
Grégory Piet, Régis Dandoy et Jeroen Joly

Comprendre le contenu des programmes électoraux. Comparaison des méthodes d'encodage manuel et automatique

Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

revues.org

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

Référence électronique

Grégory Piet, Régis Dandoy et Jeroen Joly, « Comprendre le contenu des programmes électoraux. Comparaison des méthodes d'encodage manuel et automatique », *Mots. Les langages du politique* [En ligne], 108 | 2015, mis en ligne le 06 octobre 2017, consulté le 11 septembre 2015. URL : <http://mots.revues.org/22020> ; DOI : 10.4000/mots.22020

Éditeur : ENS Éditions

<http://mots.revues.org>

<http://www.revues.org>

Document accessible en ligne sur : <http://mots.revues.org/22020>

Ce document est le fac-similé de l'édition papier.

Cet article a été téléchargé sur le portail Cairn (<http://www.cairn.info>).



Distribution électronique Cairn pour ENS Éditions et pour Revues.org (Centre pour l'édition électronique ouverte)

© ENS Éditions

Comprendre le contenu des programmes électoraux. Comparaison des méthodes d’encodage manuel et automatique

Comme le souligne Philippe Braud, le « débat d’idées permanent qui oppose les différents acteurs de la vie politique demeure un objet central de la science politique » (2012). Dans le champ d’étude des partis politiques, ce débat nous permet de mettre en exergue la compétition partisane à laquelle se livrent ces acteurs politiques, leurs positionnements sur les nombreux enjeux de politique publique, leurs intérêts politiques, leurs priorités politiques, etc. Ces luttes pour la conquête du pouvoir sont d’autant plus intéressantes à analyser en période électorale qu’elles mettent aux prises notamment des acteurs politiques par programmes électoraux interposés.

Toutefois, comment pouvons-nous faire émerger de ces textes parfois très longs (plusieurs centaines de pages, plus de 200 000 mots) des thèmes ou enjeux de politique publique – parfois nouveaux – qui rassemblent et/ou divisent les partis politiques ? Et surtout, deux méthodes distinctes (manuelle et automatique) peuvent-elles valider la production de résultats sur l’étude de ces thèmes issus de programmes électoraux et, plus largement, des discours politiques ? Si les techniques d’analyse de textes basées sur l’humain sont les plus répandues et fournissent vraisemblablement les meilleures données en ce qui concerne l’analyse du contenu des programmes électoraux (Budge, 2001, p. 50), les techniques automatisées présentent également d’incontournables avantages en termes de temps, d’absence de financement des codeurs, etc. Mais sont-elles pour autant aussi fiables que l’encodage manuel et produisent-elles des résultats similaires sur les thèmes des partis politiques issus des programmes électoraux ?

En effet, si notre recherche privilégie une analyse des discours politiques par thèmes (Pineira, Tournier, 2009 ; Angermuller, 2010) ou, autrement

appelée, par enjeu (Brouard *et al.*, 2009, 2012) et vise à faire émerger les convergences et divergences entre les partis politiques belges en période électorale, les démarches choisies entre les auteurs de cette contribution diffèrent largement. D'un côté, l'encodage manuel des programmes électoraux est privilégié et s'inscrit dans les travaux issus du *Comparative Agendas Project* (CAP). De l'autre, l'encodage automatique et le logiciel Prospéro, Programmes de sociologie pragmatique expérimentale et réflexive sur ordinateur (Chateauraynaud, 2003) sont favorisés. Il est d'emblée nécessaire de préciser que cette contribution résulte donc de la rencontre de deux groupes de chercheurs travaillant sur les mêmes corpus de textes (les programmes électoraux), le même terrain (les partis politiques belges), mais n'ayant pas la même finalité de recherche puisque l'une vise à étudier le contenu des programmes électoraux et des thèmes émergents (Dandoy, 2014 ; Joly *et al.*, 2014) tandis que l'autre envisage à terme une analyse argumentative des thèmes ou dossiers complexes identifiés dans les programmes électoraux, comme l'enjeu climatique (Piet, 2015). La volonté de comparer les résultats de ces deux recherches et de confronter *in fine* les deux méthodes est donc apparue utile, voire nécessaire, afin d'asseoir et conforter nos analyses sur les partis politiques belges.

Le choix d'un outil informatique n'a, cependant, rien d'anodin quant aux résultats produits. Pour notre démarche de recherche, nous souhaitions pouvoir conserver à tout moment un retour sur nos textes et une maîtrise de notre corpus. Christophe Lejeune (2010, p. 19) propose de classer l'informatisation de l'analyse de textes en cinq familles : l'analyse lexicale, les logiciels « concordanciers », les « automates », les outils « réflexifs » et, enfin, les dictionnaires/registres. Les logiciels associés à l'analyse lexicale s'appuient sur l'exploration statistique du corpus et le traitement des cooccurrences (Lafon, 1981 ; Mayaffre, 2008) au sein de mêmes phrases, paragraphes, etc. Les logiciels « concordanciers » sont assimilés aux « outils de recherche d'expressions » (Lejeune, 2010, p. 21), tandis que les « automates » rendent possibles l'automatisation du « calcul d'agrégats de portions du corpus et de leur assigner (automatiquement) une étiquette » (*ibid.*, p. 23). Les outils dits « réflexifs » assistent le chercheur dans l'analyse textuelle par « une simple fenêtre d'annotation » (Lejeune, 2010, p. 25). Enfin, les logiciels construits sur la base de dictionnaires et/ou registres comme Prospéro se distinguent des outils CAQDAS (*Computer Assisted Qualitative Data Analysis Software*) par leur capacité à calculer et à produire des résultats en partant de catégories et de dictionnaires améliorés au fur et à mesure par le codage du chercheur.

Encodage manuel ou automatique, le cadre d'analyse de nos deux recherches (thématiques politiques et trajectoires des dossiers complexes) n'en reste pas moins commun. En effet, pour prendre la mesure de l'évolution des systèmes politiques des sociétés occidentales, il est essentiel d'identifier les thèmes politiques et priorités des partis. Les priorités sont à ce titre

déterminantes pour comprendre la mise en place des politiques publiques et certains processus politiques. Elles permettent tant d'appréhender les budgets et dépenses étatiques, la législation, ou encore diverses politiques socio-économiques que de comprendre des phénomènes et processus politiques, comme les résultats électoraux, la formation des gouvernements ou le contenu d'accords entre partis. Pour ce faire, en nous inscrivant dans le sillage de la *Saliency Theory* (Robertson, 1976 ; Budge *et al.*, 1987 ; Budge, Farlie, 1983), la priorité politique peut être définie comme le poids relatif d'un thème dans un programme électoral. En effet, si un parti prête attention à de nombreux thèmes, la priorité n'est pas infinie et doit donc être distribuée en tenant compte de l'ensemble des thèmes identifiés dans les programmes électoraux. La priorité politique ne peut donc être identique pour tous les thèmes, sachant que ces derniers restent dépendants du contexte politique, des stratégies du leadership du parti ou tout simplement de l'idéologie partisane. Ainsi, il paraît vraisemblable qu'un parti écologiste va mettre davantage l'accent sur les thèmes environnementaux, tandis qu'un parti socialiste va se focaliser sur les thèmes sociaux. L'argument de base de la *Saliency Theory* repose sur le fait que les partis choisissent stratégiquement de mettre en avant certains thèmes afin de se différencier des autres partis, de séduire une nouvelle frange de l'électorat, etc.

Partis politiques belges et programmes électoraux

Afin de mesurer ces thèmes, le programme électoral reste au regard de la littérature en science politique l'un des documents le plus souvent analysé. À la suite du *Comparative Manifesto Project*, l'utilisation des programmes électoraux s'est imposée dans de nombreuses recherches comparatives sur les partis politiques (Budge, 2001). Ceux-ci présentent de nombreux avantages, en comparaison d'autres techniques parfois utilisées (enquêtes d'opinion, documents parlementaires, vote dans les assemblées, analyses des discours de campagne, etc.). Ils s'agit tout d'abord de documents officiels, écrits et avalisés par les structures du parti. Ils engagent ensuite le parti, ses candidats et ses différentes factions lors de la campagne et lors de la législature. Ils traitent également de tous les thèmes que le parti désire mettre en avant. Ils sont toujours rendus publics et sont destinés aux électeurs, à la société civile et aux autres partis. Ils sont enfin disponibles à intervalle régulier (lors de chaque élection) et, dans la plupart des cas, chaque parti en rédige un. Au demeurant, le principal avantage des programmes électoraux est qu'ils permettent une comparaison dans le temps et dans l'espace – entre partis – sur toute une série de thèmes politiques. Dans l'histoire politique belge, les programmes électoraux ont souvent occupé une place prépondérante, étant donné qu'ils sont à la base du contenu des accords de gouvernement (Joly *et al.*, 2014) – nécessaire pour

la stabilité des gouvernements de coalition. Qui plus est, les enquêtes d'opinion nous révèlent que près d'un tiers des électeurs belges francophones affirment baser leurs votes uniquement sur les programmes électoraux ou davantage sur les programmes que sur les qualités des candidats (Dandroy, 2007 ; Dandroy, Museur, 2015).

Enfin, au vu de la taille potentielle de notre corpus (plusieurs décennies de programmes électoraux disponibles, et cinq niveaux de pouvoir – communal, provincial, régional, fédéral et européen), une sélection des documents a été effectuée afin de rendre la comparaison des thèmes cohérente eu égard à l'évolution des systèmes partisans (Deschouwer, 2012 ; Dandroy *et al.*, 2013 ; Dandroy, De Decker, 2009). Ainsi, nous avons choisi de nous pencher sur un segment temporel assez long, allant de 1987 à 2007, et de nous concentrer sur les élections législatives fédérales. L'avantage de cette période repose tout d'abord sur le fait que la structure des deux systèmes de partis (flamand et francophone) reste très stable à travers le temps, puisqu'aucun nouveau parti politique n'est entré de manière durable au Parlement fédéral dans cette période. L'avantage du choix de se centrer sur ce niveau fédéral repose ensuite sur le fait qu'il s'agit du niveau le plus intéressant pour prendre en compte les différences linguistiques.

Nous comparerons donc, au départ de nos deux méthodes automatique et manuelle, les thèmes identifiés dans 59 programmes électoraux pour les six élections fédérales organisées entre 1987 et 2007 issus de douze partis politiques (tableau 1, synthèse, *infra*). La sélection de ces partis se base sur un critère simple : avoir obtenu au moins un siège au Parlement fédéral (Chambre et/ou Sénat) lors de deux élections parmi les six scrutins analysés. Au sein de la famille écologiste, Agalev (néerlandophone) naît en 1982 et se renomme Groen en 2003 tandis qu'Ecolo (francophone) voit le jour en mars 1980. Au sein de la famille démocrate-chrétienne, le CVP devient en 2001 le parti Christen, Democratisch en Vlaams (CD&V) tandis que le PSC devient le Centre démocrate Humaniste (cdH) en 2002. Du côté des socialistes, le SP change de nom en 2001 et devient le Socialistische Partij anders (sp.a) alors que son homologue francophone, le PS, ne change pas de nom mais fait évoluer ses statuts¹. Au sein de la famille libérale, le PVV devient le VLD (Vlaamse Liberalen en Democraten) en 1992 et se mue en Open Vld en 2007. Le parti libéral francophone devient le Mouvement réformateur (MR) en 2002. Deux partis régionalistes répondent également à notre critère de sélection sans qu'ils soient toutefois des partis frères (Dandroy, De Decker, 2009) : les FDF du côté francophone et la N-VA (Nieuw-Vlaams Alliantie) du côté néerlandophone (héritière de la Volksunie, disparue en 2001). Enfin, deux partis d'extrême-droite clôturent notre corpus : le Vlaams Blok devenu Vlaams Belang (VB) en 2004, et le Front national (FN).

1. Nous respectons ici la typographie officielle des sigles des partis : CD&V, cdH, sp.a.

Tableau 1. Présentation synthétique des partis politiques analysés

Parti politique	Acronyme	Système partisan	Parti frère	Famille politique
Mouvement réformateur	MR	Francophone	Open Vld	Libérale
Parti socialiste	PS	Francophone	sp.a	Socialiste
Centre démocrate humaniste	cdH	Francophone	CD&V	Démocrate-chrétienne
Écolo	Ecolo	Francophone	Groen	Écologiste
Fédéralistes démocrates francophones	FDF	Francophone	-	Régionaliste
Front national	FN	Francophone	-	Extrême droite
Open Vlaamse Liberalen en Democraten-Partij	Open Vld	Néerlandophone	MR	Libérale
Socialistische Partij anders – Sociaal progressief alternatief	sp.a	Néerlandophone	PS	Socialiste
Christen, Democratisch en Vlaams	CD&V	Néerlandophone	cdH	Démocrate-chrétienne
Groen	Groen	Néerlandophone	Ecolo	Écologiste
Nieuw-Vlaams Alliantie	N-VA	Néerlandophone	-	Régionaliste
Vlaams Belang	VB	Néerlandophone	-	Extrême droite

Au total, ces 59 programmes électoraux représentent un corpus de 1 284 468 mots : 625 736 mots pour les 23 programmes francophones et 658 732 mots pour les 36 programmes néerlandophones². Quant à la longueur des programmes électoraux analysés, elle varie entre 3 031 mots dans le cas du VB en 1987 et 208 803 mots dans le cas d'Ecolo en 2007.

Un même *codebook*, deux approches distinctes

À la suite de la *Saliency Theory*, le *Comparative Agendas Project* (CAP) (Baumgartner, Jones, 1993 ; Soroka, 2002 ; Brouard *et al.*, 2009 ; Wilkerson *et al.*, 2009 ; Green-Perdersen, Walgrave, 2014) a développé un *codebook* thématique permettant de catégoriser différents thèmes afin de tracer l'évolution

- Le corpus comptait initialement 64 programmes électoraux pour cette période. Toutefois, nos deux corpus et méthodes ont utilisé deux versions différentes pour pas moins de cinq programmes, étant donné qu'il circule une version complète et une version résumée des programmes électoraux du cdH (2003 et 2007), d'Ecolo (2003 et 2007) et du PS (2007).

de ceux-ci dans les programmes des différents acteurs politiques. Bien que créé initialement pour coder la législation et les budgets nord-américains, ce *codebook* est désormais utilisé pour encoder diverses activités politiques dans de nombreux pays³.

Le *codebook* CAP comprend 249 thèmes regroupés en 21 thématiques (tableau 2) et permet de tracer l'évolution temporelle de ceux-ci pour chaque acteur partisan. Ce *codebook* est commun aux deux recherches et aux deux méthodes manuelles CAP et automatique Prospéro/CAP que nous proposons à présent de développer.

Encodage manuel : l'approche par les agendas comparés

La méthode CAP, utilisée ici pour mesurer les priorités politiques au sein d'un programme électoral, est simple. Dans un premier temps, chaque phrase du document se voit attribuer jusqu'à trois codes thématiques. Dans un second temps, le nombre de références attribuées à chaque thème est comptabilisé. La mesure de l'importance (relative) de chaque enjeu politique pour chaque acteur est calculée sur la base de la proportion de chaque thème par rapport à la totalité du programme électoral – ce que l'on appelle communément un « agenda partisan ».

Le premier défi pour composer l'agenda des partis politiques consiste à obtenir le matériel brut, c'est-à-dire les programmes électoraux tels qu'ils ont été présentés aux électeurs et discutés dans les médias lors de la campagne. Ensuite, ces documents faisant parfois plusieurs centaines de pages ont été digitalisés. Cette digitalisation est une étape cruciale puisque les étudiants formés et rémunérés encodent directement en ligne ces programmes au moyen du logiciel *AmCat*, conçu spécialement pour ce type d'encodage par l'Université d'Amsterdam. L'encodage des programmes se fait phrase par phrase. Les phrases contenant plusieurs idées sont divisées en sous-unités, appelées quasi-phrases selon les techniques appliquées par le *Comparative Manifestos Project*. Les étudiants attribuent alors jusqu'à trois thèmes par quasi-phrase, selon la thématique mise en avant dans la phrase – d'abord le moyen, puis le but de la politique mis en exergue⁴.

Une fois le programme électoral encodé dans sa totalité, des statistiques descriptives nous permettent d'observer combien de fois un thème a été mis en avant par programme et par parti. En calculant la proportion d'attention par

3. Pour plus d'informations, voir le site internet du projet : <http://www.comparativeagendas.info> (consulté le 24 mars 2015).

4. Afin de pouvoir comparer les priorités à travers différents agendas politiques, on encode d'abord le moyen, puis le but de la promesse électorale ou de la politique à mettre en place. Si l'on veut, par exemple, baisser la vitesse maximale des voitures à des fins écologiques, on encodera d'abord le thème « transport » et ensuite le thème « écologique ».

Tableau 2. Présentation des 21 thématiques du *codebook* CAP

Code	Thématique	Code	Thématique
1	Polit. économique et fiscale	13	Affaires sociales
2	Droits et libertés	14	Développement local et politique du logement
3	Santé	15	Entreprises, PME et secteur bancaire
4	Agriculture et pêche	16	Défense
5	Emploi	17	Politique scientifique et des nouvelles technologies
6	Enseignement	18	Commerce extérieur
7	Environnement	19	Affaires européennes, étrangères et coopération au développement
8	Politique énergétique	20	Questions administratives et institutionnelles
9	Immigration	21	Aménagement du territoire
10	Mobilité	23	Culture et loisirs
12	Justice		

thème, il nous est possible de mesurer les priorités politiques de chaque parti. Comme l'indique son nom, l'approche CAP est une approche comparative à plusieurs niveaux : entre partis pour une même élection ou pour un même parti à travers le temps. Ainsi, il est possible de comprendre la manière dont les priorités politiques partisans ont évolué et quand/comment un thème spécifique s'est imposé sur l'agenda d'un parti ou de plusieurs partis. Par la suite, il est possible de positionner les partis sur différents clivages (par exemple sur une échelle gauche-droite) selon les enjeux politiques qui sont mis en avant dans leurs programmes. Il est dès lors possible de voir quels programmes – et quels partis – sont les plus compatibles au niveau de leurs priorités politiques.

Enfin, le concept d'attention et l'utilisation de la saillance relative permettent de comparer les priorités de divers acteurs de la société civile ou au sein d'un même type d'acteur comme les partis politiques, et ce, à travers différents stades du processus décisionnel. Le choix de se concentrer sur la saillance et non sur une sélection de propositions concrètes se justifie dans cette optique comparative. Ainsi, plusieurs auteurs ont constaté les effets des partis politiques sur les politiques des gouvernements dans plusieurs pays ou l'effet des discours publics sur les politiques suivantes (Laver, Budge, 1992 ; Klingemann *et al.*, 1994 ; McDonald, Budge, 2005).

Encodage automatique : Prospéro et l'analyse des dossiers complexes

Le choix du logiciel Prospéro s'est imposé initialement dans notre démarche de recherche, d'une part parce qu'il permettait d'analyser le corpus par acteur et par thématique et, d'autre part, parce qu'il permettait d'identifier les différents registres argumentatifs – qu'ils soient ou non rattachés à un thème étudié – ainsi que les cooccurrences entre plusieurs thématiques, thèmes, voire représentants, au sein d'un même énoncé (une même phrase) (Piet, 2013). Comme le souligne Patrick Trabal, Prospéro a davantage été pensé en termes de « dossiers » qu'en termes d'agendas puisque ce qui occupe le chercheur au travers de ce logiciel repose sur la trajectoire des thèmes étudiés eu égard à l'évolution des jeux d'acteurs et d'arguments : « on va regarder le corpus comme un *dossier*, c'est-à-dire comme un ensemble de textes dans lesquels se discutent des notions, s'opposent des points de vue et se déploient des argumentations » (2005).

Si la prise en main du logiciel demande une petite formation, sa compréhension et son utilisation sont relativement simples. Prospéro s'appuie déjà sur un ensemble de registres préconstruits par les précédents utilisateurs permettant un usage immédiat et une première exploration du corpus. Pour notre recherche, nous sommes toutefois repartis d'une page blanche ; nous avons encodé les 21 thématiques du *codebook* CAP ainsi que leurs 249 thèmes. Pour chacun d'entre eux, au fur à et mesure de nos lectures, nous avons ensuite introduit plusieurs représentants (mots et expressions), et ce, jusqu'au moment où les registres semblent stabilisés – autrement dit, jusqu'au moment où le logiciel nous indique que chaque nouveau représentant potentiel identifié dans le corpus durant notre lecture se trouve déjà associé à un thème. Une fois les registres par thème stabilisés, nous avons relevé l'ensemble des occurrences présentées par Prospéro pour chacune des 21 thématiques du *codebook* CAP. Dans le logiciel, les thèmes construits sont donc « formés de plusieurs représentants, dont la composition est sous le contrôle de l'utilisateur » (Chateauraynaud, 2003, p. 215). Une fois le codage des représentants réalisé, il est alors facile d'identifier les mots et expressions présents et absents dans l'ensemble du dossier étudié. Dans le cadre de cet article, nous avons travaillé sur une base de 16 862 représentants qualifiant les 21 thématiques et les 249 thèmes du *codebook* Prospéro/CAP.

Ainsi, les thématiques, thèmes et représentants se présentent sur quatre colonnes (figure 1) sous la forme d'une hiérarchie. La première colonne (intitulée « êtres-fictifs ») reprend l'ensemble des 21 thématiques. La deuxième colonne (« type ») englobe l'ensemble des thèmes pour chaque thématique. Les troisième et quatrième colonnes rassemblent tous les représentants (mots et expressions) encodés par thème afin que le logiciel puisse l'identifier dans le corpus de textes.

La présentation des résultats peut se faire par thématique, par thème, voire par représentant « présent » dans le corpus (figure 2) avec, comme indication, le nom de l'auteur, la date de publication du document et le nombre d'occurrences relevé par texte. Par ailleurs, il est toujours possible de retourner au texte et à ses énoncés apparentés à une phrase que le logiciel identifie d'un point à un autre.

Convergences entre les deux approches

Par ailleurs, si les deux méthodes sont différentes, la manière d'appréhender les enjeux politiques ou les problèmes publics l'est nettement moins. En effet, d'une part, Prospéro est un outil informatique qui vise à étudier « des processus historiques, des controverses, des affaires, des crises, des polémiques, autant de configurations que nous rangeons sous l'appellation commode de “dossiers complexes” » (Chateauraynaud, 2003, p. 17-18). D'autre part, le projet CAP repose sur trois questions majeures : 1) Comment les problèmes parviennent-ils à l'agenda politique ? 2) Comment ces agendas politiques évoluent-ils au cours du temps et pourquoi ? 3) Dans quelle mesure les priorités politiques reflètent-elles les priorités des citoyens ? (Wilkerson *et al.*, 2009). Nous retrouvons en partie ces interrogations dans le développement du logiciel Prospéro et de la sociologie argumentative à travers plusieurs débats théoriques, épistémologiques et méthodologiques depuis la fin des années quatre-vingt.

Au niveau théorique, l'objectif de Chateauraynaud est de répondre à un besoin de la sociologie contemporaine en dépassant une certaine forme de dualité entre « sociologie de la critique » et « théorie de l'acteur-réseau » en proposant une méthode d'analyse des changements et turbulences au travers des interactions entre les acteurs, les objets, les controverses, les crises, etc. (Chateauraynaud, Tornay, 1999 ; Chateauraynaud, 2011). Cette démarche, complétée d'une analyse des récits et des arguments des différents acteurs, s'inscrit aisément à la suite des travaux de la théorie de la saillance et de l'étude des évolutions des enjeux politiques telle qu'envisagée, notamment, par Carmines et Stimson (1989) et Baumgartner et Jones (1993).

Au niveau épistémologique, le débat porte sur la dualité au sein des sciences sociales entre l'approche quantitative et l'approche qualitative. Le statut de l'interprétation (Chateauraynaud, 2003) est au cœur de la réflexion sociologique entamée à travers le logiciel Prospéro. Il se voit donner un rôle d'appui à l'interprétation des changements, ruptures et turbulences conjoncturelles d'un dossier complexe étudié par l'utilisateur à travers les jeux d'arguments et les formes de la critique sur lesquels les acteurs prennent appui. L'analyse quantitative préalable travaille, d'une part, à la mise en contexte des dossiers complexes et longs pour lesquels nous pouvons par la suite produire

des analyses qualitatives discursives et argumentatives. Le logiciel permet, d'autre part, de dépasser l'opposition désuète entre approches quantitatives et qualitatives (*ibid.*, p. 63). Nous retrouvons également dans ce débat épistémologique les tenants de la méthode CAP permettant de passer de l'analyse du contenu à l'analyse du discours pour étudier tant les évolutions des enjeux que la manière dont chaque acteur politique s'y positionne (voir par exemple Budge *et al.*, 2001).

Le débat est, enfin, méthodologique. Prospéro propose cinq niveaux d'analyse différents : sur l'ensemble du corpus, sur les acteurs, les textes, les discours et les arguments. À nouveau, les deux méthodes ont des points de convergence. Tout d'abord, il est possible de travailler sur une grande variété de textes et sources différentes (programmes électoraux, médias, législation, etc.). L'analyse peut, ensuite, se concentrer sur les thématiques et thèmes qui émergent dans un dossier, et ce, soit texte par texte, soit acteur par acteur. Les deux méthodes envisagent également d'étudier l'« environnement » lexical dans lequel se développent les entités et les acteurs au sein des énoncés (définis ici comme une phrase). Les différences essentielles se concentrent donc au niveau de la considération des enchaînements argumentatifs et de la sociologie argumentative (Chateauraynaud, 2011) absente dans la méthode CAP.

Comparaison des méthodes manuelle et automatique

Pour cet article, la comparaison se concentre au niveau des résultats des thématiques issues des deux méthodes et identifiées dans les programmes électoraux⁵. Plusieurs analyses sont à ce titre proposées pour juger de la qualité des résultats et de leur proximité : l'analyse des similarités (Green-Pedersen, Mortensen, 2009), l'analyse des corrélations ainsi que l'écart des rangs et leur corrélation. La première nous informe sur la similarité des thématiques présentes dans chaque texte et la faible variabilité des données. La deuxième nous informe sur la liaison entre la distribution des enjeux de chaque programme électoral tandis que les troisième et quatrième analyses nous indiquent la variation de la hiérarchie ou du rang des thématiques pour chaque programme. Le tableau 3 illustre ces mesures appliquées au programme électoral du parti libéral francophone (PRL) en 1987.

De manière générale, la similarité moyenne entre les résultats des deux méthodes est de 83,6 % sur l'ensemble du corpus. Autrement dit, un peu plus de 16 % des priorités politiques attribuées aux 21 ensembles de politiques publiques fluctuent entre les deux méthodes. La corrélation moyenne (0,86)

5. L'objectif du présent article étant méthodologique, nous renvoyons l'auteur intéressé par les résultats empiriques publiés sur le blog « Élections fédérales, régionales et européennes en Belgique - Mai 2014 » : <http://electionsbelges2014.blogspot.be>.

Tableau 3. Exemple : le programme électoral du PRL (1987)

	Prospéro (N= 2201)	CAP (N= 2344)	Diffé- rence	Rang Pros- péro	Rang CAP
Administration publique	26,3 %	24,9 %	- 1,4 %	1	1
Macroéconomie	12,3 %	12,4 %	0,1 %	2	2
Entreprises	8,9 %	6,4 %	- 2,5 %	3	4
Affaires étrangères	7,5 %	4,2 %	- 3,3 %	4	8
Travail	6,5 %	3,5 %	- 3 %	5	11
Droits et libertés	5,3 %	4,3 %	- 1 %	6	6
Enseignement	5 %	6,8 %	1,8 %	7	3
Justice	5 %	5,5 %	0,5 %	8	5
Culture	3,8 %	3,7 %	- 0,1 %	9	10
Affaires sociales	3,2 %	3,0 %	- 0,2 %	10	13
Santé	2,6 %	4,2 %	1,6 %	11	7
Mobilité	2,5 %	3,0 %	0,5 %	12	12
Commerce extérieur	2,3 %	2,9 %	0,6 %	13	15
Agriculture	1,5 %	2,4 %	0,9 %	14	16
Défense	1,5 %	4,1 %	2,6 %	15	9
Développement local	1,4 %	1,9 %	0,5 %	16	17
Recherche scientifique	1,3 %	1,8 %	0,5 %	17	18
Immigration	1,3 %	2,9 %	1,6 %	18	14
Environnement	1 %	0,4 %	- 0,8 %	19	21
Énergie	0,5 %	1,3 %	0,8 %	20	19
Territoire	0,2 %	0,5 %	0,3 %	21	20
Total	100,0 %	100,0 %			
Similarités des attentions politiques					87,8 %
Corrélation des attentions politiques					.97
Corrélation de l'écart des rangs					.89

entre les 59 textes confirme le précédent résultat et la forte ressemblance entre l'encodage des mêmes programmes électoraux au travers des deux méthodes. Nous relevons également des similarités moyennes relativement proches entre les programmes francophones et néerlandophones, et ce, même si les versions francophones révèlent une similarité moyenne légèrement plus élevée (84,3 %) par rapport aux programmes électoraux des partis flamands (83,2 %). Les corrélations moyennes des versions francophones (0,88) et néerlandophones (0,85) confirment également cette tendance. Ces résultats donnent déjà une première indication positive quant à la qualité et la validité des deux méthodes.

En dehors de ces quelques cas particuliers, l'ensemble des similarités moyennes par parti se situe entre 80,4 % (FN) et 86,5 % (cdH), du côté francophone, et entre 81,01 % (VB) – suivi directement de Groen (81,3 %) – et 86,6 % (CD&V), du côté néerlandophone (voir tableau 4 pour l'ensemble des données). Au regard de cette analyse, les familles politiques écologiste et d'extrême-droite révèlent le plus fort taux de redistribution des attentions politiques (19 %) entre les deux méthodes pour un même texte. L'analyse des corrélations moyennes confirme en partie ce résultat : entre 0,70 (Groen) et 0,92 (N-VA) chez les partis néerlandophones, et entre 0,77 (FN) et 0,94 (FDF) chez les partis francophones. Ces importantes différences ne sont ainsi pas spécifiques aux petits partis et aux partis « de niche », puisque les corrélations maximales sont observées dans le cas de partis correspondant aux mêmes caractéristiques. Entre ces extrêmes, nous retrouvons tous les principaux partis traditionnels belges (tableau 4).

Enfin, l'analyse de l'écart moyen de rang et leur corrélation confirme les précédents constats sur les deux méthodes. L'écart moyen des rangs est de 2,2 et révèle une corrélation moyenne de 0,87. Les thématiques varient en moyenne de deux rangs entre l'analyse produite par les deux méthodes, et ce, tant au sein des programmes électoraux francophones (2,3) que néerlandophones (2,2). Les corrélations moyennes des rangs sont, quant à elles, identiques (respectivement 0,86 et 0,87).

Exemple de comparaison des résultats produits : les thématiques écologiques

Si les deux méthodes testées ici semblent proches sur l'ensemble des textes d'un même corpus, produisent-elles pour autant les mêmes résultats et les mêmes trajectoires lorsque nous étudions l'évolution temporelle d'une thématique ? L'hypothèse que nous soulevons à ce stade est la suivante : si les méthodes d'encodage automatique et manuelle s'équivalent et valident chacune les résultats produits par l'autre, elles devraient reproduire des évolutions similaires par thématique ou par regroupement thématique. Nous proposons,

Tableau 4. Mesures de similarités et de corrélation par texte et entre méthodes

Année	Parti politique	Similarité (%)	corrélation	Écart de rang (similarité)	Écart de rang (corrélation)
1987	PS	85,2	.96	2,7	.82
	PRL	87,8	.97	2,2	.89
	PSC	86	.94	2,2	.89
	Ecolo	83,2	.92	1,1	.95
	DFD	80,1	.90	2,7	.82
	SP	87,9	.93	1,6	.94
	PVV	80,1	.91	2	.90
	CVP	87,8	.96	3,2	.78
	Agalev	79,9	.58	2,2	.88
	VU	84,6	.96	2,7	.84
	VB	76,5	.82	3,1	.79
1991	PS	92,4	.97	1,8	.93
	PRL	81,7	.84	2	.90
	PSC	87,4	.92	1,8	.92
	Ecolo	88,9	.92	2,3	.89
	DFD	84,1	.97	1,7	.93
	SP	89,2	.85	2,3	.85
	PVV	84,6	.87	4	.65
	CVP	80,3	.82	2	.91
	Agalev	66,8	.42	2,6	.86
	VU	80,5	.89	2,8	.83
	VB	86,5	.97	2,6	.87
1995	PS	80,9	.81	2,6	.86
	PRL	86,6	.97	1,9	.91
	PSC	87	.92	1,1	.97
	ECOLO	86	.92	2,3	.88
	SP	85,9	.87	2,2	.89
	VLD	87,7	.94	2,2	.85
	CVP	88,8	.96	2,3	.87
	Agalev	88,7	.82	1,9	.89
	VU	78,1	.80	2,7	.82
	VB	83,4	.85	2	.91

Année	Parti politique	Similarité (%)	corrélation	Écart de rang (similarité)	Écart de rang (corrélation)
1999	PS	81,2	.77	3,7	.67
	PRL	84,9	.87	2,7	.78
	PSC	85,6	.90	2	.91
	Ecolo	79,5	.67	2,2	.88
	SP	81,9	.83	1,9	.89
	VLD	88,9	.92	2,1	.90
	CVP	88,8	.93	1,7	.92
	Agalev	86,1	.78	3	.79
	VU	83,6	.95	2,4	.82
	VB	71,4	.76	2,5	.84
2003	PS	85,4	.86	1,7	.91
	MR	82,7	.84	1,3	.95
	FN	81,9	.79	2,2	.87
	SP	79,4	.80	1,5	.94
	VLD	72,8	.72	2	.87
	CD&V	90,7	.96	2,4	.90
	Agalev	81,4	.75	1,3	.96
	N-VA	84,9	.96	2,9	.83
	VB	85,4	.92	2,5	.82
2007	MR	80,8	.75	2,3	.89
	FN	78,8	.75	1,5	.92
	SP	83	.77	3,2	.75
	VLD	80,9	.74	1,8	.93
	CD&V	83,2	.85	1,4	.93
	Groen	84,9	.84	2,6	.88
	N-VA	87,4	.95	2,4	.87
	VB	83	.84	2,4	.86
MOYENNE		83,6	.86	2,2	.87

Tableau 5. Thèmes des trois thématiques écologiques

Thématiques	Thèmes
Agriculture et pêche	400 - Général 401 - Exportation et importation agricole 402 - Subventions de l'État 403 - Politique alimentaire 404 - Promotion de l'agriculture et information du citoyen 405 - Lutte contre maladie plantes, animaux 406 - Bien-être des animaux dans l'agriculture 407 - Problématique environnement-agriculture 408 - Pêche 498 - Recherche sur l'agriculture 499 - Autres
Environnement	700 - Général 701 - Eau, qualité, approvisionnement 703 - Problématique des déchets 704 - Substances toxiques, déchets 705 - Pollution atmosphérique, effet de serre et pollution sonore 707 - Recyclage 708 - Dangers écologiques intérieurs 710 - Pollution de la côte, mer du Nord 711 - Protection des sols, eau douce 712 - Risques techniques 720 - Protection des bois, arbres, plantes 721 - Protection des espèces animales 797 - Catastrophes naturelles 798 - Recherche et développement 799 - Autres
Politique énergétique	800 - Général 801 - Énergie nucléaire 802 - Électricité, électricité hydrolique 803 - Gaz et pétrole 805 - Charbon 806 - Sources d'énergie alternative et renouvelable 807 - Économie d'énergie 898 - Recherche et évolution 899 - Autres

pour ce faire, de tester une trajectoire en particulier : celle des thématiques écologiques. Celles-ci englobent trois distinctes : l'environnement, la politique énergétique et l'agriculture (voir tableau 5 pour la présentation des thèmes par thématique). Ces thématiques sont essentielles pour comprendre le système partisan belge. Tout d'abord l'environnement est un enjeu-clé pour les deux partis écologistes et est devenu, suite au succès de ces partis, un réel élément de campagne. La politique énergétique – et surtout son volet nucléaire – est omniprésente au sein de ces campagnes électorales, tandis que l'agriculture est un enjeu clé pour les partis chrétiens-démocrates, principalement en Flandre.

Au total, dix trajectoires de partis politiques sont analysées⁶. Parmi celles-ci, cinq trajectoires suivent les mêmes évolutions. C'est le cas du sp.a, du CD&V, de la N-VA, d'Ecolo et du MR (figure 3). Deux trajectoires sont très similaires (figure 4), mais rencontrent une évolution différente : au VLD (1995-1999) et au VB (1987-1991). Enfin, trois trajectoires sont relativement peu semblables, même si elles restent dans un ordre de magnitude relativement proche comme relevé au PS, au cdH et chez Groen (figure 5).

Le relevé des trajectoires est, en soi, très intéressant car il permet de confirmer la qualité et la validité externe des deux méthodes et de leurs résultats. Certaines différences persistent entre les deux approches automatique et manuelle, mais elles ne remettent en cause ni les tendances des trajectoires, ni l'évolution des enjeux. Elles se situent davantage au niveau de l'un ou l'autre programme électoral. Tout comme le démontrait déjà l'analyse des similarités, des corrélations et des écarts de rangs, 11 programmes électoraux sur les 55 mobilisés pour l'analyse des trajectoires des thématiques écologiques affichent des tendances différentes, soit à la baisse, soit à la hausse, en fonction des deux méthodes utilisées.

Discussion

Les diverses analyses (similarité, corrélation, écart des rangs et trajectoires de thématiques) ont démontré que les méthodes d'encodage manuel et automatisé sont relativement proches sur l'ensemble du corpus, même si quelques différences et variations sont relevées. Ainsi, en moyenne, seules 16 % des attentions politiques sont redistribuées différemment d'une méthode à l'autre. De même, l'analyse des trajectoires montre que seul un programme électoral sur cinq présente des tendances sensiblement différentes. Plusieurs éléments permettent certainement d'expliquer ces écarts.

Premièrement, l'élément principal repose sur la méthode d'encodage elle-même. En effet, même si les 21 thématiques traitées sont similaires, comptabiliser des données au regard de phrases ou de quasi-phrases est clairement distinct d'un codage sur la base de mots et d'expressions repris dans des répertoires construits par le chercheur. Le nombre total de données encodées par chaque méthode varie d'ailleurs, en moyenne, de 41 % sur l'ensemble des 59 programmes de notre corpus. La taille du document n'a toutefois pas d'influence sur cette variation des données. Celle-ci s'explique principalement

6. Étant donné que les deux méthodes se sont basées pour certains partis sur des documents différents, les données relatives au PS (2007), au cdH (2003-2007) et à Ecolo (2003-2007) ne sont pas incluses dans les graphes. Les programmes des FDF et du FN ont été laissés de côté pour cette analyse puisqu'ils n'ont présenté un programme électoral qu'à deux reprises, ne permettant donc pas d'établir une quelconque trajectoire.

Figure 3. Cinq mêmes trajectoires : sp.a, CD&V, N-VA, Ecolo et MR

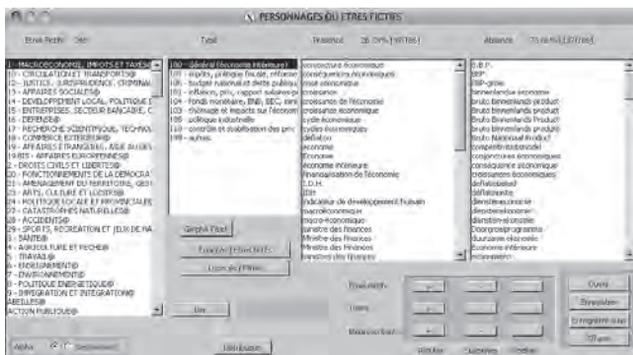


Figure 4. Deux trajectoires similaires, mais une évolution différente entre deux élections : Open Vld et VB

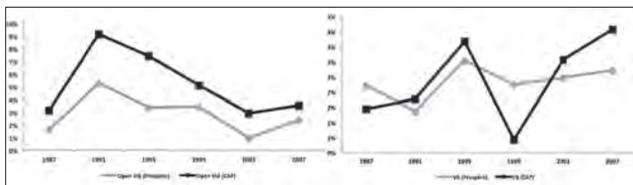
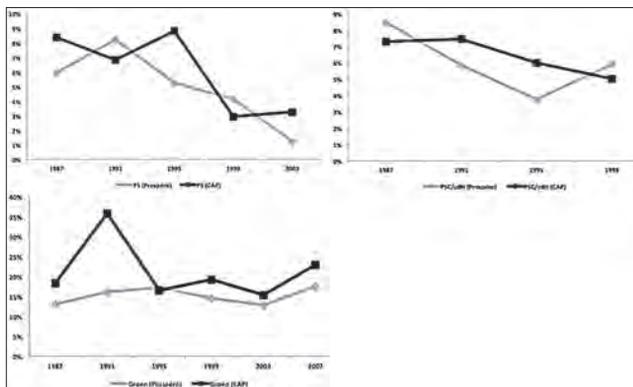


Figure 5. Trois trajectoires différentes : PS, cdH, Groen



par le fait que le logiciel ne prend pas en compte les anaphores grammaticales dans le comptage des thématiques ; ce que permet de faire l'encodage manuel.

Deuxièmement, nous constatons que la manière de rédiger un programme électoral ou le vocabulaire particulier utilisé par certains partis politiques joue en partie un rôle dans la proximité des résultats entre les deux méthodes d'encodage. Du côté des partis francophones, par exemple, les programmes du FN testés sont les plus éloignés : 80,4 % de similarité moyenne et 0,77 de corrélation moyenne. Du côté néerlandophone, les programmes électoraux du parti écologique (Groen) connaissent la plus faible corrélation moyenne (0,70). Ce résultat est toutefois à nuancer, d'une part en indiquant que cette faible corrélation est essentiellement due à deux liaisons très faibles entre les programmes de 1987 (0,58 de corrélation) et les programmes de 1991 (0,42) et d'autre part, par une similarité moyenne de 81,3 % qui reste fort acceptable pour l'analyse d'un même texte. Nous constatons également que, tant chez les partis politiques francophones que néerlandophones, les partis de la famille démocrate-chrétienne présentent les niveaux les plus hauts en termes de similarité moyenne (86,6 % pour les programmes du CD&V et 86,5 % pour ceux du cdH).

Troisièmement, les métaphores ou les expressions particulières limite la capacité de la méthode automatique à identifier certaines données que la méthode manuelle, par l'intervention d'un codeur, n'a aucune difficulté à prendre en compte.

Quatrièmement, les deux méthodes influent sur la production des résultats. Par exemple, le grand nombre de codeurs dans l'approche par « agendas » multiplie les risques d'un codage différent d'un programme électoral à l'autre. Pas moins de 17 codeurs ont participé à l'encodage de l'ensemble des programmes électoraux allant de 1987 à 2007, à la différence de la méthode automatique qui repose sur un seul codeur – le chercheur et son logiciel. Par ailleurs, selon nos analyses préliminaires, la fiabilité entre codeurs est équivalente au niveau de fiabilité relevée entre les deux méthodes. De plus, la systématisme de l'approche par logiciel tranche, d'une part, immanquablement par rapport au choix laissé aux chercheurs dans la méthode manuelle d'encoder ou non une donnée dans une thématique. Elle a, d'autre part, des conséquences sur les résultats produits. À titre d'exemple et de comparaison, nous avons analysé, pour valider notre méthode automatique (dans sa version française et néerlandaise), un programme électoral produit dans les deux langues par un petit parti régionaliste (ProBruxsel) qui se présentait aux élections régionales bruxelloises de mai 2014. Les cantons électoraux de la Région Bruxelles-Capitale étant bilingues, ce parti avait décidé de rédiger une version française et une version néerlandaise du même programme électoral. Il ressortait alors

7. Une majorité des programmes électoraux ont été encodés par deux codeurs différents et leurs encodages comparés.

une corrélation de 0,99, une similarité de 92,7 % et un écart des rangs de 0,7 entre les deux versions du même programme.

Cinquièmement, la variation entre certains résultats s'explique par les éléments des programmes électoraux qui ont été effectivement encodés. Dans certains cas, les titres, les tables de matières, les introductions, les conclusions, les « mots du président », etc., ont été laissés ou retirés de l'encodage des programmes en fonction de la méthode choisie. Les textes traités ne sont donc pas totalement identiques lors de leur traitement.

Enfin, sixièmement, la méthode automatique cherche impérativement à éviter les doublons. Autrement dit, le chercheur tente d'éviter qu'un même représentant se retrouve dans deux catégories différentes. Les mots et expressions comme « développement durable », « environnement » ou « climat » ont été, dans un premier temps, isolés, puis, dans un second temps, réintroduits, au cas par cas, dans les thèmes correspondant au contexte de l'énoncé. Il peut donc subsister quelques erreurs minimales dans le recodage de certains mots et expressions.

L'objectif de cet article était, d'une part, de comparer les résultats issus de l'analyse d'un même corpus de textes (les programmes électoraux des principaux partis politiques belges entre 1987 et 2007) au moyen de deux méthodes distinctes (une méthode d'encodage manuel et une méthode d'encodage automatisé) et, d'autre part, de voir si ces résultats pouvaient être validés par deux méthodes distinctes. Nos analyses démontrent que les deux méthodes mènent à des résultats (très) proches tant au niveau de l'étude des programmes électoraux et des thèmes qu'au regard de la tendance générale des trajectoires des thématiques.

Par ailleurs, il est possible que d'autres outils informatiques valident également nos résultats, comme par exemple *Wordfish* ou *Wordscores*, deux logiciels qui ont souvent été utilisés dans l'analyse des programmes électoraux. L'étude comparée de Albaugh *et al.* (2013), par exemple, s'appuyant sur le même *codebook* CAP et un autre logiciel (RTextTools), conclut sur un corpus d'articles de presse à des résultats relativement convergents entre les méthodes manuelle et automatique, même s'ils restent un peu moins convaincants que la présente analyse sur les programmes électoraux.

Bien entendu, chacune de ces deux méthodes comporte des avantages et des inconvénients. Basée sur un *codebook* et une méthodologie communs relativement invariants dans le temps et l'espace, la méthode CAP permet la comparaison de textes issus de nombreux pays et de différentes périodes. Son principal avantage est qu'elle permet ainsi de concilier une analyse systématique d'un document avec une expertise et une connaissance approfondie d'un parti et/ou d'une thématique. Mais cette méthode présente également d'importantes faiblesses. Tout d'abord, elle requiert un investissement intensif en temps et en ressources humaines et financières. Ensuite, elle peut par-

fois poser des problèmes en termes de fiabilité entre différents codeurs, voire entre le même codeur à deux points dans le temps. Ce désaccord peut porter non seulement sur le découpage du texte en phrases et en quasi-phrases mais également sur le ou les codes à attribuer à chacune de ces unités. La méthode automatique permet, quant à elle, un gain de temps considérable, l'analyse simultanée de textes bilingues lorsque les thématiques sont encodées dans deux ou plusieurs langues, la reproduction rapide des analyses textuelles, l'absence de codeurs multiples, l'amendement du ou des *codebooks*, etc. L'effet pervers de cette méthode repose sur la possibilité de recoder continuellement les thématiques. Dans notre cas, la transposition du *codebook* CAP dans le logiciel Propéro a nécessité quelque 18 mois d'encodage et de phases de tests.

Les prochains défis de nos recherches aidées de l'encodage automatique seront doubles. Premièrement, il sera très intéressant d'élargir le spectre de nos recherches en introduisant d'autres langues (anglais, allemand, espagnol, etc.) dans le cadre d'une étude comparée des programmes électoraux au niveau des pays européens. Deuxièmement, il serait primordial d'arriver à prendre en compte les anaphores grammaticales qui représentent une perte d'information importante pour nos propres analyses discursives et argumentatives. Pour l'heure, nous sommes toujours contraints de repasser par les énoncés pour les identifier et faire, ensuite, l'étude argumentative sur les thématiques (Piet, 2015).

Références

- ALBAUGH Quinn, SOROKA Stuart, JOLY Jeroen, LOEWEN Peter, SEVENANS Julie, WALGRAVE Stefaan, 2013, « Comparing and combining machine learning and dictionary-based approaches to topic coding », *6th Annual Comparative Agendas Project (CAP) Conference*, Anvers, 27-29 juin 2013.
- ANGERMULLER Johannes, 2010, « Analyser le discours politique en Allemagne (1980-2010) », *Mots. Les langages du politique*, n° 94, *Trente ans d'étude des langages du politique (1980-2010)*.
- BAUMGARTNER Frank R., JONES Bryan D., 1993, *Agendas and Instability in American Politics*, Chicago, University of Chicago Press.
- BRAUD Philippe, 2010, « L'apport de la science politique à l'étude des langages du politique », *Mots. Les langages du politique*, n° 94, *Trente ans d'étude des langages du politique (1980-2010)*.
- BROUARD Sylvain, CHAQUÉS Laura, GREEN-PEDERSEN Christoffer, TIMMERMANS Arco, 2009, « Comparer les agendas politiques. Une nouvelle voie pour la politique comparée », *Revue internationale de politique comparée*, n° 16 (3), p. 361-364.
- BROUARD Sylvain, GROSSMAN Emiliano, GUINAUDEAU Isabelle, 2012, « La compétition partisane française au prisme des priorités électorales. Compétition sur enjeux

- et appropriations thématiques», *Revue française de science politique*, n° 62 (2), p. 255-276.
- BUDGE Ian, 2001, « Validating the Manifesto Research Group approach. Theoretical assumptions and empirical confirmations », *Estimating the Policy Positions of Political Actors*, M. Laver éd., Oxford/New York, Routledge, p. 50-65.
- BUDGE Ian, FARLIE Dennis, 1983, *Explaining and Predicting Elections. Issue Effects and Party Strategies in Twenty-Three Democracies*, Londres, George Allen & Unwi.
- BUDGE Ian, ROBERTSON David, HEARL Derek éd., 1987, *Ideology, Strategy and Party Change. Spatial Analyses of Post-War Election Programmes in 19 Democracies*, Cambridge, Cambridge University Press.
- BUDGE Ian, KLINGEMAN Hans-Dieter, VOLKENS Andrea, BARA Judith, TANENBAUM Eric éd., 2001, *Mapping Policy Preferences. Estimates for Partis, Electors and Governments*, Oxford, Oxford University Press.
- CARMINES Edward G., STIMSON James A., 1989, *Issue Evolution. Race and the Transformation of American Politics*, Princeton, Princeton University Press.
- CHATEAURAYNAUD François, 2003, *Prospéro. Une technologie littéraire pour les sciences humaines*, Paris, CNRS.
- 2011, *Argumenter dans un champ de forces. Essai de balistique sociologique*, Paris, Pétra.
- CHATEAURAYNAUD François, TORNY Didier, 1999, *Les sombres précurseurs. Une sociologie pragmatique de l'alerte et du risque*, Paris, EHESS.
- DANDOUY Régis, 2007, « L'analyse des programmes de partis », *Élections : le reflux ? Comportements et attitudes lors des élections en Belgique*, A.-P. Frogner, L. de Winter, P. Baudewyns éd., Bruxelles, De Boeck, p. 141-156.
- 2014, « The impact of government participation and prospects on party policy preferences in Belgium », *Government and Opposition*, n° 49 (4), p. 630-657.
- DANDOUY Régis, DECKER Nicolas (DE), 2009, « Peut-on encore parler de partis frères en Belgique ? », J.-B. Pilet, J.-M. De Waele, S. Jaumain éd., *L'absence de partis nationaux. Menace ou opportunité ?*, Bruxelles, Université de Bruxelles, p. 19-35.
- DANDOUY Régis, MATAGNE Geoffroy, VAN WYNSBERGHE Caroline éd., 2013, *Le fédéralisme belge. Enjeux institutionnels, acteurs sociopolitiques et opinions publiques*, Louvain-la-Neuve, Academia - L'Harmattan.
- DANDOUY Régis, MUSEUR Clément, à paraître (2015), « Les programmes électoraux des partis francophones en 2007 et 2010 », *Être électeur en Wallonie*, P. Baudewyns éd.
- DESCHOUWER Kris, 2012, *The Politics of Belgium. Governing a Divided Society*, 2^e édition, New York - Basingstoke, Palgrave - MacMillan.
- GREEN-PEDERSEN Christoffer, MORTENSEN Peter B., 2009, « Issue competition and election campaigns. Avoidance and engagement », *Potsdam, ECPR Conference*, en ligne [<http://www.agendasetting.dk/papers/uploaded/115201044905PM.pdf>], consulté le 15 août 2014.
- GREEN-PEDERSEN Christoffer, WALGRAVE Stefaan éd., 2014, *Agenda Setting, Policies, and Political Systems. A Comparative Approach*, Chicago, University of Chicago Press.
- JOLY Jeroen, ZICHA Brandon, DANDOUY Régis, 2014, « Does the government agreement's grip on policy fade over time? An analysis of policy drift in Belgium », *Acta Poli-*

- tica, en ligne [http://www.palgrave-journals.com/ap/journal/vaop/ncurrent/full/ap201418a.html], consulté le 19 mars 2015.
- KLINGEMAN Hans-Dieter, HOFFERBERT Richard I., BUDGE Ian éd., 1994, *Parties, Policies and Democracy*, Boulder, Westview Press.
- LAFON Pierre, 1981, « Analyse lexicométrique et recherche des cooccurrences », *Mots / Ordinateurs / Textes / Sociétés*, n° 3, p. 95-148.
- LAVIER Michael, BUDGE Ian, 1992, *Party Policy and Government Coalitions*, Londres, Macmillan.
- LEJEUNE Christophe, 2010, « Montrer, calculer, explorer, analyser. Ce que l'informatique fait (faire) à l'analyse qualitative », *Recherches qualitatives*, hors-série, n° 9, p. 15-32.
- MAYAFFRE Damon, 2008, « De l'occurrence à l'isotopie. Les co-occurrences en lexicométrie », *Sémantique & Syntaxe*, n° 9, p. 53-72.
- MCDONALD Michael D., BUDGE Ian, 2005, *Elections, Parties, Democracy. Conferring the Median Mandate*, Oxford, Oxford University Press.
- PINEIRA Carmen, TOURNIER Maurice, 2009, « Ségolène Royal entre François Bayrou et Nicolas Sarkozy. Approche lexicométrique », *Mots. Les langages du politique*, n° 89, 2007. *Débats pour l'Élysée*, p. 83-104.
- PIET Grégory, 2013, « Derrière les mots du politique. *Le-pire-est-à-venir*, un marqueur temporel particulier dans le discours politique en Belgique », *Mots. Les langages du politiques*, n° 103, *Le silence en politique*, p. 121-136.
- 2015, « Compétition partisane et processus de transformation politique d'enjeux de politique publique. De l'attention politique à la priorité politique de l'enjeu climatique en Belgique (1980-2014) », thèse de doctorat, Liège, Université de Liège.
- ROBERTSON David B., 1976, *A Theory of Party Competition*, Londres - New York, Wiley.
- SOROKA Stuart, 2002, *Agenda-setting Dynamics in Canada*, Vancouver, UBC Press.
- TRABAL Patrick, 2005, « Le logiciel Prospéro à l'épreuve d'un corpus de résumés sociologiques », *Bulletin de méthodologie sociologique*, n° 85, en ligne [http://bms.revues.org/993], consulté le 11 février 2015.
- WILKERSON John, BAUMGARTNER Frank R., BROUARD Sylvain, CHAQUÉS Laura, GREEN-PEDERSEN Christoffer, GROSSMAN Emiliano, JONES Bryan D., TIMMERMANS Arco, WALGRAVE Stefaan, 2009, « Le Projet Agendas Comparés. Objectifs et contenus », *Revue internationale de politique comparée*, n° 2 (10), p. 365-379.