



HAL
open science

Les thèmes dans le discours électoral de candidature à la députation sous la Cinquième République. Perspective de genre (1958-2007)

Magali Guaresi

► To cite this version:

Magali Guaresi. Les thèmes dans le discours électoral de candidature à la députation sous la Cinquième République. Perspective de genre (1958-2007). Mots: les langages du politique, 2015, Thèmes et Thématiques dans le discours politique, 108, pp.15 - 37. 10.4000/mots.21977 . hal-01228476

HAL Id: hal-01228476

<https://hal.science/hal-01228476v1>

Submitted on 26 Jun 2019

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.

Les thèmes dans le discours électoral de candidature à la députation sous la Cinquième République. Perspective de genre (1958-2007)

Key electoral issues in parliamentary campaign discourse under the Fifth Republic. A gender approach (1958-2007)

Los temas dentro de los discursos electorales de la candidatura a diputado bajo la Quinta República. Perspectivas desde el género (1958-2007)

Magali Guaresi



Édition électronique

URL : <http://journals.openedition.org/mots/21977>

DOI : 10.4000/mots.21977

ISSN : 1960-6001

Éditeur

ENS Éditions

Édition imprimée

Date de publication : 6 octobre 2015

Pagination : 15-37

ISBN : 978-2-84788-727-3

ISSN : 0243-6450

Référence électronique

Magali Guaresi, « Les thèmes dans le discours électoral de candidature à la députation sous la Cinquième République. Perspective de genre (1958-2007) », *Mots. Les langages du politique* [En ligne], 108 | 2015, mis en ligne le 06 octobre 2017, consulté le 17 mai 2019. URL : <http://journals.openedition.org/mots/21977> ; DOI : 10.4000/mots.21977

Les thèmes dans le discours électoral de candidature à la députation sous la Cinquième République. Perspective de genre (1958-2007)

La mise en place de la parité a, ces dernières années, suscité de nombreuses études consacrées aux représentations discursives des fonctions politiques exercées par les femmes (Bonnafous, 2001 ; Lévêque, Dulong, 2002 ; Guionnet, 2002 ; Wodak, 2003 ; Achin *et al.*, 2007 ; Desmarchelier, Rennes, 2005 ; Coulomb-Gully, 2009)¹. Si les manières de dire et de faire de la politique des nouvelles élues ont été, à raison, placées au centre des analyses, les thématiques défendues par les représentantes de la nation dans leurs discours officiels ont été quelque peu ignorées². Pourtant, dès les années quatre-vingt, les travaux pionniers de politologues telle Mariette Sineau (1988) émettaient l'hypothèse de l'inclusion, de l'exclusion ou de la modulation de sujets politiques consécutifs à la féminisation de la représentation politique.

Si la question des contenus, des sujets ou des thématiques du discours politique des locutrices n'a pas fait l'objet de développements importants, c'est que la mise en exergue des thèmes dans un texte demeure en histoire du discours une difficulté. La logométrie, et ses dernières avancées épistémologiques, méthodologiques et techniques, semblent néanmoins autoriser des protocoles de repérage thématique toujours plus perfectionnés. Établissant les textes non plus seulement comme des sacs de mots (Muller, 1977) mais comme des réseaux d'unités linguistiques organisées (Adam, 2006 ; Brunet, 2007 ; Mellet, Longrée, 2012 ; Mayaffre, 2014), la logométrie entend permettre le prolongement d'une approche statistique paradigmatique par un traitement plus global de la textualité, plus à même de rendre compte de son organisation séquentielle et, ici, thématique. Le postulat de départ est simple : si les mots fréquents des corpus sont des indices des sujets ou des *topics* du discours,

1. Des travaux sur l'anglais, dans une perspective de traitement automatique des langues, se sont également emparés de la question du genre dans le discours politique. Voir récemment Yu (2013).
2. Ainsi, les travaux de science ou d'histoire politiques traitent-ils essentiellement des ressources politiques ou des trajectoires des actrices politiques qu'ils considèrent (Sineau, 2011 ; Tricaud, 2012).

la recherche de thèmes ne saurait uniquement procéder de mots isolés mais semble devoir passer par l'observation de l'organisation lexicale des textes. En analyse du discours, on connaît l'importance de la co(n)textualisation dans l'émergence du sens (Rastier, 2001, 2011). Comme tout autre fait textuel, les thématiques ne sauraient s'appréhender exclusivement à partir des mots seuls, quand bien même ils seraient les plus redondants. C'est, au contraire, dans la rencontre entre les mots dans l'entrelacs de la textualité pensée comme une combinaison entre la sélection des unités du discours sur l'axe paradigmatique et leur combinaison sur l'axe syntagmatique (Adam, 2006 ; Mayaffre, 2008) que se construisent et se saisissent les thématiques.

Méthodologiquement, l'une des voies du dépassement de l'analyse d'unités décontextualisées pour atteindre celle des relations entre les items du texte passe par la cooccurrence. Quantifiant la coprésence des termes d'un corpus deux à deux, l'approche cooccurrence constitue une avancée dans le balisage de parcours interprétatifs (con)textuels, sémantiques ou thématiques. Précisément, les propositions de Max Reinert (1983, 1990), Jean-Marie Viprey (1997, 2006) ou Étienne Brunet (2012), implémentées dans les logiciels Hyperbase ou Iramuteq³ utilisés dans cette contribution, visent à saisir par les proximités statistiques les proximités sémantiques.

La recherche thématique et la réflexion méthodologique proposées ici s'effectuent, comme toujours en logométrie, au sein d'un corpus rigoureusement constitué sur la base des hypothèses de travail de l'analyste. Concrètement, le balisage des thématiques du discours politiques des élues s'effectuera sur un corpus de professions de foi rassemblant la quasi-totalité des textes de femmes élues au Palais Bourbon entre 1958 et 2007 et un échantillon de textes d'hommes rédigés dans des conditions de production comparables⁴. Source précieuse de l'histoire politique (Prost, 1974), la proclamation électorale est un lieu privilégié de la construction identitaire et programmatique des futurs élus⁵ (Déloye, 1999 ; Lévêque, 2005). Si elle est un texte de présentation de soi, elle est aussi un discours de politique générale dans lequel les candidates s'expriment sur l'ensemble des enjeux électoraux qu'elles jugent déterminants

3. Hyperbase a été conçu et développé par Étienne Brunet à l'Université Nice Sophia Antipolis. Iramuteq (Interface de r pour les analyses multidimensionnelles de textes et de questionnaires) a été conçu et développé par Pierre Ratinaud à l'Université Toulouse Le Mirail. Il propose en accès libre les méthodes et calculs de Max Reinert implémentés dans Alceste.
4. Précisément, le corpus de travail rassemble la quasi-exhaustivité des professions de foi – conservées aux Archives de l'Assemblée nationale dans le *Barodet* depuis 1881 – des candidates élues aux élections législatives tenues au scrutin uninominal majoritaire (l'élection de 1986 au scrutin de liste est écartée puisqu'elle ne permet pas l'analyse de la variable « sexe » de chaque candidature). Un échantillon raisonné de textes d'hommes sélectionné selon une logique chronologique, politique et géographique a été constitué pour la comparaison : pour un texte de femme élue a été retenu un texte d'homme élu à la même date, du même bord politique et de la même zone géographique. L'ensemble compte 719 textes pour 463 236 occurrences.
5. Conformément aux conventions orthographiques adoptées par la revue *Mots. Les langages du politique*, le masculin générique sera utilisé pour désigner les locutrices et les locuteurs du corpus.

dans la campagne et servir au mieux leur candidature. Pour ces raisons, et parce que le corpus en permet la comparaison sexuée, les professions de foi sont un observatoire pertinent des thématiques développées par les futures élues de la République.

Cette contribution se fixe deux objectifs. Le premier vise à éclairer l'économie thématique du discours électoral à l'aune d'une perspective de genre. Après avoir établi la structuration fondamentale des thèmes sur l'ensemble du corpus, l'analyse réintroduira le contraste sexué entre les professions de foi : les femmes et les hommes développent-ils les mêmes enjeux électoraux ? Quelles sont les thématiques spécifiquement portées par les unes ou par les autres ? Poursuivi en parallèle, le second objectif concerne l'explicitation méthodologique des protocoles de repérage thématique développés ces dernières années en logométrie.

La structure thématique des professions de foi des femmes sous la Cinquième République

Cooccurrences généralisées

Au cœur du programme d'une statistique contextualisante, la cooccurrence apparaît comme l'une des voies privilégiées de la recherche des thèmes dans la pratique logométrique. Récemment caractérisée à plusieurs occasions (Mayaffre, Viprey, 2012 ; Mayaffre, 2014) comme fait statistique et unité textuelle, la cooccurrence permet une approche de la relation entre les mots sans conduire à l'abandon précoce du critère quantitatif comme garant de la significativité du phénomène observé et interprété. En considérant les unités du texte deux à deux, en fonction de la régularité et de la significativité de leur association au sein d'une fenêtre textuelle prédéfinie (souvent le paragraphe)⁶, la statistique cooccurrence entend traiter de paires de mots comme de véritables noyaux de sens⁷ susceptibles de rendre compte de la textualité dans sa double dimension paradigmatique et syntagmatique, décontextualisée et (re-)contextualisée, quantitative et qualitative⁸. Généralisée (Viprey, 1997, 2006), la statistique cooccurrence

6. Sans prétendre clore les discussions sur le choix et l'impact de la taille de l'empan contextuel, nous posons ici le paragraphe comme l'unité la plus adaptée au calcul cooccurrence. Traduction topographique d'une unité thématique dans le discours politique, le paragraphe s'est révélé, de manière empirique, être l'une des options les plus favorables au repérage des attractions et des répulsions d'ordre sémantique, tout en limitant le poids de la syntaxe que soulignerait le choix d'une fenêtre textuelle plus courte.
7. La paire de mots cooccurents peut en effet être comprise comme l'unité minimale formalisable de contextualisation d'un mot par un autre (Mayaffre, 2008).
8. La recherche des cooccurents d'un terme procède d'une décontextualisation puis d'une recontextualisation. La statistique repère d'abord systématiquement les mots très utilisés (en fréquence absolue) dans un texte comme autant de mots-pôles. Chaque mot-pivot ainsi défini

établit le voisinage entre tous les mots. Pour chacun d’entre eux, un profil cooccurrentiel est défini, laissant apparaître le contexte minimal d’emploi du mot-pôle, déjà porteur de sens. Pris dans leur ensemble, et comparés, les profils d’association construisent la trame sémantique du corpus. Ainsi, dans la combinaison de la fréquence et de la séquence, la cooccurrence généralisée permet l’émergence inductive de saillances, de polarisations ou d’isotropies (Viprey, 2006) informant la structuration thématique d’un texte.

Techniquement, la fréquence de la coprésence des mots choisis, ici les substantifs privilégiés du discours⁹, est reportée dans une matrice de cooccurrence dont on trouvera un exemple dans le tableau 1 :

Tableau 1. Exemple de matrice de co-occurrences généralisées

	accord	achat	action	agriculteur	aide	allocation	etc.
accord	XXX	a*	b	c	d	e	...
achat	a	XXX	f	g	h	i	...
action	b	f	XXX	j	k	l	...
agriculteur	c	g	j	XXX	m	n	...
aide	d	h	k	m	XXX	o	...
allocation	e	i	l	n	o	XXX	...
etc.	XXX

*Dans les cellules, les lettres a, b, c, etc., chiffrent la coprésence des deux mots considérés en tête de ligne et de colonne dans les paragraphes du corpus.

Chaque cellule du tableau comptabilise la rencontre entre les lemmes, considérés deux à deux, dans le corpus ; chaque ligne (ou chaque colonne) donne le profil cooccurrentiel de chaque substantif sélectionné. L’ensemble de la matrice formalise le texte comme un tissage où les réseaux lexicaux, thématiques et sémantiques s’élaborent dans la rencontre entre les mots (Viprey, 2006 ; Ben Hamed, Mayaffre, 2014 ; Mayaffre, 2014).

L’analyse factorielle des correspondances ou AFC (Benzécri, 1981) permet, dans un dernier temps, de traiter la matrice et de représenter les relations cooccur-

est ensuite replacé dans son contexte (tous les paragraphes où il est utilisé) dans lequel sont déterminées ses attirances ou répulsions lexicales, qui sont précisément quantifiées et hiérarchisées. Le mode de calcul hypergéométrique implémenté dans Hyperbase sera utilisé ici (Brunet, 2007). Soit s = nombre de paragraphes, f = fréquence du mot-pôle dans le texte, g = fréquence du mot cooccurrent dans le texte et k = cooccurrence observée. Alors : $\text{Prob}(x=k) = \frac{(f! (s+g)! g! (f+s)!)}{(k! (f-k)! (g-k)! (s+k)! (f+g+s)!)}$

- Implémenté dans Hyperbase, ce traitement embrasse jusqu’à 300 mots. L’analyse exhaustive de tous les mots de corpus conséquents demeure aujourd’hui techniquement difficile. Dans notre corpus, de taille moyenne en logométrie mais déjà riche de plus de mille vocables, le nombre de paires à observer serait considérable. Aussi Hyperbase se contente-t-il, pour des raisons de durée du traitement mais également pour la lisibilité du graphique qui en résulte, de sélectionner 300 lemmes sur critères grammaticaux et de fréquence. Ici, compte tenu de l’objectif de recherche fixé, ce sont les substantifs qui seront choisis.

rentielles de chaque nom, figurées par des proximités ou des distances entre les mots. Les mots *géographiquement* proches sur le graphique, c'est-à-dire les items ayant des profils associatifs semblables, témoignent d'une saillance sémantique significative¹⁰ prenant sens en soi mais également dans le phénomène de répulsion des autres polarisations lexicales formées sur la représentation factorielle.

Panorama thématique du corpus : les fondamentaux de la représentation parlementaire

Ainsi, appliquée aux 300 substantifs les plus récurrents du corpus des professions électorales, la fonction « corrélats » implémentée dans Hyperbase produit une carte de l'organisation lexicale des candidatures à la députation sous les douze législatures de la Cinquième République considérées. Quatre regroupements thématiques se dessinent dans les quatre quadrants de l'AFC pour définir les grands pôles de la parole de candidature aux élections législatives. Le nuage de points produit par Hyperbase décrit la structure fondamentale et robuste du discours de profession de foi contemporain (voir figure 1).

La première lecture de la représentation factorielle s'organise autour de l'axe horizontal qui regroupe 13 % de l'information et oppose un vocabulaire politique à des mots socio-économiques. À l'extrême gauche du graphique se rassemblent les termes de l'élection (*candidature, suffrage, rassemblement...*) quand l'extrême droite réunit un vocabulaire axé sur les questions économiques et sociales des programmes politiques : *crédit, commerçant, retraité, fiscalité*. Cette vue horizontale doit être complexifiée par une lecture verticale le long du second axe (5 % de l'information). Le deuxième facteur de l'AFC oppose la partie haute de la figure constituée d'un vocabulaire local de la partie basse organisée par un vocabulaire national. En haut, se tiennent les références au *quartier* à la *commune* ou à la *ville* d'appartenance des locuteurs et des locutrices. En bas, c'est davantage de la *France*, des *Français* voire de l'*Europe* dont il est question. Au croisement de ces deux axes (de gauche à droite, le politique *versus* le socio-économique, et de haut en bas, le local *versus* le national), quatre quadrants se détachent pour former les quatre facettes du discours de candidature parlementaire sous la Cinquième République.

Le quadrant supérieur gauche (figure 2) traite de la dimension locale de l'élection et du rôle de *député*.

Dans cette partie du graphique se rassemblent les ressources politiques (*maire, conseiller, expérience*) et personnelles (*conviction, présence, écoute*)

10. Autrement dit, si la position sur le graphique informe l'analyste sur la parenté des profils cooccurrentiels des mots proches, elle ne dit rien de la coprésence effective des mots dans les paragraphes. Deux mots peuvent avoir deux environnements lexicaux similaires – et concourir à la formation d'une thématique – sans jamais être cooccurrents entre eux.

Figure 1. Analyse factorielle des cooccurrences du corpus « Professions de foi (1958-2007) »

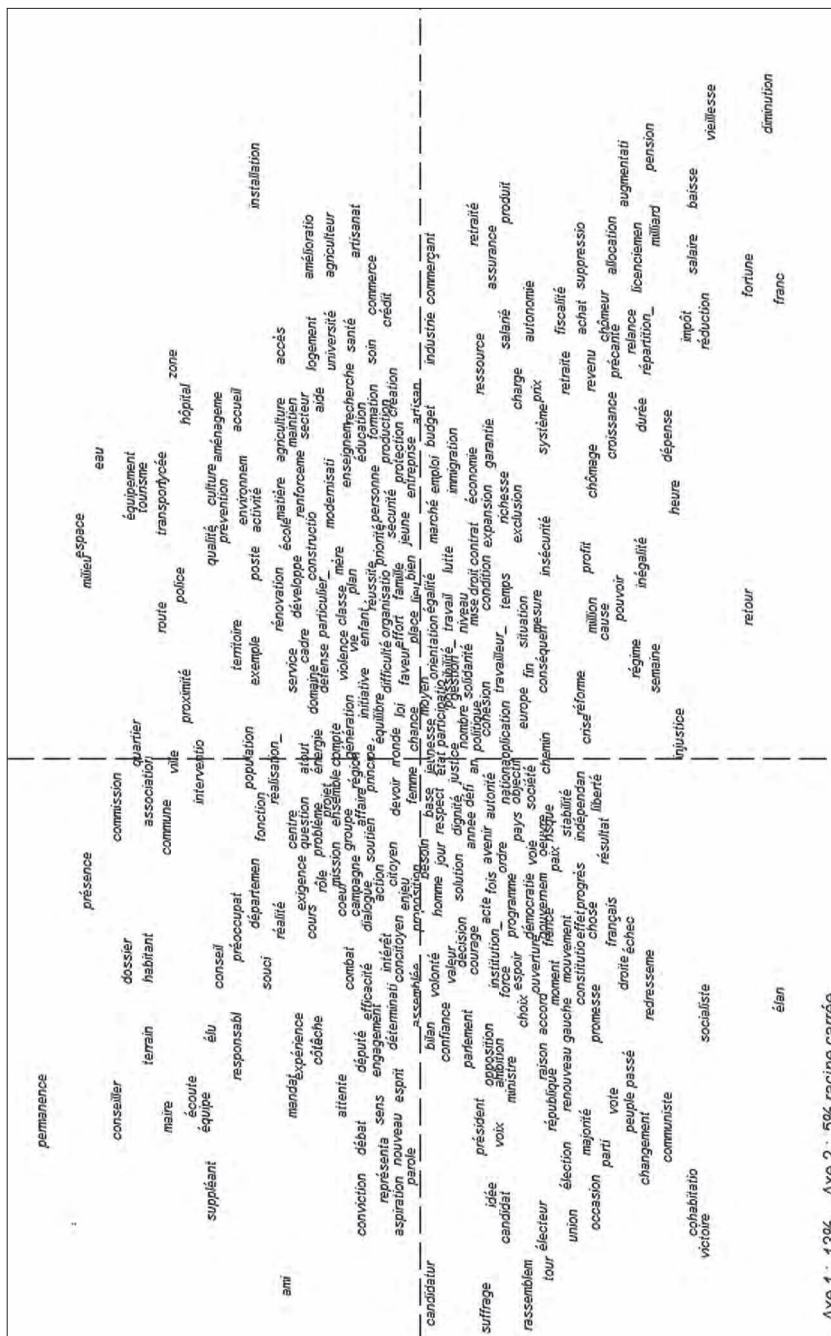
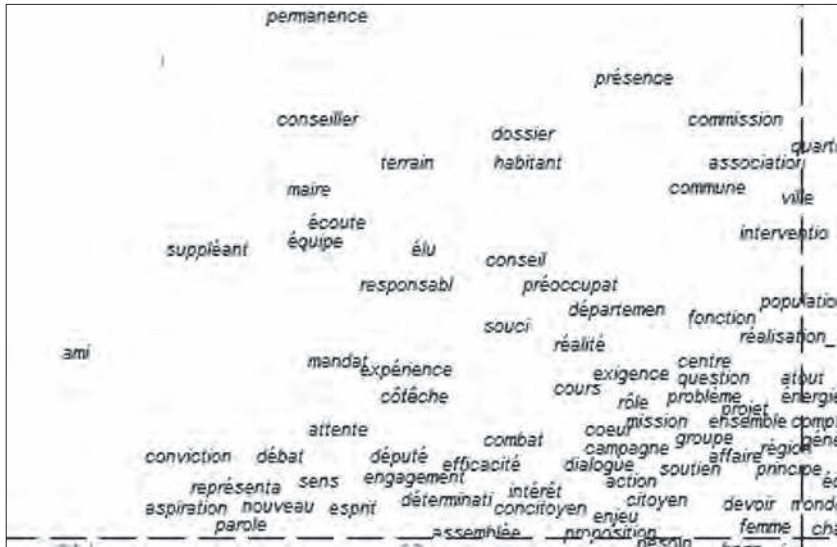


Figure 2. Thématique « Politique locale » dans le corpus « Professions de foi (1958-2007) »



communiquées en soutien aux candidatures et qui s'imposent comme les qualités indépassables des élus dans leur circonscription. La présence et l'importance de ce groupement lexical informe d'un aspect de la représentation sous la Cinquième République : l'insistance sur la dimension locale, c'est-à-dire aux prises avec les *réalité[s]*, de l'engagement politique des locuteurs. Dans cette thématique, la proximité avec les (*con*)citoyens passe par la *présence* physique dans les territoires, sur le *terrain* (la *permanence* étant l'un des lieux phares de la rencontre représentés/représentants) mais également par le *souci* accordé aux *attentes* et aux *préoccupations* pour les *dossiers* ou les *problèmes* des *habitants*. C'est dans cet espace factoriel que se rassemblent les mots de l'*engagement* et des qualités et compétences personnelles des candidats : *détermination*, *action*, *efficacité*¹¹ viennent construire le *combat* politique non seulement comme un affrontement d'idées mais comme une opposition de personnalités soucieuses d'incarner un porte-*parole* crédible des *aspirations* des électeurs et des électrices. Pour exemple dans le corpus des textes de cette thématique, on pourra lire les engagements de Nicole Ameline, candidate dans le Calvados en 2002 et de Jacqueline Thome-Patenôtre, candidate en Seine et Oise en 1958 :

11. Notons néanmoins que ces lemmes sont tirés vers le bas du graphique consacré à la dimension nationale de l'engagement politique.

Privilégier la réussite. Doter la région d'un ensemble de filières de formations cohérentes, tournées vers l'avenir, en prise directe avec la réalité environnante [...]. Rénover le lycée Paul Éluard de Saint-Denis et créer un lycée polyvalent à Pierrefitte-Stains. Ouvrir les universités Paris 8 et Paris 13 à tous les bacheliers et non bacheliers, et garantir leurs conditions d'études. (Berthelot, 1988, PC, Seine-Saint-Denis)

Plutôt situées dans la partie haute du graphique, ces questions d'aménagement s'enracinent dans des *territoire[s]*, des *zone[s]*, des *espace[s]* particuliers (souvent locaux) et renouent avec la valorisation de l'action locale des candidats : « Avec vous et pour vous, j'ai travaillé pour que le quartier développe ses équipements et améliore son environnement », fait valoir Hélène Missoffe, candidate à Paris en 1981.

À ce niveau de la textualité (c'est-à-dire l'environnement lexical des 300 substantifs les plus redondants dans les professions de foi), la représentation factorielle de l'organisation lexicale du corpus ci-dessus donne à voir les soubassements structurants de la proclamation électorale aux scrutins législatifs. L'application du calcul de la cooccurrence généralisée sur chaque corpus sexué (les textes des femmes d'un côté, ceux des hommes d'un autre) confirme le primat de l'influence du genre discursif des professions de foi sur celle du genre (féminin/masculin) dans l'organisation thématique du corpus. Les nuages de points tirés de l'analyse de chaque sous-corpus sexué – que nous ne reproduisons pas ici tant ils sont proches de celui reproduit en figure 1 – ne montrent pas de différence fondamentale, preuve de la solidité thématique des professions de foi davantage contrainte par l'institution politique que par les rapports sociaux de sexes dans le champ parlementaire.

La force de l'analyse cooccurentielle menée ici réside dans la présentation d'une vue synthétique de l'économie thématique du corpus et, par là, de l'originalité du discours des proclamations électorales qu'A. Prost (1974, p. 10) avait qualifié de « genre littéraire particulier, à la fois solennel et laconique, personnel et stéréotypé, mais unique ». Organisées en quatre grands pôles, décrits par l'AFC, les thématiques du corpus électoral fixent le cadre de la représentation parlementaire contemporaine. Deux groupements lexicaux, eux-mêmes subdivisibles en deux, mettent en forme les rôles parlementaires sous le régime quinto-républicain : d'une part, une présentation de soi des candidats et d'autre part, le détail d'un programme politique ; chacun des deux éléments recouvrant une dimension nationale et un ancrage local. C'est la tension entre le territoire et la nation, qui fonde la représentation depuis Sieyès, qui se manifeste ici (Gaxie, 2000 ; Manin, 1995), les députés étant à la fois appelés à former une majorité gouvernementale tout en étant liés à une circonscription, lieu de l'élection¹² ; à se distinguer pour se faire l'interprète de l'intérêt

12. Notons que si la structure « national/local » ou « présentation de soi/programme politique » est caractéristique de la profession de foi à la députation, le poids de ces pôles de la parole

général tout en restant proches des intérêts d'un territoire et d'une population. En plaçant le curseur diversement entre le local et le national, nombreux sont les paragraphes du corpus qui évoquent explicitement l'équilibre (ou les hésitations) de la représentation parlementaire sous la Cinquième République :

Nous voici arrivés au terme de cette législature. Depuis près de cinq ans que je vous représente à l'Assemblée nationale, je me suis efforcée d'y accomplir avec conscience la mission que vous m'avez confiée. Comme à mon sens la tâche d'un député est aussi de représenter la circonscription, j'ai voulu garder un contact régulier avec vous, afin de rapporter à Paris les doléances, les souhaits et les indications que chacun d'entre vous m'a apportés lors de plus de 300 permanences que j'ai tenues dans les 55 communes de ma circonscription et lors des visites que j'ai reçues à la mairie de Pont-de-Buis, ou chez moi à Plomodiern. (Dienesch, 1967, UNR, Côtes-du-Nord)

Femme de parole et femme d'action, je m'engage, élue député, à relayer avec force et conviction au Parlement vos demandes, vos projets, à faire entendre votre voix pour améliorer votre vie quotidienne, à porter haut les couleurs du Bessin et de la Côte-de-Nacre, à les faire connaître et reconnaître nationalement. (Dumont, 1997, PS, Calvados)

Un investissement thématique genré ?

Sur l'ensemble de la Cinquième République, l'analyse cooccurrentielle établit la structuration solide de la profession de foi et la révèle globalement invariante au sexe des candidats. Il ne s'agit dès lors plus de se demander si l'architecture thématique des discours électoraux est renouvelée par la féminisation des instances électives, mais plutôt de s'intéresser aux éventuels thèmes structurants de la profession de foi que les locutrices investiraient de manière privilégiée. Parmi les indépassables de la parole élective contemporaine décrits *supra*, y en a-t-il qui sont spécifiquement, c'est-à-dire statistiquement, surdéveloppés par les femmes ? Repérées dans le corpus global, les thématiques seront examinées à l'aune de leur « poids » dans chaque ensemble discursif sexué.

Mondes lexicaux

Pour poursuivre l'analyse dans ce sens, c'est le logiciel Iramuteq qui est choisi. Toujours à partir d'un calcul de cooccurrences mais selon une philosophie un peu différente, Iramuteq permet une classification lexicale rendant compte de regroupements isotropiques comme autant de mondes lexicaux (Reinert, 1993). Ces classes de vocabulaire, aussi stables et homogènes que possible, sont définies à l'aide d'une classification hiérarchique descendante appliquée au corpus

électorale varie selon les époques (Le Bart, Lefebvre, 2005 ; de manière générale, sur la contingence des ressources politiques, lire Guionnet, 2013). Concrètement, la logique de constitution de notre corpus (suivant le processus de féminisation de l'Hémicycle) implique une surreprésentation en fréquence absolue des textes de la fin de la période (les élues étant plus nombreuses qu'au début du régime) dans la vue synchronique de l'organisation thématique des discours.

de professions de foi ramené à ses quelque 300 substantifs les plus récurrents¹³. L'algorithme proposé par Reinert repose sur une série de bipartitions construites sur la base d'un tableau croisant des unités textuelles, également appelées unités de contexte élémentaire – ici, le paragraphe défini comme un segment de texte de 7 items¹⁴ – et les mots sélectionnés – les 300 substantifs lemmatisés. Une classe de vocabulaire est constituée et séparée des autres lorsque l'inertie interclasse la plus importante est atteinte. Ainsi les mots rassemblés au sein d'une même classe ont-ils un profil cooccurentiel similaire entre eux et aussi différent que possible des mots d'une autre classe (Reinert, 2008).

Menée sur le corpus de professions électorales, la classification hiérarchique descendante discrimine 5 classes¹⁵ de vocabulaire mettant au jour les structures thématiques des textes de candidature. Avant d'en décrire le contenu lexical, observons les classes dans leur organisation hiérarchique donnée par la figure 6. Le graphique place la plus importante division – c'est-à-dire hétérogénéité – entre deux groupements lexicaux séparant les classes 3 et 5, d'une part, et les classes 4, 1 et 2, d'autre part.

De manière plus incarnée, il est possible de consulter les mots qui ont contribué de manière significative à la construction des classes. La figure 7 synthétise les « profils » de chaque catégorie en détaillant les mots qui les composent. Les mondes lexicaux rassemblent plusieurs dizaines de mots ; pour la lisibilité du graphique, seuls les plus spécifiques de chaque groupement ont été reportés¹⁶. Ils suffisent souvent à pressentir le contenu thématique de la classe de vocabulaire.

La lecture du graphique informe de la teneur thématique de chaque classe de mot. Les classes 3 et 5 traitent de l'aspect politico-institutionnel du mandat brigué par les locuteurs et les locutrices, dans sa dimension nationale pour la première et dans sa dimension locale pour la seconde. Les trois autres classes distinguées sur la seconde branche principale du dendrogramme s'attachent aux affaires socio-économiques du discours électoral. Les mondes lexicaux 1 et 2 concernent les questions de société et d'économie dans les territoires ; la classe 4 embrasse des thématiques socio-économiques nationales. Complexifiée, la structuration en quatre grands pôles exposée plus haut est néanmoins confirmée, comme l'illustre la projection factorielle des classes de vocabulaire (figure 8).

La structuration selon l'axe horizontal entre un vocabulaire politique et les mots de l'économie et du social est avérée ici. À la verticale, le clivage local

13. Le corpus est réduit pour ne conserver que les 300 noms communs les plus fréquents (comme précédemment dans « corrélats »). Les verbes, adverbes, adjectifs et autres mots outils sont supprimés. L'ordre syntagmatique ou séquentiel du corpus est en revanche maintenu.

14. Empiriquement, le segment de 7 items (ou de 7 substantifs consécutifs dans le corpus vidé de toutes les autres parties du discours) apparaît le mieux adapté à l'appréhension du paragraphe, unité du traitement cooccurentiel choisie ici.

15. Le nombre de classes discriminé par le logiciel varie selon les choix de paramétrage de l'analyste. Ici, un seuil de 10 classes maximum a été imposé.

16. Dans Iramuteq, la relation spécifique, au sens statistique, d'un terme à une classe constituée est signifiée par un Chiz.

Figure 6. Classification hiérarchique descendante sur le vocabulaire du corpus « Professions de foi (1958-2007) »

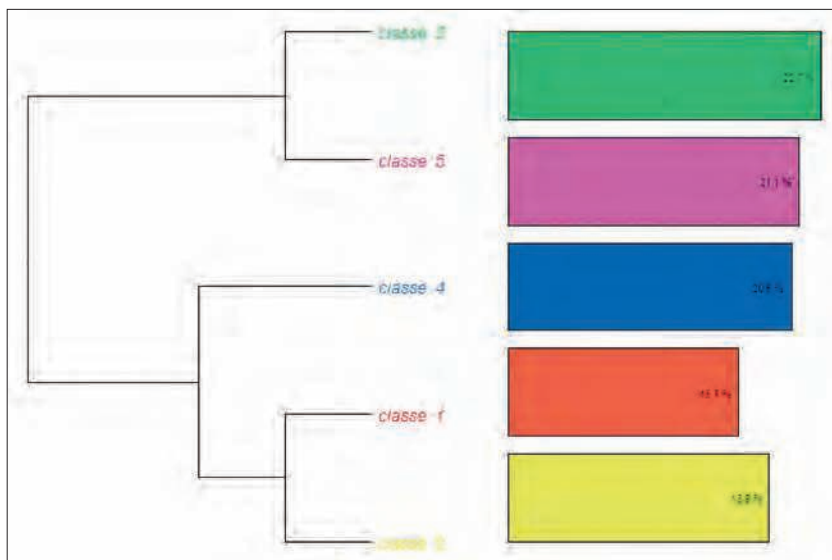
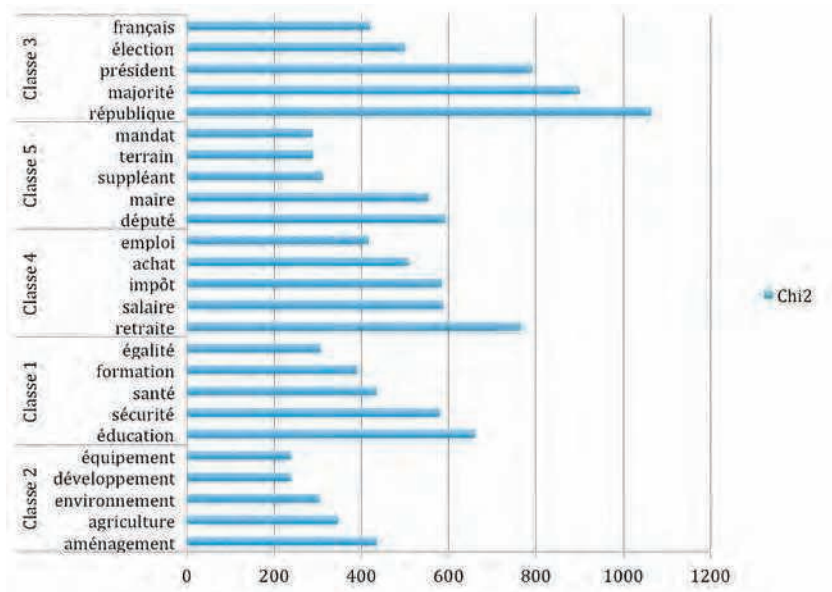


Figure 7. Substantifs les plus spécifiques des classes de vocabulaire du corpus « Professions de foi (1958-2007) »



(en bas) et national (en haut) est également opérant. La seule nuance apportée par la classification hiérarchique descendante concerne les termes économiques et sociaux. Là où deux regroupements avaient été effectués à la lecture de la figure produite par Hyperbase, Iramuteq en discrimine statistiquement trois. Les mots des propositions d'économie territoriale et les questions sociales sont répartis en deux catégories lexicales différentes (dans les classes 1 et 2) mais restent néanmoins étroitement imbriqués sur l'AFC et intimement affiliés sur le dendrogramme.

Les thèmes électoraux au féminin : rejet des mots institutionnels et poids de la socio-économie

De manière heuristique pour notre analyse, Iramuteq permet de mesurer les relations entre une variable – ici, le sexe des locuteurs – et les classes de vocabulaire. Toujours à l'aide d'un indice statistique de spécificité, le Chi2, le logiciel calcule la force de la participation des textes de chaque groupe de candidats à la thématique considérée. Alors que les deux ensembles de textes n'ont pas participé à la classification, le calcul exige néanmoins qu'ils choisissent leur camp. La figure 9 résume le degré de spécificité de chaque corpus sexué en fonction des classes lexicales.

Si la variable sexuée n'apparaît pas comme pertinente au regard de la statistique pour trois des cinq classes de substantifs (classes 1, 2 et 5), deux d'entre elles se distinguent par un lien privilégié avec l'une des modalités sexuées. En effet, alors que les textes des hommes sont significativement reliés au monde lexical de la politique nationale et institutionnelle (classe 3), les candidates surinvestissent le vocabulaire de l'économie dans leur proclamation électorale. Comme toujours dans une comparaison binaire, la distribution inverse est vraie, c'est-à-dire que les textes des femmes rejettent la classe des mots institutionnels et ceux des hommes les termes de l'économie. Bien sûr – les graphiques tirés d'Hyperbase sont là pour le rappeler – femmes et hommes évoquent ces sujets centraux dans leur profession de foi. Énoncés à partir des substantifs les plus fréquents du corpus, ces thèmes sont des indépassables du discours électoral aux scrutins législatifs. Ici, la plus-value du traitement est quantitative. L'approche du corpus par la cooccurrence généralisée tend à montrer que les hommes et les femmes construisent leur discours de candidature selon une structure commune que nous qualifierons de mixte ; mais l'investissement est divers au regard du critère chiffré¹⁷.

17. Évidemment, ces grandes tendances thématiques structurelles seraient précisées par une approche chronologique et/ou partisane ; le discours des candidats évoluant au gré des évolutions des rapports de pouvoir générés sur la scène politique ou dans la société (Achin, Bereni, 2013 ; Guionnet, 2013).

Si le moindre recours des femmes au vocabulaire politico-institutionnel est un poncif pour décrire la parole féminine (Sineau, 1988 ; Mossuz-Lavau, de Kervasdoué, 1997), le surinvestissement de la thématique socio-économique nationale du discours électoral par les candidates est moins attendu. Pourquoi des mots aussi généralistes et inévitables que *retraite*, *salaire*, *allocation*, *augmentation* ou encore *réduction*, comptés parmi les plus caractéristiques du monde lexical socio-économique, sont-ils particulièrement présents dans les énoncés féminins ?

Par ailleurs, à la consultation fine de l'ensemble des mots regroupés dans le profil de la classe 4, la présence du mot *femme* dont la spécificité est chiffrée à 5,28 attire l'attention de l'analyste. Le nom, pourtant rattaché à la classe de vocabulaire de politique locale (comme l'attestent l'AFC, figure 8, et son indice de spécificité, Chiz, dans la classe 5 chiffré à 29,6), est également lié de manière significative à la classe du vocabulaire socio-économique. La parenté cooccurrence entre les mots *femme* mais également *condition*, *cause* ou *lutte* et les mots de la thématique économique et sociale instruit le traitement original qu'en font les candidates dans leur programme. En plus des enjeux socio-économiques incontournables de chaque élection législative, les futures élues portent à l'agenda électoral des revendications spécifiques, comme en témoignent les passages suivants comprenant une cooccurrence entre *femme* et un terme appartenant à la classe socio-économique (*travail*, *allocation*, *salaires*, *augmentation*...) :

L'expansion économique française, atteignant le plus haut niveau des pays d'Europe, a permis bien des transformations [...]. Je citerai : pour les travailleurs, la mensualisation, les congés payés, la reconnaissance de la section syndicale d'entreprise, la formation professionnelle permanente. Les retraités : retraite basée sur les dix meilleures années, prise en compte des années au-delà de la trentième, augmentation du taux de pension, pensions de réversion pour les femmes à 55 ans, prise en compte pour la retraite des années passées sans travailler, à élever leurs enfants, possibilité de travail à mi-temps. (Ploux, 1973, UDV^e, Finistère)

Avec les salariés et leurs syndicats, votre députée agit pour le droit de travailler, pour une sécurité des revenus, d'emploi et de formation, pour l'augmentation du SMIG, des salaires, des retraites et des minima sociaux et pour l'égalité professionnelle entre les hommes et les femmes. (Buffet, 2007, PC, Seine Saint-Denis)

Les candidates semblent investir le thème, bien établi, de la socio-économie nationale en le chargeant – comparativement aux hommes – d'une dimension spécifique : les intérêts des femmes. Quantitativement et qualitativement, la différence de traitement des questions économiques entre les locuteurs et les locutrices tiendrait moins à l'emploi du lexique indépassable des programmes politiques (*retraite*, *salaires*, *emploi*...) qu'au développement de sous-thématiques originales par les femmes, moins à des macro-variations dans l'architecture thématique qu'à des structurations locales particulières. Il est à ce titre intéressant de noter qu'en synchronie, sous la Cinquième République,

c'est dans la perspective économique que se développent le plus les propositions « (pro)féminines »¹⁸ des futures députées.

La centralité du mot *femme* dans le corpus de candidatures féminin

Les différentes méthodes de repérage thématique mises en œuvre ci-dessus le montrent : aucune saillance thématique exclusivement féminine ne se détache du corpus de professions de foi ; en revanche, des spécificités lexicales originisent qualitativement et quantitativement les pôles thématiques structurants du discours électoral. Ainsi, la thématique socio-économique est-elle réélaboree en quantité (significative) et en qualité, en poids et en contenu, dans le corpus des futures députées par le traitement des « questions féminines » (de rapports entre les sexes, ou d'égalité).

Si, répétons-le, la parenté cooccurentielle des mots de la socio-économie nationale et des intérêts féminins est une spécificité du corpus féminin attestée par la valeur du *Chiz*, la position topographique centrale du mot *femme* dans le nuage de points « corrélats » (figure 1) ou encore la présence du terme dans plusieurs classes définies par *Iramuteq* tend à montrer le refus des futures représentantes de la nation de faire des questions féminines une thématique indépendante et particulière du discours électoral pour en faire au contraire un axe inter-thématique de leur candidature. Pour tester cette hypothèse, des logiciels de réseaux viennent compléter l'approche discriminante des thèmes menée plus haut¹⁹ pour représenter non seulement des groupements lexicaux saillants mais également les liens entre les mots qui les composent. S'il est particulièrement pertinent de penser les thèmes comme des polarisations paradigmatiques et syntagmatiques homogènes en elles-mêmes et hétérogènes par rapport aux autres, il est illusoire de postuler l'usage exclusif d'un mot dans une thématique donnée. C'est pourquoi il convient d'observer, au-delà des relations des mots au sein d'une classe, celles qui transcendent les mondes lexicaux ou les isotopies définies par la statistique cooccurentielle.

Concrètement, des logiciels de représentation de réseaux, comme *Gephi*

18. Nous choisissons d'utiliser ici l'expression « questions/propositions féminines » malgré son ambiguïté. Il s'agit de désigner des thèmes construits essentiellement autour du mot-pôle *femme* (voir *infra* : La centralité du mot *femme* dans le corpus de candidatures féminin) et concernant directement les femmes pour des raisons de sociabilisation, de rôles sociaux de sexes, de discrimination, etc. Néanmoins, rappelons avec Tremblay (1999, p. 277) que ces enjeux s'inscrivent dans des rapports de pouvoir genrés, médiatisent les rapports des femmes à l'État et sont eux-mêmes source de pouvoir entre les sexes. Ils concernent donc de fait également les hommes.
19. L'opposition entre approche discriminante et approche relationnelle ou réticulaire doit cependant être nuancée. En effet, l'AFC proposée par *Hyperbase*, notamment, établit un continuum lexical entre les thématiques et certaines positions (centrales) rendent compte de la faible teneur thématique d'un mot.

Au sein des quatre communautés qui peuvent être établies par le logiciel²⁰ sont mises en surbrillance les relations directes – c’est-à-dire les coprésences attestées dans les paragraphes du corpus – entre *femme* et les autres substantifs analysés. Si le lemme *femme* se range dans la communauté du vocabulaire de politique locale²¹ (en vert sur le réseau), il entretient de nombreux liens avec le lexique d’autres communautés. Ainsi les liens qui partent du nœud *femme* relient-ils le mot à des termes de sa communauté : *terrain*, *mandat* ou encore *expérience* mais également avec des mots de la politique nationale tels *assemblée* ou *progrès* ou encore avec les classes de la socio-économie (*salaires*, *allocation*, *cause*) et de la société (*mère*, *enfant*, *droit*, *égalité*, *famille*, *violence*)²².

Plus qu’à une thématique très structurée et polarisée, le réseau lexical de *femme* renvoie à la fois à un positionnement énonciatif (ou argumentatif, à un *ethos*) des candidates et participe à l’enrichissement des grands thèmes du discours électoral. Les relations lexicales du lemme *femme* témoignent de l’intrication dans les professions de foi entre les mots de la présentation de soi et ceux du discours programmatique et mettent en lumière l’une des spécificités majeures des candidatures de femmes aux législatives sous la Cinquième République : le lien entre présentation de soi explicitement « genrée » et la représentation des intérêts des femmes.

Les progrès récents de la logométrie dans le balisage de parcours interprétatifs cooccurentiels offrent des perspectives dans le programme ambitieux de la recherche thématique dans le discours politique. Avec le recours à la cooccurrence, la logométrie actuelle permet l’émergence inductive de polarisations lexicales, de groupements isotropiques ou encore de saillances cooccurentielles statistiquement et sémantiquement significatives.

Dans sa dimension généralisée, la statistique cooccurentielle parvient à produire des vues synthétiques des parentés et des contrastes de centaines de profils associatifs de substantifs et réussit, par là, à établir la structuration thématique de corpus importants de manière non supervisée. Par cette approche, ce sont les grands thèmes structurants et invariables aux sexes des locuteurs de la profession de foi contemporaine qui ont été distingués. Informant la représentation politique telle qu’elle s’impose sous le régime quinto-

20. Là encore le nombre final de communautés est paramétrable.

21. La projection factorielle d’Iramuteq avait également classé le mot *femme* dans la partie inférieure gauche du graphique rassemblant les mots de la classe « politique locale ».

22. Dans le corpus des candidats élus, le mot *femme* prend une position moins centrale. Les liens directs avec les mots de la politique locale sont évidemment absents. Sont attestés dans le réseau du lemme dans les discours des candidats les mots suivants : *travailleur*, *union*, *pays*, *jeune*, *jeunesse*, *vie*, *cadre*, *retraite*, *an*, *homme*, *constitution*, *progrès*, *indépendance*, *génération*, *droit*, *famille*, *égalité*. Les mots des politiques sociales du genre sont moins nombreux au profit d’une approche plutôt politique générique. Par exemple, un extrait de la profession de foi de Gaëtan Gorce (2007, PS, Nièvre) : « Le 10 juin, faites entendre votre voix, celle de femmes et d’hommes qui veulent assurer l’avenir de leurs familles dans la sécurité et la solidarité ! [...] »

républicain dans sa double dimension nationale et locale, personnelle et programmatique, les saillances thématiques repérées participent également à la caractérisation des discours. Ainsi, dans notre corpus, les passages consacrés à la socio-économie nationale se révèlent-ils quantitativement surreprésentés dans les candidatures féminines, là où la thématique institutionnelle est davantage rejetée par les députées que par leurs homologues. Enfin, sans s'en tenir à la nécessaire et informative perspective discriminante pour établir les thèmes d'un texte, les méthodes et les outils à disposition du logomètre insistent également sur l'aspect réticulaire des textes et la porosité des structururations thématiques. Il serait, en effet, caricatural d'imaginer cloisonner chacun des mots observés dans une thématique. Certains contribuent à la lexicalisation de plusieurs thèmes. C'est dans la combinaison des deux approches de repérage thématique, dichotomique et réticulaire, que se saisit l'originalité de l'organisation lexicale des candidatures des femmes.

En synchronie, la mise en exergue des thématiques du discours électoral des futures élues de la nation a mis au jour plusieurs caractéristiques de leurs promesses de représentation²³. En investissant moins que leurs homologues la dimension institutionnelle du mandat, les femmes « démasculinisent » le style parlementaire (Tremblay, 1997) ; en revanche, elles enrichissent le pôle socio-économique des problématiques « féminines » dont l'un des mots centraux, *femme* », irradie l'ensemble des textes pour informer et relier les deux pôles majeurs du discours électoral législatif : les présentations de soi et les contenus programmatiques des candidates.

Références

- ACHIN Catherine *et al.*, 2007, *Sexes, genre et politique*, Paris, Economica.
- ACHIN Catherine, BERENI Laure, 2013, *Dictionnaire Genre et Science politique. Concepts, objets, problèmes*, Paris, Presses de Sciences Po.
- ADAM Jean-Michel, 2006, « Autour du concept de texte. Pour un dialogue des disciplines de l'analyse des données textuelles », Conférence d'ouverture des JADT 2006.
- BENZECRI Jean-Paul, 1981, *Pratique de l'analyse des données. Linguistique et lexicologie*, Paris, Dunod.
- BLONDEL Vincent D. *et al.*, 2008, « Fast unfolding of communities in large networks », *Journal of Statistical Mechanics. Theory and Experiment*, n° 10 (P10008).
- BONNAFOUS Simone, 2003, « Femme politique. Une question de genre », *Réseaux*, n° 120, p. 119-145.

23. Si l'approche synchronique a le mérite de révéler les tendances structurantes des discours, elle masque ses évolutions. Or, on sait la dimension contextuelle du genre en politique. Une analyse dans une perspective chronologique devra permettre de saisir les évolutions de l'organisation thématique des professions de foi et d'affiner les constats globaux décrits ici.

- BRUNET Étienne, 2007, « Fréquences et séquences. Mise en œuvre dans Hyperbase », *Lexicométrie. Topographie et topologie textuelles*, n° 7, p. 1-20.
- 2012, « Nouveau traitement des cooccurrences dans Hyperbase », *Corpus*, n° 11, p. 219-248.
- COULOMB-GULLY Marlène éd., 2009, *Mots. Les langages du politique*, n° 90, *Présidentielles 2007. Scènes de genre*.
- DÉLOYE Yves, 1999, « Se présenter pour représenter. Enquête sur les professions de foi électorales de 1848 », *La profession politique, XIX^e-XX^e siècles*, M. Offerlé éd., Paris, Belin, p. 231-254.
- DESMARCHELIER Dominique, RENNES Juliette éd., 2005, *Mots, les langages du politique*, n° 78, *Usages politiques du genre*.
- GAXIE Daniel, 2000, *La démocratie représentative*, Paris, Montchrestien (3^e édition).
- GUIONNET Christine, 2002, « Entrées de femmes en politique. L'irréductibilité du genre à l'heure de la parité », *Politix*, n° 60, p. 113-146.
- 2013, « Fluctuation et légitimation des ressources politiques. Ce que nous apprennent les études sur le genre », *Quoi de neuf depuis la parité ? Du genre dans la construction des rôles politiques*, M. Gateau, M. Navarre, F. Schepens éd., Dijon, EUD.
- LE BART Christian, LEFEBVRE Rémy éd., 2005, *La proximité en politique. Usages, rhéoriques, pratiques*, Rennes, Presses universitaires de Rennes.
- LÉVÊQUE Sandrine, DULONG Delphine, 2002, « Une ressource contingente. Les conditions de reconversion du genre en ressource politique », *Politix*, n° 60, p. 81-111.
- LÉVÊQUE Sandrine, 2005, « Une féminité dépassée ? Usages et non-usages du genre dans les professions de foi des candidat(e)s parisien(ne)s aux élections législatives de 2002 », *Revue française de science politique*, n° 55 (3), p. 501-520.
- MANIN Bernard, 1995, *Principes du gouvernement représentatif*, Paris, Flammarion.
- MAYAFFRE Damon, 2008, « De l'occurrence à l'isotopie. Les co-occurrences en lexicométrie », *Sémantique & Syntaxe*, n° 9, p. 53-72.
- 2014, « Plaidoyer en faveur de l'Analyse de Données co(n)Textuelles. Parcours cooccurrentiels dans le discours présidentiel français (1958-2014) », *JADT 2014, Proceedings of the 12th International Conference on Textual Data Statistical Analysis*, E. Née, M. Valette, J.-M. Daube, S. Fleury éd., Paris, Inalco-Sorbonne nouvelle, p. 15-32, en ligne [<http://lexicometrica.univ-paris3.fr/jadt/jadt2014/01-ACTES/01-JADT2014.pdf>], consulté le 3 mars 2015.
- MAYAFFRE Damon, BEN HAMED Mahé, 2014, « Récits de mort et souvenir traumatique. Trames et traces lexicales des témoignages sur la Shoah », *Analyse et Argumentation du discours*, n° 13, en ligne [<http://aad.revues.org/1836>], consulté le 3 mars 2015.
- MAYAFFRE Damon, VIPREY Jean-Marie, 2012, « Introduction », *Corpus*, n° 11, *La cooccurrence. Du fait statistique au fait textuel*, p. 7-19.
- MELLET Sylvie, LONGRÉE Dominique, 2012, « Légitimité d'une unité textométrique : le motif », *JADT 2012*, A. Dister, D. Longrée, G. Purnelle éd., Bruxelles, Université de Liège / Facultés Saint-Louis, p. 715-728.
- MOSSUZ-LAVAU Janine, KERVASDOUÉ Anne (DE), 1997, *Les femmes ne sont pas des hommes comme les autres*, Paris, Odile Jacob.
- MULLER Charles, 1977, *Principes et méthodes de statistique lexicale*, Paris, Hachette.

- NEWMAN Marc E. J., GIRVAN Michelle, 2004, « Finding and evaluating community structure in networks », *Physics Reviews*, n° 69 (026113).
- PROST Antoine, 1974, *Vocabulaire des proclamations électorales de 1881, 1885, 1889*, Paris, PUF.
- RASTIER François, 2001, *Arts et sciences du texte*, Paris, PUF.
- 2011, *La mesure et le grain. Sémantique de corpus*, Paris, Champion.
- REINERT Max, 1983, « Une méthode de classification descendante hiérarchique. Application à l'analyse lexicale par contexte », *Les cahiers de l'analyse des données*, vol. VIII (2), p. 187-198.
- 1990, « Une méthodologie d'analyse des données textuelles et une application. Aurélia de Gérard de Nerval », *Bulletin de méthodologie sociologique*, n° 26, p. 24-54.
- 1993, « Les mondes lexicaux et leur logique à travers l'analyse statistique d'un corpus de récits de cauchemars », *Langage et société*, n° 66, p. 5-39.
- 2008, « Mondes lexicaux stabilisés et analyse statistique de discours », *JADT 2008*, S. Heiden, B. Pincemin éd., Lyon, PUL, p. 811-822.
- SINEAU Mariette, 1988, *Des femmes en politique*, Paris, Economica.
- 2011, *Femmes et pouvoir sous la Cinquième République. De l'exclusion à l'entrée dans la course présidentielle*, Paris, Presses de Sciences Po.
- TREMBLAY Manon, 1997, « La représentation des femmes par la voie (x) d'une démasculinisation du style parlementaire », *Femmes et représentation politique au Québec et au Canada*, M. Tremblay, C. Andrew éd., Montréal, Remue-Ménage, p. 69-100.
- 1999, *Des femmes au Parlement. Une stratégie féministe*, Montréal, Remue-Ménage.
- TRICAUD Sabrina, 2012, « Candidates et élues gaullistes au Palais Bourbon de 1958 à 1973 », *Histoire@Politique*, n° 17, p. 25-36.
- VIPREY Jean-Marie, 1997, *Dynamique du vocabulaire des Fleurs du Mal*, Paris, Champion.
- 2006, « Structure non séquentielle des textes », *Langages*, n° 163, p. 71-85.
- YU Bei, 2014, « Language and gender in Congressional speech », *Literary and Linguistic Computing*, n° 29 (1), p. 118-132.
- WODAK Ruth, 2003, « Multiple identities. The roles of female parliamentarians in the Eu Parliament », *The handbook of Language and Gender*, J. Holmes, M. Meyerhoff éd., Oxford, Blackwell Publishing, p. 671-698.