



HAL
open science

**L'approche co-occurentielle, un bond qualitatif?
L'environnement lexical du lemme “ député ” dans les
Professions de foi des candidates à la députation (1958 –
2002)**

Magali Guaresi

► **To cite this version:**

Magali Guaresi. L'approche co-occurentielle, un bond qualitatif? L'environnement lexical du lemme “ député ” dans les Professions de foi des candidates à la députation (1958 – 2002). CORELA - COgnition, REprésentation, LAngage, 2014, Complémentarité des approches qualitatives et quantitatives dans les Analyses de discours (HS 15), 10.4000/corela.3586 . hal-01362764

HAL Id: hal-01362764

<https://hal.univ-cotedazur.fr/hal-01362764>

Submitted on 16 Apr 2020

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.



Corela

Cognition, représentation, langage

HS-15 | 2014

Complémentarité des approches qualitatives et quantitatives dans l'analyse des discours

L'approche co-occurentielle, un bond qualitatif ? L'environnement lexical du lemme « député » dans les Professions de foi des candidates à la députation (1958 – 2002)

Magali Guaresi



Édition électronique

URL : <http://journals.openedition.org/corela/3586>

DOI : 10.4000/corela.3586

ISSN : 1638-573X

Éditeur

Cerle linguistique du Centre et de l'Ouest - CerLICO

Référence électronique

Magali Guaresi, « L'approche co-occurentielle, un bond qualitatif ? L'environnement lexical du lemme « député » dans les Professions de foi des candidates à la députation (1958 – 2002) », *Corela* [En ligne], HS-15 | 2014, mis en ligne le 15 octobre 2014, consulté le 02 mai 2019. URL : <http://journals.openedition.org/corela/3586> ; DOI : 10.4000/corela.3586

Ce document a été généré automatiquement le 2 mai 2019.



Corela – cognition, représentation, langage est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

L'approche co-occurrence, un bond qualitatif ? L'environnement lexical du lemme « député » dans les Professions de foi des candidates à la députation (1958 – 2002)

Magali Guaresi

Introduction

- 1 En sciences humaines, toute approche quantitative du discours est encadrée en amont - au moment de la constitution du corpus textuel - et en aval - au moment de l'interprétation du chercheur - par une démarche nécessairement qualitative. Néanmoins, si l'on veut bien admettre le postulat de la non-transparence du sens et reconnaître la plus-value interprétative, sans cesse démontrée depuis les années 1960-1970, de la lecture d'importants corpus textuels assistée par ordinateur, nous comprenons que l'enjeu n'est plus de plaider la supériorité d'une approche sur une autre mais de donner à l'analyste du discours les moyens de s'approcher au plus près d'une bonne compréhension du texte. Ces dernières années de nombreux travaux (Brunet 2007 ; Viprey 2006 ; Mayaffre 2007 ; Mellet & Barthélémy 2007 ; Longrée et al. 2008) dans le champ de l'analyse des données textuelles ont témoigné du dépassement de l'opposition désuète entre qualitativistes et quantitativistes dans une volonté commune de penser la textualité comme « la combinaison de parcours linéaires et réticulaires » (Adam, 2006). Le texte, perçu comme l'entrelacement de deux fils, est embrassé dans sa double dimension (Mayaffre, 2008): il s'agit, d'une part, de considérer le fil de chaîne ou la linéarité du texte qui rendrait compte des enchaînements de données linguistiques examinées dans leur contexte immédiat¹ et d'autre part, d'appréhender le fil de trame ou la non-linéarité du texte à travers l'édification de listes d'index de formes, hiérarchisées, détachées de leur

contexte, révélant la sélection opérée par le locuteur pour produire son discours. Historiquement attachée à une statistique fréquentielle fondée sur le décompte d'unités textuelles, l'analyse de données textuelles tente aujourd'hui de se donner les moyens d'une statistique syntagmatique ou contextualisante. L'approche co-occurrence, permise par le logiciel HYPERBASE utilisé pour l'étude, tend à incarner l'une des modalités de cette avancée qualitative. Il s'agit de dépasser l'étude de la seule fréquence – néanmoins souvent informative – de termes décontextualisés dans un corpus divisé en sous-ensembles pour aboutir à l'étude du rapport fréquentiel de deux termes présents simultanément dans une fenêtre textuelle préalablement définie². Par là, la statistique textuelle contextualisante entend traiter de paires co-occurrence, véritables « noyaux sémantiques » susceptibles de nourrir l'interprétation.

Les enjeux de la statistique co-occurrence

- 2 Nous considérons, en effet, la co-occurrence comme le contexte minimal d'un terme nécessaire à sa compréhension. Dire que deux mots sont co-occurrence revient à contextualiser l'un par l'autre³. En généralisant le propos, le profil co-occurrence d'un mot, donné par le calcul systématique de toutes ses attirances et répulsions lexicales significatives, constitue son contexte minimal formalisable objectivé par la statistique (Mayaffre, 2008b).
- 3 Pour l'analyse du discours, le sens naît en contexte. L'implémentation ancienne et complémentaire d'outils statistiques et documentaires dans les logiciels⁴ de logométrie (également appelée textométrie) témoigne de l'exigence historique de la contextualisation comme condition *sine qua non* de l'interprétation. Loin d'apporter des réponses ou des preuves, les tendances statistiques décrites par les logiciels d'analyse de données textuelles interrogent et enjoignent le chercheur à poursuivre des parcours de lecture quantitativement balisés. Mais là où l'analyste était contraint d'abandonner le soutien de l'outillage statistique pour une lecture naturelle ou qualitative – à travers la convocation de concordanciers ou le retour direct au texte – la pratique co-occurrence parvient à réintroduire les contextes dans l'analyse sans renoncer trop tôt à la plus-value d'une lecture assistée par ordinateur.
- 4 C'est bien là que réside l'atout majeur de l'approche co-occurrence : la définition de contextes porteurs de sens, relevant d'une approche qualitative, ne se soustrait pas au contrôle de la statistique. La recherche des co-occurrence d'un terme procède, en effet, d'une décontextualisation puis d'une recontextualisation : la statistique repère d'abord systématiquement les mots sur-utilisés par un locuteur donné comme autant de mots-pôles. Chaque mot-pôle, ainsi défini, est ensuite replacé dans son contexte dans lequel sont déterminées ses attirances (ou répulsions) lexicales. Mieux, le traitement co-occurrence aboutit lui-même à un effet paradigmatique puisque les co-occurrence sont classés hiérarchiquement selon le degré de significativité de leur attraction avec le mot-pôle. Le traitement co-occurrence, par ce va-et-vient entre sélection et combinaison, entre lecture paradigmatique et lecture syntagmatique, semble incarner le dépassement de l'opposition artificielle des méthodes qualitatives et quantitatives dans la perspective d'avancer toujours plus loin dans l'objectivation de parcours de lecture susceptibles d'encadrer l'acte final d'interprétation du sens. En rendant compte de la mise en résonance du vocabulaire d'un texte, de l'organisation des réseaux isotopiques et plus globalement de la textualité, les divers calculs et fonctions de représentation⁵

implémentés dans les logiciels de textométrie permettent à l'analyse de données lexicales de renouer avec sa vocation herméneutique (Mayaffre, 2002).

Un corpus de Professions de foi électorales

- 5 C'est du point de vue de l'historienne du discours politique que nous nous proposons de réfléchir à la plus-value qualitative de l'approche co-occurrence appliquée à l'analyse d'un corpus de professions de foi électorales⁶ de femmes élues députées sous la Cinquième République (1958-2002) en France. Rassemblé sur la base d'hypothèses de travail relatives au genre (Scott, 1988) en politique, cet ensemble de textes met en contraste la quasi-exhaustivité⁷ des professions de foi des élues députées avec un corpus symétrique et échantillonné de professions de foi d'hommes également élus députés. La collecte exhaustive des professions de foi des milliers d'hommes élus sous la Cinquième République étant matériellement impossible et méthodologiquement problématique⁸, deux logiques ont présidé à la constitution de l'échantillon textuel masculin. Une logique géographique d'abord : pour limiter les éventuels effets des variations discursives liées à des phénomènes géographiques, les textes de cet ensemble ont été choisis parmi les candidatures d'hommes déclarées dans chaque département ayant porté une femme à la députation. Une logique politique ensuite : pour éviter les éventuels contrastes discursifs liés à l'appartenance politique des locuteurs et des locutrices, pour chaque texte de femme élue a été collecté un texte rédigé par un homme du même bord politique. Par exemple, pour l'année 2002 en Seine Saint Denis, la profession de foi de Claude Bartolone a été retenue pour faire contraste avec la profession de foi de la candidate socialiste dans le même département, Elisabeth Guigou. L'ensemble du corpus compte 566 textes pour plus de 370 000 occurrences.
- 6 Présentations de soi et textes programmatiques, les professions de foi sont un lieu d'observation privilégié des rapports qu'entretiennent le genre et la politique en France. Acte discursif fondateur de la candidature, la profession de foi ouvre la campagne et transforme, dans un mouvement performatif, le locuteur ou la locutrice en candidat-e officiel-le. A travers les justifications invoquées pour expliquer l'engagement électoral ou les modalités de présentation de soi, elles révèlent en creux la promesse de représentation que les candidat-e-s formulent (Deloye, 1999). Pour tenter d'appréhender les éventuelles modalités genrées de la construction d'identités politiques dans notre corpus, nous nous concentrerons sur l'étude du lemme « député » (658 occurrences dans le corpus global) qui, on le devine aisément, est l'un des mots-pôles autour desquels les candidat-e-s énoncent les représentations qu'ils se font de leurs rôles parlementaires (Lagroye, 1994)⁹. Cette contribution a un objectif double : il s'agira de cerner les spécificités discursives de la construction des rôles politiques au féminin tout en explicitant les parcours de lecture dessinés par l'approche co-occurrence.

1. Une approche fréquentielle limitée du mot « députée »

- 7 Si l'approche occurrence classique en analyse de données textuelles est le plus souvent informative, elle touche parfois à ses limites. C'est le cas pour l'analyse du mot « député » dans ses déclinaisons féminine et masculine. Présent en moyenne 1,2 fois dans chaque profession de foi, le mot désignant la fonction politique à laquelle prétendent les

candidat-e-s parcourt l'ensemble du corpus à proportion quasi égale entre les textes de femmes et ceux des hommes. Si le décompte du mot « député » permet de conclure au moins provisoirement à une légère sur-utilisation du terme dans les textes de candidats (Figure 1), il ne permet pas de mener plus loin l'interprétation. Le calcul des spécificités¹⁰, outil historique de la statistique textuelle, nous informe néanmoins davantage sur la distribution du mot « députée » dans sa forme féminine. Avec un écart réduit chiffré à +8,9, il apparaît dans les premiers rangs des termes spécifiques du corpus de professions de foi rédigées par les candidates (Figure 1).

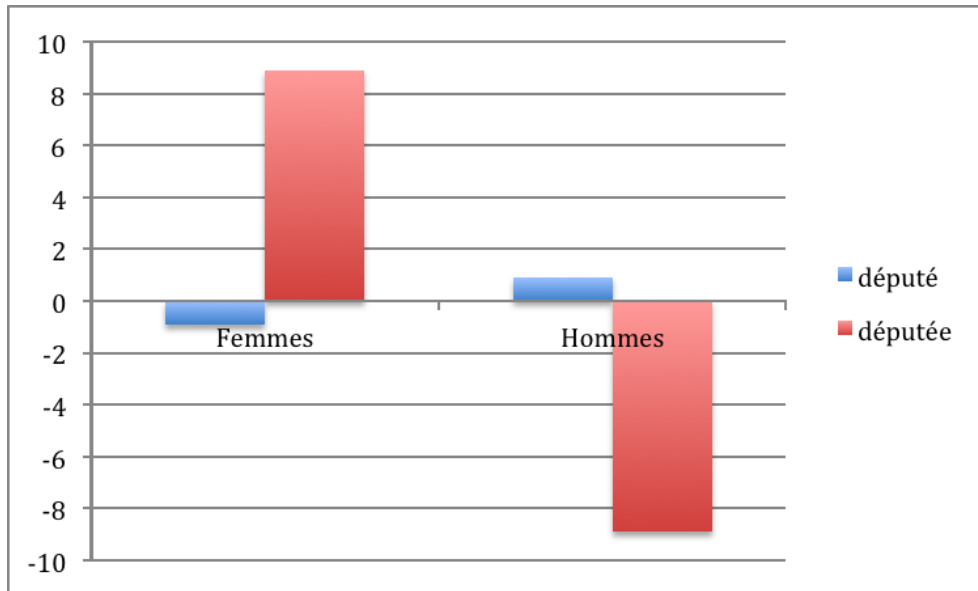


Figure 1 : Distribution des mots « député » et « députée » dans les Professions de foi des candidates et des candidats aux élections législatives (1958-2002)

- 8 Une sur-utilisation bien « naturelle » du point de vue grammatical puisque la présence du « e » final, marque linguistique du féminin, témoigne d'un emploi quasi exclusif dans les textes des candidates qui s'auto-désignent comme prétendantes à un mandat de « députée »¹¹. Au contraire, le mot « député » dans sa forme masculine n'est pas attesté dans la liste des termes repérés comme spécifiques du corpus masculin bien qu'il apparaisse à 310 reprises sous la plume des candidats. Cette asymétrie s'explique par l'usage presque généralisé que font les candidates à la députation de la forme masculine du mot « député » jusqu'au début des années 1990 (cf. infra 3.1).
- 9 Pour achever la description fréquentielle du lemme « député », précisons que l'index des formes d'HYPERBASE répertorie 91 occurrences de « députés » au pluriel. 49 sont comptés dans les textes des femmes ; 42 dans les textes de leurs homologues. Là encore, l'approche occurrence classique ne permet pas de s'avancer plus loin dans la compréhension de l'édification des identités politiques des hommes et des femmes candidat-e-s à l'obtention d'un siège au Palais-Bourbon. Au contraire, en ne révélant pas de saillances quantitatives heuristiques, cette lecture conduit même à masquer l'intérêt de l'étude du lemme « député ». Dès lors, la compréhension des modalités de construction d'une identité politique au féminin à partir du mot « député-e » semble nécessairement passer par l'analyse contrôlée de ses contextes d'emploi.

2. Contextualisation minimale du lemme « député » dans les textes de candidates et de candidats

- 10 HYPERBASE propose un premier outil, très précieux et éprouvé depuis longtemps, appelé « fonction Thème » (Lebart & Salem, 1994). Il s'agit d'extraire du corpus global de manière systématique les paragraphes contenant le mot-pôle demandé pour les reconstituer en sous-corpus. L'index fréquentiel établi dans le sous-ensemble est comparé au dictionnaire des fréquences du corpus entier. HYPERBASE associe alors à chaque co-occurrence un indice de spécificité indiquant la force d'attraction avec le mot choisi pour pôle. Cette opération est successivement réalisée pour les contextes de « député » dans le corpus des textes de candidates et dans le corpus de textes des candidats. Deux listes sont soumises à l'analyse du chercheur (Tableau 1).

2.1. L'élection, acte fondateur de la représentation politique

- 11 Si quelques mots sont communs aux environnements lexicaux du mot « député » dans les deux sous-corpus étudiés, un certain nombre de co-occurrences différencient nettement les deux profils co-occurrence :

Co-occurrences de « député » (Corpus Femmes)				Co-occurrences de « député » (Corpus Hommes)			
votre	11,75	fidèle	3,38	votre	37,58	communiste	3,95
sortant	11,07	engage	3,32	circonscription	11,61	mandat	3,8
élire	10,16	être	3,22	sortant	11,11	avez	3,78
allez	7,41	action	3,13	élire	10,7	envoyer	3,76
circonscription	6,81	dévouement	3,11	depuis	6,28	conseiller	3,74
élue	6,55	dynamique	3,04	choisissez	6,22	présent	3,44
mandat	5,05	efficace	3,02	élu	6,04	législature	3,41
actif	4,72	suppléant	2,98	élirez	5,65	nationale	3,29
choisissez	4,62	suffrages	2,92	appelés	5,44	voté	3,293,35
assemblée	4,38	nationale	2,87	assemblée	5,34	nouveau	3,11
disponible	4,33	proche	2,84	serai	5,26	prochain	3,09
appelés	4,31	présent	2,84	ai	5,04	désigner	3,08
serai	4,26	projets	2,79	maire	4,87	proche	2,98
représentera	3,99	voter	2,73	êtes	4,75	honneur	2,86
lois	3,97	représenter	2,55	actif	4,51	UNR	2,71
maire	3,88	écoute	2,51	réélire	4,47	conseil	2,65
femme	3,48	habitants	2,43	allez	4,19	terrain	

Tableau 1 : Co-occurrences de « député » dans les Professions de foi de candidates et de candidats à la députation

- 12 Le tableau montre que les douze co-occurrences privilégiées du mot « député » dans le corpus des textes de femmes figurent en bonne position dans la liste des termes associés au même mot dans le sous-ensemble masculin. L'environnement lexical du mot « député » commun aux locutrices et aux locuteurs correspond à un vocabulaire institutionnel nécessairement présent dans un discours de campagne visant à l'élection. « Elire », « assemblée », « loi », « suffrages », « législature », ou encore « suppléant » sont autant de mots articulés à la référence au mandat politique brigué dans la course électorale. Ce vocabulaire traduit, en même temps qu'il perpétue, l'importance fondatrice du vote dans la représentation politique dans nos démocraties contemporaines (Rosanvallon, 1992). Être « député », c'est avant tout être « élu(e) ». Aussi, les candidats et les candidates, de manière indifférenciée et dans un souci pédagogique, commencent-ils leur profession de foi en rappelant l'enjeu institutionnel des élections législatives :

« Le 5 mars, vous allez élire le député qui vous représentera à l'assemblée nationale. » (Jacques Chambaz, 1967)

- 13 Le moment électoral est l'expression d'une promesse de représentation du candidat ou de la candidate à son électorat. Conformément à loi, la représentation parlementaire décrite autour du mot « député » dans les deux corpus articule les deux dimensions du mandat : la représentation de la nation (« assemblée », « nationale », « représenter ») et l'ancrage local (« circonscription », « maire », « conseiller », « commune », « permanences »). Cette double dimension représentative présente dans le discours électoral témoigne des prescriptions institutionnelles qui contraignent l'exercice d'une charge politique. Surtout, l'institutionnel semble informer littéralement le discours en lui imposant les thématiques incontournables de la représentation parlementaire : les prétendant-e-s à la députation – quel que soit leur sexe – sont contraint-e-s de démontrer leurs atouts aux deux niveaux de la représentation parlementaire puisque, de fait, leur action, une fois élu-e se partagera entre leur circonscription d'élection et l'Assemblée nationale (Costa et Kerrouche, 2007). Les verbes conjugués à la seconde personne du pluriel et/ou au futur, également présents dans les deux environnements lexicaux décrits dans le tableau 1, expriment encore la place centrale du vote dans la délégation de souveraineté du peuple à ses représentant-e-s. Ces verbes exhortent les électeurs au vote en faveur du locuteur-candidat ou de la locutrice-candidate. Les électeurs sont directement interpellés, dans un souci phatique et perlocutoire, par l'usage de la seconde personne du pluriel :

« Le 25 mai, vous irez voter et vous choisirez votre député » (Hélène Mignon, 1997)

2.2. Député et « femme député »

- 14 Malgré un stock de mots qui semble indépasseable lorsque un-e candidat-e évoque le mandat de « député », les listes des co-occurrences du mot « député » dans les deux sous-corpus ne se recoupent que partiellement. La confrontation de la liste des co-occurrences de « député » dans les textes des femmes et dans les textes des hommes donnée plus haut (Tableau 1) avec le graphique ci-dessous (Figure 2) représentant les liens d'attractions entre le mot « député » et ses différents co-occurrences dans l'ensemble du corpus révèle les spécificités de la production discursive des candidates.

2.3. « Députée » et féminisation des rôles politiques

- 20 La fonction « Thème », de nouveau appliquée au corpus de textes de femmes, autour du mot-pôle féminin « députée » nous conforte dans l'hypothèse de la pertinence de la variable « genre » dans la construction des rôles parlementaires dans les Proclamations électorales. La liste des co-occurents statistiquement significatifs de « députée » (Tableau 2) se distingue nettement de la liste des co-occurents de « député » dans le corpus masculin et partiellement de celle des co-occurents de « député » dans le corpus féminin.

être	7,13	candidature	2,9
votre	6,02	déterminée	2,77
serai	4,19	avez	2,73
comme	4,15	présente	2,67
sortante	4,14	préoccupation	2,6
servir	4,11	mandat	2,6
compte	3,86	accordez	2,57
proche	3,85	honneur	2,54
efforcée	3,77	contribuer	2,54
élue	3,73	service	2,5
devenue	3,52	côtés	2,5
engagement	3,4	écouter	2,41
prouvé	3,39	efficace	2,21
compétente	3,19	tâche	2,16
circonscription	3,19	convictions	2,14
aurai	3,19	combat	2,14
depuis	3,12	femme	2,08
entièrement	2,98	socialiste	2,07
attentive	2,98		

Tableau 2 : Co-occurents du mot « députée » dans le corpus des professions de foi de candidates (1958-2002)

- 21 Après la présence presque inévitable de mots que l'on devine composer des phrases telles que « Je serai votre députée », la liste des termes associés à « députée » se singularise par une plus grande rareté du vocabulaire institutionnel et une présence plus importante d'unités linguistiques faisant référence à des activités de soin à l'égard de l'électorat et plus généralement à l'éthique du *care* théorisée par la psychologue américaine Carol Gilligan dans son ouvrage, *Une voix différente. Pour une éthique du care* (2008). Pour reprendre les mots de Bernadette Isaac-Sibille, candidate en 1997, la représentation politique ne se limiterait pas à « s'occuper des électeurs uniquement au moment des élections ». Elle semble plutôt s'abstraire du seul cadre des scansion électorales pour se fonder sur un lien tissé et entretenu au fil du mandat. Ce lien se construit d'abord sur l'affirmation de dispositions à l'attention comme la reconnaissance d'un besoin : les

candidates se disent « attentive(s) », pétries de « préoccupations » à l'égard de leurs représenté-e-s qui se voient reconnaître une part de vulnérabilité.

« Etre députée, c'est rencontrer, informer et écouter les citoyens pour faire vivre la démocratie. Avec votre soutien, je serai disponible, assidue, attentive, proche de vos préoccupations » (Catherine Picard, 1997)

« Députée, je serai toujours proche de vous, attentive à vos problèmes et j'aurai la même énergie pour me battre pour vous, à vos côtés, pour améliorer la vie de chacun et porter vos préoccupations au plus haut niveau ». (Bernadette Isaac-Sibille, 1997)

- 22 Un ensemble de verbes exprime le mandat de « députée » comme un service (« servir », « contribuer », « écouter ») à la population motivé par une implication personnelle forte (« engagement », « convictions », « entièrement », « déterminée », « présente », « combat »). Plus qu'un service, le mandat parlementaire est associé aux mots « tâche » ou « efforcée » qui lui confèrent la dimension d'une charge dont les candidates se font une obligation, presque un devoir auquel elles ne pourraient pas échapper et qui justifierait leur présence dans la sphère politique. L'adjectif « compétente » ou encore le participe passé « prouvé » – qui n'ont pas d'équivalent dans la liste des co-occurents de « député » dans les professions de foi d'hommes – contribuent à la construction d'une image d'élue en capacité d'assumer ses fonctions. Sans cesse soupçonnées d'usurpation et d'incompétence, les femmes soulignent davantage que leurs homologues masculins, leurs compétences dans la construction discursive de leur légitimité.

« Votez pour une députée dynamique, compétente, femme d'action, femme de terrain très proche de vous » (Catherine Genisson, 1997)

- 23 Comme en témoignent les extraits de professions électorales citées plus haut, les femmes construisent une identité politique tournée vers les trois aspects fondamentaux – et étroitement liés dans le discours produit autour du mot féminin « députée » – de l'éthique du *care* : la reconnaissance du besoin de l'autre, une mise au service des électeurs, l'affirmation de compétences spécifiques pour le mandat convoité.

2.4. Co-occurrences généralisées autour du lemme « député » dans le corpus des candidates

- 24 L'application du troisième outil proposé par les logiciels de textométrie, et en particulier par HYPERBASE – le calcul des co-occurrences généralisées d'après le modèle exposé par Jean-Marie Viprey (1997) – donne à voir l'organisation des fréquentations lexicales des co-occurents du lemme « député » (Figure 3) dans le corpus de candidates. HYPERBASE établit une matrice carrée qui croise tous les co-occurents du lemme « député » entre eux. Le graphique d'analyse factorielle des correspondances (AFC) ainsi produit représente le résultat des comparaisons des profils co-occurrence de chacune des unités lexicales recensées dans la matrice. Plus deux vocables ont un profil associatif semblable, plus leurs points seront proches sur la cartographie AFC. A partir du constat statistique de la coprésence de deux unités dans un paragraphe, l'AFC témoigne de l'organisation textuelle du corpus et ouvre la voie à l'interprétation sémantique. Appliquée à notre corpus, la représentation AFC de l'attraction généralisée de tous les co-occurents du lemme « député » distingue deux grands types d'emploi. A gauche de l'axe horizontal, un vocabulaire institutionnel combinant verbes électoraux et désignations des électeurs et des électrices se tient à distance du centre du graphique. A droite du graphique, un vocabulaire plus dense s'articule autour du mot « députée » au féminin,

atteindre les 10%. Dans cette « République des mâles » (Sineau, 2011), les rôles politiques, à travers les noms des titres et des charges, ne se déclinaient qu'au masculin, prétendument neutre. Les candidates à la députation en situation de minorité numérique adoptaient en même temps que le titre, les codes parlementaires dominants. La consultation des index de fréquences du corpus indique qu'avant les années 1980, aucune occurrence de la forme féminine de « députée » n'est attestée dans le corpus. Ce constat quantitatif confirme les études menées sur la féminisation des noms de métiers qui concluaient à un début de mouvement de féminisation « spontanée » au début des années 1980, après l'ouverture du premier débat public sur la question en 1983 (Fujimura, 2005). Dans notre corpus, deux emplois du mot féminin sont attestés dans les professions de foi de candidates socialistes, Edwige Avice et Véronique Neiertz ; le nom est féminisé dans les mêmes proportions en 1988 par Huguette Bouchardeau et Nicole Cacheux et en 1993 par Ségolène Royal et Huguette Bello. Il faut attendre 1997 pour que 21 occurrences soient comptées et 2002 pour qu'un tiers environ des élues l'emploie (38 occurrences).

- 27 Le mot « députée » caractérise donc particulièrement les deux dernières élections législatives étudiées, qui marquent du point de vue de la question de la représentation des femmes en politique, deux évolutions historiques¹³. Ces élections sont d'abord le théâtre d'une féminisation sans précédent de l'hémicycle : plus de 60 femmes sont élues en 1997 et 107 en 2002. Pour la première fois, la barre symbolique des 10% de femmes au Palais-Bourbon est dépassée. La seconde évolution, d'une certaine manière cause de la première, est la reconnaissance dans le corps de la Constitution d'une nation souveraine composée d'hommes et de femmes qui se voient conférer un égal accès aux charges électives. Etre une femme devient, dans certaines circonscriptions, un enjeu politique et partisan et une ressource stratégique. Les politologues constatent alors une valorisation des qualités dites féminines dans la course électorale et émettent l'hypothèse du « salto du stigmate » (Paoletti & Achin, 2002) pour questionner ce renversement.

3.2. Sexe de la langue et représentations politiques

- 28 Les différences contextuelles mesurées par HYPERBASE révèlent le poids du « sexe » de la langue dans la représentation du mandat de député-e, étant entendu que la langue est à la fois le symptôme des rapports de pouvoir et l'un des moyens de sa mise en œuvre (Chetcuti, 2012). Dans notre corpus électoral, les candidates qui utilisent le masculin-générique « député » projettent une image de la représentation parlementaire conforme, en partie au moins, aux codes masculins qui y correspondent. Mais si elles ne féminisent pas le nom du mandat qu'elles briguent, elles lui attribuent néanmoins une charge sémantique différente par la référence plus ou moins explicite à leur sexe.
- 29 Les modifications sémantico-politiques apportées au mandat législatif sont encore plus importantes lorsque les candidates marquent leur sexe dans la langue à travers la féminisation grammaticale du mot « députée »¹⁴. Ce n'est pas seulement le titre de la fonction qui est féminisé mais l'ensemble des rôles politiques auxquels il fait référence. Les candidates qui évoquent leur futur mandat au féminin déploient, de manière plus ou moins stratégique, une identité politique procédant de stéréotypes féminins. Ce positionnement discursif témoigne à la fois des contraintes qui pèsent sur la construction de la légitimité des candidates et de la diffusion de conceptions et valeurs nouvelles dans le champ politique. Pour être tolérées dans l'espace politique, les candidates doivent faire montre de leur féminité préservée, notamment dans le cas qui nous intéresse en

décrivant des qualités traditionnellement reconnues aux femmes. Dans le contexte de la crise de la représentation politique telle qu'elle s'est diffusée comme représentation sociale dominante dans l'espace public, « faire de la politique autrement » (c'est-à-dire différemment des hommes qui incarnent la norme) est l'une des attentes formulées par les appareils partisans à l'endroit des femmes politiques. Depuis la réforme paritaire, les femmes sont recrutées comme candidates titulaires aux élections législatives par leur parti respectif dans l'espoir d'une « plus-value de la différence » (Sénac-Slawinski, 2012). Etre, dire et faire différemment des hommes politiques est une des conditions de leur recrutement, les candidates sont donc (plus ou moins) contraintes d'élaborer une identité politique, dans une visée rhétorico-pragmatique, en adéquation avec ces attentes. La représentation politique relationnelle fondée sur une approche empathique, proche et concrète de la réalité des citoyen-ne-s érige les qualités que les femmes auraient acquises dans le champ domestique au rang de ressources dans le discours électoral. Les candidates se posent en intermédiaires entre la sphère publique et la sphère privée, entre l'Etat et la société civile. Cette éthique du *care* dont les femmes seraient porteuses – et qu'elles auraient développée dans la famille ou au fil de leurs expériences de vie spécifiques –, est présentée comme un des moyens appelés à résorber le fossé entre le corps électoral et les élu-e-s. Une position discursive qui ne va pas sans risquer de réactiver une vision essentialiste des femmes ou d'enfermer les candidates dans une image sexuée.

- 30 Alors que le débat électoral se dépolitise au fil de la Cinquième République, le discours se recentre sur les impétrant-e-s et leur conception du lien de représentation. A ce titre, la féminisation grammaticale de la fonction législative peut se comprendre comme une volonté militante et/ou stratégique de créer de nouvelles modalités d'incarnation des rôles parlementaires. Elle apparaît comme un catalyseur d'évolutions de l'ordre politique à un moment où la démocratie représentative est en question.

Conclusion

- 31 En permettant une étude du vocabulaire en contexte et du maillage textuel, l'approche co-occurrence assistée par ordinateur constitue une avancée qualitative. Cette pratique d'essence contextualisante apparaît comme le moyen de s'avancer plus loin dans des parcours de lecture contrôlés, susceptibles d'orienter l'interprétation du chercheur. Par la co-occurrence, l'analyse de données textuelles quitte la lexicographie fondée sur une approche décontextualisée des unités textuelles pour une science lexicologique du vocabulaire en usage et en corpus (Mayaffre, 2008b). L'analyse exhaustive et contrôlée des contextes minimaux d'emploi du lemme « député » dans le corpus électoral, par des allers-retours entre sélection et combinaison, entre décontextualisation et recontextualisation, entre lecture quantitative et lecture qualitative, a permis d'approcher les modalités de la construction genrée des identités parlementaires. C'est précisément la mixité des outils proposés par HYPERBASE et leur constant dialogue qui a permis la compréhension de phénomènes discursifs a priori difficilement observables.
- 32 Concluons l'étude de cas en remarquant que la féminisation du nom de la charge parlementaire contribue à la diffusion de valeurs jusque là radicalement éloignées du champ politique. Plus que la féminisation en nombre de l'hémicycle, c'est la féminisation en genre (grammatical) dans le discours politique qui semble opérer un renouvellement des contours de la représentation politique.

BIBLIOGRAPHIE

- Achin C., Bargel L., Dulong D. et al., 2007. *Sexes, genre et politique*. Paris : Economica.
- Adam J.-M., 2006. *Autour du concept de texte. Pour un dialogue des disciplines de l'analyse des données textuelles*, Conférence plénière, JADT 2006. Besançon, 19-21 avril.
- Amossy R., 1999. *Images de soi dans le discours. La construction de l'ethos*. Lausanne : Delachaux et Niestlé.
- Ben Hamed M. & Mayaffre D., 2012. Saisir le sens dans les deux sens : Exploration de la portée interprétative de l'énergie et de la disponibilité. *JADT 2012*, Liège. Accessible à l'URL : <http://lexicometrica.univ-paris3.fr/jadt/jadt2012/Communications/Ben%20Hamed,%20Mahe%20et%20al.%20-%20Saisir%20le%20sens%20dans%20les%20deux%20sens.pdf>
- Bonneau J., 2012. La co-occurrence asymétrique : propriétés quantitatives en disponibilité et énergie. *JADT 2012*, Liège. Accessible à l'URL : <http://lexicometrica.univ-paris3.fr/jadt/jadt2012/Communications/Bonneau,%20Julien%20-%20La%20cooccurrence%20asymetrique.pdf>
- Brunet E., 2007. Séquences et fréquences. Mises en œuvre dans Hyperbase. *Lexicometrica 7, Topographie et topologie textuelles*. Accessible à l'URL : <http://www.cavi.univ-paris3.fr/lexicometrica/numspeciaux/special9/brunet.pdf>
- Costa O. & Kerrouche E., 2007. *Qui sont les députés français. Enquête sur une élite inconnue*. Paris : Presses de Sciences po.
- Deloye Y., 1999. Se présenter pour représenter. In Offerlé, J. (dir.), *La profession politique*. Paris : Belin.
- Gilligan C., 2008. *Une voix différente. Pour une éthique du care*. Paris : Flammarion.
- Heiden S., 2004. Interface hypertextuelle à un espace de cooccurrences : implémentation dans Weblex. *JADT04, Le poids des Mots*, Louvain. Accessible à l'URL : http://lexicometrica.univ-paris3.fr/jadt/jadt2004/pdf/JADT_055.pdf
- Fujimura I., 2005. La féminisation des noms de métiers et des titres dans la presse française (1988-2001). *Mots. Les langages du politique*, 78, 37-52.
- Lagroye J., 1994. Etre du métier. *Politix*, 7, 5-15.
- Lebart L. & Salem A., 1994. *Statistique textuelle*. Paris : Dunod.
- Longrée D., Luong X. & Mellet S., 2008. Les motifs : un outils pour la caractérisation topologique des textes. *Actes des JADT 2008*, Lyon. Accessible à l'URL : <http://lexicometrica.univ-paris3.fr/jadt/jadt2008/pdf/longree-luong-mellet.pdf>
- Luong X., Brunet E., Longrée D., Mayaffre D., Mellet S. & Poudat C., 2010. La cooccurrence, une relation asymétrique ? *Actes des JADT 2010*, Milan.
- Mayaffre D., 2002. L'Herméneutique numérique, *L'Astrolabe. Recherche littéraire et Informatique*. Accessible à l'URL : http://hal.archives-ouvertes.fr/docs/00/58/65/12/PDF/11_Mayaffre_Astrolabe_2002.pdf

- Mayaffre D., 2007. L'analyse de données textuelles aujourd'hui : du corpus comme une urne au corpus comme un plan. Retour sur les travaux actuels de topographie/topologie textuelle. *Lexicometrica 7, Topographie et topologie textuelles*. Accessible à l'URL : <http://lexicometrica.univ-paris3.fr/numspeciaux/special9/mayaffre.pdf>
- Mayaffre D., 2008a. Quand travail, famille, patrie co-occurent dans le discours de Nicolas Sarkozy. Etude de cas et réflexion sur la co-occurrence. *JADT 2008*. Accessible en ligne à l'URL : <http://lexicometrica.univ-paris3.fr/jadt/jadt2008/pdf/mayaffre.pdf>
- Mayaffre D., 2008b. De l'occurrence à l'isotopie. Les co-occurrences en lexicométrie. *Syntaxe et Sémantique. Textes, documents numériques, corpus...*, 9, 53-72.
- Mellet S. & Barthelemy J.-P., 2007. La topologie textuelle : légitimation d'une notion. *Lexicometrica 7, Topographie et topologie textuelles*. Accessible à l'URL : <http://lexicometrica.univ-paris3.fr/numspeciaux/special9/mellet.pdf>
- Paoletti M. & Achin C., 2002. Le salto du stigmaté. Genre et construction des listes aux municipales de 2001. *Politix*, 60 (15), 33-54.
- Rosanvallon P., 1992. *Le sacre du citoyen : histoire du suffrage universel en France*. Paris : Gallimard.
- Scott J. W., 1988. Genre : une catégorie utile d'analyse historique. *Les Cahiers du GRIF*, 37-38, 125-153.
- Sénac-Slawinski R., 2012. *L'Invention de la diversité*. Paris : PUF.
- Sineau M., 2011. *Femmes et pouvoir. De l'exclusion à la course pour la présidentielle*. Paris : Presses de Sciences po.
- Viprey J.-M., 1997. *Dynamique du vocabulaire des « Fleurs du Mal »*. Paris : H. Champion.
- Viprey J.-M., 2006. Structure non séquentielle des textes. *Langages*, 163, 71-85.

NOTES

1. Notons dès à présent que le mot « contexte » sera utilisé dans cet article pour désigner le contexte linguistique immédiat d'une unité textuelle, également appelé cotexte (le syntagme, la phrase, le paragraphe...). Il est le lieu de la formation du sens ; son appréhension est la condition de l'interprétation.
2. Sans prétendre clore les discussions sur la taille de l'empan co-textuel le plus adapté pour un traitement co-occurrence, nous indiquons que la fenêtre textuelle retenue pour cette étude est le paragraphe. 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 qu'imposerait le choix d'une fenêtre d'étude plus courte.
3. Sans rentrer dans les détails, nous mentionnons l'apport de travaux récents sur la question de l'asymétrie des relations co-occurrence (Luong et al. 2010 ; Ben Hamed & Mayaffre 2012 ; Bonneau, 2012)
4. HYPERBASE, Lexico 3, TXM, pour ne citer qu'eux, permettent un retour organisé au texte. Précisons que seul HYPERBASE (E. Brunet, BCL, Université de Nice Sophia-Antipolis) a été utilisé pour cette étude.
5. 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 co-occurent dans le texte et k = co-occurrence observée. Alors : $\text{Prob}(x=k) = (f! (s+g)! g! (f+s)!) / (k! (f-k)! (g-k)! (s+k)! (f+g+s)!)$

6. Les professions de foi sont les documents électoraux adressés aux électeurs à la veille des élections. Les candidat(e)s s'y présentent et exposent leur programme politique.
7. L'élection de 1986 s'est déroulée de manière exceptionnelle, sous la Cinquième République, au scrutin de liste. Les professions de foi de cette élection ont donc été écartées en raison de leur rédaction collective.
8. La différence de taille entre le corpus féminin et l'écrasant corpus masculin aurait interdit un traitement statistique fiable par les logiciels de textométrie.
9. Nous évoquerons les rôles politiques comme le résultat des prescriptions institutionnelles et socio-politiques dont ils sont l'objet et des usages qu'en font ceux qui les endossent.
10. Une spécificité positive (ou négative) est un terme sur-utilisé (ou sous-utilisé) par un locuteur au regard d'une norme d'utilisation endogène au corpus global. Elle est indiquée par une valeur positive (ou négative) nommée « écart-réduit ».
11. Les candidats n'emploient le mot « députée » dans sa forme féminine qu'à deux reprises pour évoquer la « députée sortante » ou pour évoquer le danger que constituerait l'élection d'une adversaire, « une députée socialiste » ; les impétrantes quant à elles l'utilisent 65 fois pour s'auto-désigner.
12. Nous considérerons l'*ethos* comme l'ensemble des modalités du dire contribuant à la construction d'une image du locuteur ou de la locutrice. Pour un développement des aspects épistémologiques de la notion, nous renvoyons le lecteur à la consultation de l'ouvrage de R. Amossy (1999).
13. Précisons que la chronologie n'explique que partiellement les variations discursives constatées autour des mots « député » et « députée » puisque le calcul du profil co-occurrence de « député » dans le corpus masculin réduit à un empan chronologique restreint (1997-2002) reste très proche du profil co-occurrence du mot pour l'ensemble de la période (Figure 1).
14. La féminisation des noms de métiers ou des charges politiques reste un enjeu actuel si l'on en croit les récents propos de Cécile Duflot, à l'Assemblée nationale, demandant au député Bernard Accoyer de la nommer « Madame LA ministre » (*Le Monde*, 27/09/2012).

RÉSUMÉS

En définissant de manière contrôlée par la statistique des contextes minimaux porteurs de sens, l'approche co-occurrence incarne une avancée qualitative dans le champ de l'analyse des données textuelles. Par là, l'ADT renoue avec sa vocation herméneutique. Cette contribution montrera que l'analyse en contexte du lemme « député » dans les discours électoraux de candidat-e-s à la députation permet à l'historien-ne d'accéder à la compréhension sémantique et historique des modalités de la représentation parlementaire des femmes sous la V^e République.

Cooccurrence stands as a qualitative progress in the field of textual data analysis (ADT) since it statistically determines minimal meaningful contexts. Hence, ADT reasserts its hermeneutic perspective. The paper analyses the contextualization of the lemma « député » in the candidates statements for parliamentary elections during the Vth Republic. Results give historians a semantic and historical understanding of women parliamentary representation.

INDEX

Mots-clés : Analyse de données textuelles, Co-occurrences, Contextualisation, Corpus politique, Député-e.

Keywords : Textual data analysis, Co-occurrences, Contextualization, Political discourse, (Women) Members of parliament.

AUTEUR

MAGALI GUARESÌ

BCL (UMR 7320) Université de Nice-Sophia Antipolis