stratégie de communication des candidats, spécialement des deux finalistes. N. Sarkozy a mené la campagne la plus intense et il a été constamment au centre du débat électoral à la fois dans son discours et dans ceux de ses adversaires. A l'inverse, durant toute la campagne, sauf lors du débat télévisé, il a choisi d'accorder peu d'attention à ses concurrents. Il a adopté un style de communication familier en faisant des phrases plus courtes et moins complexes que ses rivaux. Il a trouvé - grâce à ses plumes de l'ombre ? - les phrases « coup de poing » qui cristallisent les idées dans la mémoire des électeurs. Après le premier tour, ses discours étaient aussi plus ancrés dans l'espace et le temps. Il a fait un usage plus abondant de chiffres pour mieux faire ressortir ses compétences gestionnaires et sa plus grande maîtrise des dossiers. Il a aussi privilégié un vocabulaire plus spécialisé. À cet égard, l'analyse des indices de spécialisation révèle deux stratégies différentes de communication. N. Sarkozy a mobilisé un vocabulaire spécifique plus important pour chacun des thèmes abordés. Il a privilégié l'expression simple et juste alors que ses adversaires ont préféré l'hyper généralité (les mêmes mots sont utilisés pour parler de différents sujets). De cette façon, peut-être a-t-il ainsi réussi à incarner l'idée de rupture avec la langue de bois politicienne ?

¹⁸ Cressot M., *Le style et ses techniques*, Paris, PUF, 1963, p. 207-208. Voir aussi notre article : « Les styles discursifs des premiers ministres québécois de Jean Lesage à Jean Charest », *Revue canadienne de science politique*, mars 2008.

Les deux candidats finalistes se partagent également quant aux thèmes de campagne. Celle de S. Royal a été dominée par la défense des services publics et par les thèmes sociaux : sécurité sociale, logement, niveau des salaires, temps et conditions de travail, dialogue social ainsi que la recherche et l'éducation supérieure. Même si elle a voulu prendre une certaine distance vis-à-vis de sa famille idéologique pour moderniser son discours et ratisser plus large, elle est demeurée fidèle aux thèmes traditionnels de la gauche. N. Sarkozy a misé sur des enjeux à la fois concrets et symboliques : pouvoir d'achat, plein emploi, valeur du travail mais aussi : identité nationale, sécurité, anti-fiscalisme. Ainsi, a-t-il pu braconner à la fois sur les terres de l'extrême-droite et du centre gauche.

Cette recherche montre la pertinence de la lexicométrie pour éclairer des questions que se pose la science politique. Elle procure une description rigoureuse et exhaustive de ce que disent les politiciens permettant d'aborder d'autres questions, notamment celle concernant les effets de la communication électorale sur les électeurs. Évidemment, le but d'une campagne électorale est de persuader le maximum d'électeurs en proposant une vision de l'avenir, en mettant en valeur le candidat et en dépréciant ses adversaires. Bref, les mots sont choisis pour obtenir les votes des électeurs. Mais, malgré ce qu'affirment les conseillers en communication, aucune corrélation n'a pu être sérieusement établie entre le contenu des discours et la réussite ou l'échec des candidats. On peut certes tester, par un protocole expérimental, les effets des discours auprès de petits groupes, mais ces expériences sont artificielles et ne reproduisent pas la dynamique d'une campagne électorale. Pour ce faire, il faudrait mettre en relation les discours réellement prononcés et l'évolution de l'opinion publique. On pourrait par exemple, d'une part, tenter de vérifier si les sondages influencent la production des discours politiques en repérant les ruptures thématiques ou les changements stratégiques et en les corrélant aux résultats des sondages. On pourrait d'autre part relier les variations des intentions de vote, exprimées par les sondages, et le contenu des discours prononcés. Cependant, il faudrait tenir compte du processus de percolation, c'est-à-dire du décalage inhérent entre l'émission du message, sa transmission - notamment par les leaders et les relais d'opinion - et son éventuelle absorption par les électeurs. La lexicométrie produit les données nécessaires à la première étape de cette vérification en établissant le contenu de ce qui est émis à un moment donné mais aussi, comme nous

venons de le montrer, les thématiques et les stratégies de communication choisies par l'émetteur. Il reste donc à mesurer la relation entre ces données et l'évolution de l'opinion publique pour pouvoir saisir les effets des discours.