

Quelques méthodes pour l'étude des relations entre classifications lexicales de corpus hétérogènes : application aux débats à l'Assemblée Nationale et aux sites web de partis politiques*

Pierre Ratinaud¹, Pascal Marchand²

¹Laboratoire LERASS, Université de Toulouse – France

²Laboratoire LERASS, Université de Toulouse – France

Abstract

This work compares different methods for studying the relationship between lexical classifications from different corpus. We restrict here to classifications that use the method Reinert (1983, 1990). We try to make salient the proximities and distances between the identified themes and to specify the nature of these relationships. To do this, we compare the advantages and limitations of three methods: a correspondence analysis on the complete lexical table (Lebart & Salem, 1994), a measure of textual distanced using Labbé' index (Labbé & Labbé, 2003) on the sets of segments grouped in classes and use of Tgens (Lamalle & Salem, 2002) inspired by Marty, Marchand & Ratinaud (2012) work. We apply these methods to the study of the relationship between the contents of the websites of political parties (the National Front, the Socialist Party, Europe Ecologie Les Verts) and the analysis on overall debates of the National Assembly between 1998 and 2014 (Ratinaud & Marchand, 2015). The analysis are carried out from IRaMuTeQ and R software.

Key words: clustering, method Reinert, lexical distance, CA, TGen

Résumé

Ce travail compare différentes méthodes permettant d'étudier les relations entre des classifications lexicales, que nous restreindrons ici aux classifications utilisant la méthode Reinert (1983, 1990), issues de corpus différents. Nous chercherons à rendre saillantes les proximités et les distances entre les thématiques identifiées et à spécifier la nature de ces relations. Pour ce faire, nous comparerons les apports et les limites de 3 méthodes : une analyse factorielle des correspondances sur le tableau lexical complet (Lebart & Salem, 1994), une mesure des distances textuelles avec l'indice de Labbé (Labbé & Labbé, 2003) sur les ensembles de segments regroupés dans les classes et une utilisation des Tgens (Lamalle & Salem, 2002) inspirée des travaux de Marty, Marchand & Ratinaud (2012). Nous appliquerons ces méthodes à l'étude des relations entre les contenus des sites web de partis politiques (Parti socialiste, Europe Ecologie Les Verts, Front national) et une analyse sur l'ensemble des débats de l'Assemblée Nationale entre 1998 et 2014 (Ratinaud & Marchand, 2015). Les analyses seront réalisées à partir des logiciels IRaMuTeQ et R.

Mots clés : classification, méthode Reinert, distance lexicale, AFC, TGen.

* Ce travail a été réalisé dans le cadre du Labex SMS, portant la référence ANR-11-LABX-0066, il a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre du programme Investissements d'avenir portant la référence ANR-11-IDEX-0002-02.

1. Introduction

La diversification des origines possibles des corpus travaillés en Sciences Humaines et la relative simplification de leur constitution lorsqu'ils relèvent de données numériques amènent de plus en plus souvent les chercheurs à manipuler des ensembles textuels hétérogènes pour traiter une même problématique. Par exemple, la tentation est grande de comparer le discours

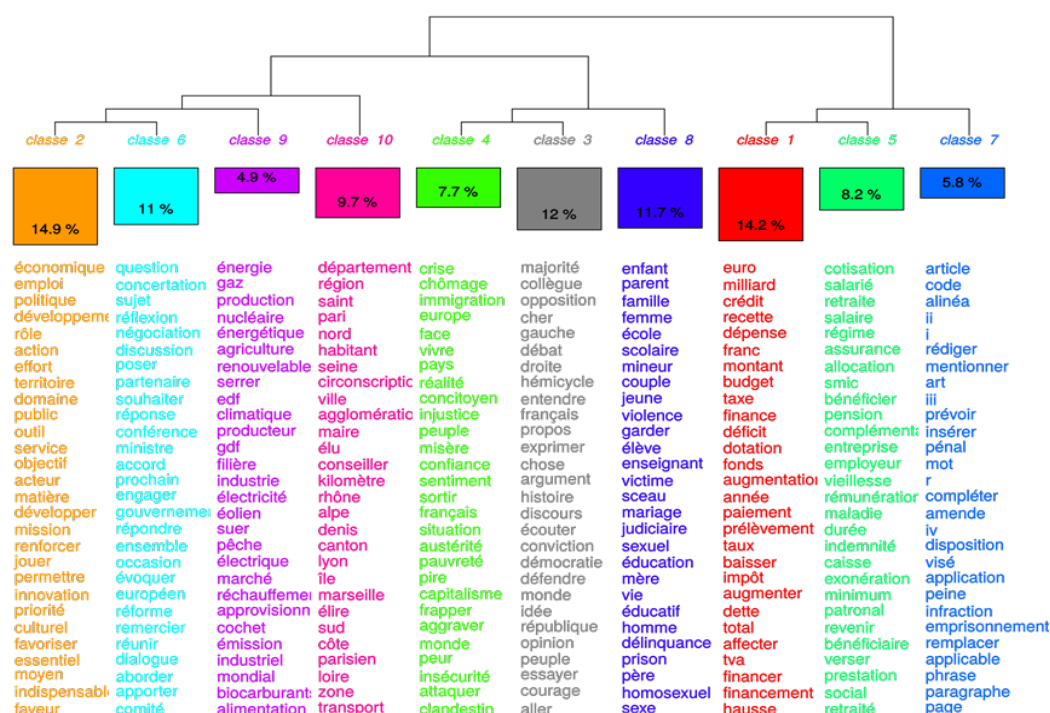
médiatique (par l'analyse d'un corpus de presse par exemple) avec celui de certains groupes sociaux (par l'intermédiaire du traitement d'entretiens semi-directifs ou par l'étude de production relevant des réseaux socio-numériques par exemple). Dans ces situations, il peut être envisagé de constituer un corpus unique réunissant l'ensemble des sources ciblées pour en réaliser une analyse unique. Toutefois, cette démarche ne semble raisonnable qu'à la condition où les documents assemblés respectent un certain niveau d'homogénéité (Pincemin, 2012). Lorsqu'ils relèvent de genres différents (discours oral vs discours écrit par exemple), de temporalité différente, qu'ils sont de dimension très éloignée ou que les thématiques dont ils traitent ne se recouvrent pas complètement, il est préférable de les envisager séparément. Si l'objectif de l'analyse est de dégager les thématiques qui composent les textes, les classifications relevant de la méthode Reinert (Reinert 1983, 1990 ; Ratinaud & Marchand, 2012) peuvent être convoquées. Il devient alors possible, par une approche qualitative, de comparer les profils des classes obtenues. L'objectif de notre étude est ici de proposer des méthodes pour accompagner cette démarche de comparaison afin de compléter l'analyse qualitative des profils de classe en mobilisant les outils proposés par la lexicométrie. Nous évaluerons donc les potentialités et les limites de trois techniques différentes : une analyse factorielle des correspondances sur le tableau lexical complet regroupant les classifications à comparer, une mesure des distances textuelles entre les classes avec l'indice de Labbé et la projection des TGen des formes sur-représentées dans les classes d'une analyse sur les profils des classes d'un corpus cible. Nous commencerons par décrire les corpus sur lesquels ces analyses seront réalisées : il s'agira d'étudier les relations entre le contenu de sites web de partis politiques (Parti Socialiste, Europe Ecologie Les Verts, Front National) et un corpus de débats à l'Assemblée Nationale entre 1998 et 2014 (Ratinaud & Marchand, 2015).

2. Description des corpus

L'analyse que nous utiliserons comme cible des comparaisons a été réalisée sur un corpus des débats à l'Assemblée Nationale (Ratinaud & Marchand, 2015). A l'origine, ce corpus se compose des 4009 sessions parlementaires qui ont eu lieu entre janvier 1998 et septembre 2014, à l'exception des sessions de 2001¹. Il est constitué d'environ 138 000 000 d'occurrences. Une première analyse sur ce corpus a permis de distinguer les classes de discours qui abordent les différentes thématiques sur lesquelles portent les textes de loi discutés en séance, de celles relatives à la façon d'élaborer ces textes (vocabulaire juridique, organisation des votes, réactions des députés aux interventions...). Nous avons extrait ces classes pour refaire une analyse uniquement sur les thématiques. C'est ce sous-corpus, d'environ 68 000 000 d'occurrences, que nous utiliserons ici. La classification que nous avons retenue présente 10 classes terminales qui sont encore fortement hétérogènes (chacune contient en fait différentes sous-thématiques). Les résultats sont présentés dans le dendrogramme 1.

¹ Le corpus a été extrait automatiquement du site web de l'assemblée nationale. Les sessions de l'année 2001 se présentaient dans un format différent des autres qui n'a pas permis leur extraction.

QUELQUES MÉTHODES POUR L'ÉTUDE DES RELATIONS ENTRE CLASSIFICATIONS DE CORPUS HÉTÉROGÈNES : APPLICATIONS AUX DÉBATS À L'ASSEMBLÉE NATIONALE ET AUX SITES WEB DE PARTIS POLITIQUES



Dendrogramme 1 : Classification du sous-corpus « thématique » de l'Assemblée Nationale

De gauche à droite sur ce dendrogramme, nous voyons une classe sur l'emploi, les territoires et le développement économique (classe 2), puis une classe sur la négociation et la concertation (classe 6), une classe autour de l'énergie, de l'agriculture et de l'écologie (classe 9) et une classe sur l'aménagement du territoire (classe 10). Une seconde branche du dendrogramme différencie une classe sur les crises et le chômage (classe 4), une classe sur les débats entre majorité et opposition (classe 3) et une classe sur la famille, l'éducation et la justice (classe 8). La dernière partie du dendrogramme rend compte d'une classe sur le budget, la finance et la fiscalité (classe 1), une classe sur le système de protection sociale (classe 5) et une petite classe contenant du lexique lié à l'élaboration des textes juridiques (classe 7).

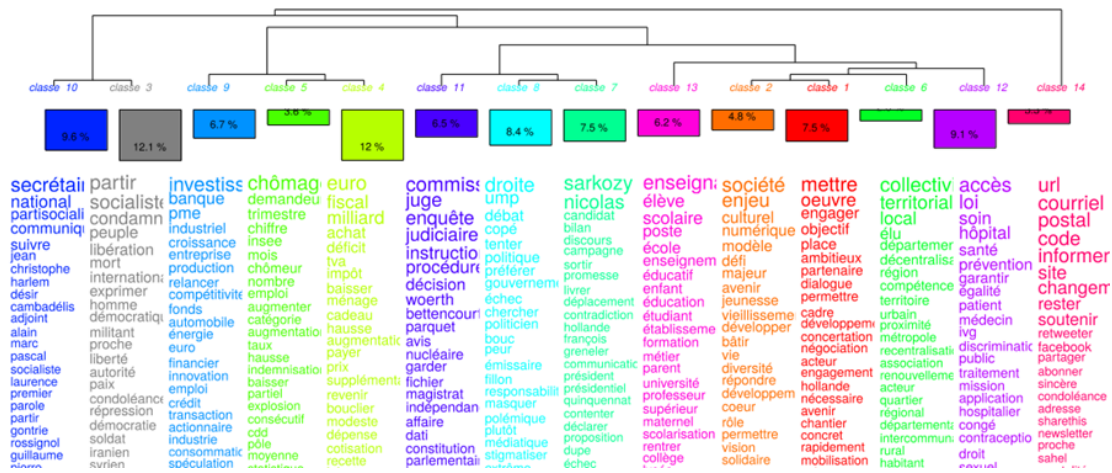
Les analyses que nous souhaitons comparer aux résultats obtenus sur l'Assemblée Nationale sont issues de corpus provenant de sites web de trois partis politiques : le Parti Socialiste, Europe Ecologie Les Verts et le Front National. Ils ont été extraits à l'aide de l'outil Gromoteur (Gerdes, 2014). L'objectif de l'analyse de ces sites web étant de déterminer les thématiques que privilégie chacun de ces partis, nous avons fait le choix de ne retenir que les documents relevant de la catégorie « actualités ». Cette catégorie, commune aux trois sites Web, est la seule qui présente ici un intérêt. Toutes les autres entrées proposées sur ces sites conduisent à des formulaires d'adhésion, à des descriptions formelles de l'organisation interne des partis ou à des outils « participatifs ». Pour le Front National, nous avons également retenu la catégorie « projet ». Le tableau 1 décrit succinctement les 4 corpus concernés. Il montre assez clairement l'hétérogénéité de ces corpus. Non seulement ces textes relèvent de genres différents, mais ils présentent par ailleurs des tailles et des temporalités qui ne sont pas comparables.

	Assemblée Nationale	FN	PS	EELV
textes	4006	3267	5419	2405
segments de texte	1867369	24945	41328	25948
occurrences	67989913	897173	1487840	924289
nombre de formes	147478	34383	33255	33479
hapax	50126	13728	11673	13066
nombre de classes	10	14	14	15
dates	1998-2014	07/2011-05/2015	12/2006-05/2015	02/2011-10/2015

Tableau 1 : Description des 4 corpus

2.1 Analyse du site du Parti Socialiste

Le dendrogramme 2 présente la classification retenue pour le site du parti socialiste.



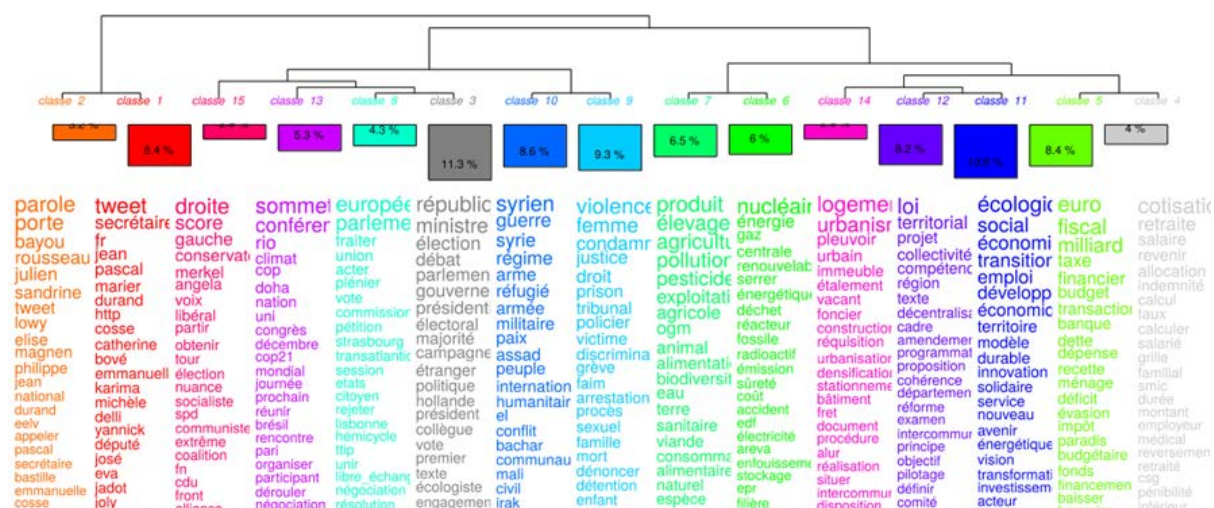
Dendrogramme 2 : Classification du site du Parti Socialiste

L'analyse met en évidence une classe listant des acteurs du parti (classe 10), proche d'une classe regroupant un lexique qui semble renvoyer aux hommages et aux condamnations liés à des événements internationaux. Nous trouvons ensuite 3 classes relevant de la dimension économique. La première (classe 9) est composée du lexique de la finance et de la production industrielle, la seconde (classe 5) traite spécifiquement de la question du chômage, enfin la troisième (classe 4) aborde les questions budgétaires et fiscales. Les classes 11, 8 et 7 regroupent les dépêches sur l'adversaire politique du PS : la droite. La classe 11 souligne la mise en avant des affaires judiciaires impliquant des membres de l'UMP, la classe 8 cible le parti et la classe 7 Nicolas Sarkozy. Le dendrogramme présente ensuite une classe sur l'éducation au sens large (classe 13), une classe traitant de questions sociétales (classe 2), puis une relative aux modalités de leur mise en œuvre (classe 1). Les deux classes suivantes abordent les collectivités territoriales (classe 6) et la politique de santé (classe 12). Enfin, la classe 14, qui est la première à se distinguer, contient le lexique « technique » typique des corpus extraits de site web. Il faut noter que la présence de cette classe est une conséquence du mode de construction du corpus. Nous avons utilisé le parser proposé par Gromoteur pour extraire le contenu textuel pertinent pour l'analyse des pages web enregistrées, ce qui revient

QUELQUES MÉTHODES POUR L'ÉTUDE DES RELATIONS ENTRE CLASSIFICATIONS DE CORPUS HÉTÉROGÈNES : APPLICATIONS AUX DÉBATS À L'ASSEMBLÉE NATIONALE ET AUX SITES WEB DE PARTIS POLITIQUES

en fait à éliminer les bannières, les menus etc... La qualité de cette extraction dépend du formatage du site.

2.2 Analyse du site d'Europe Ecologie Les Vert

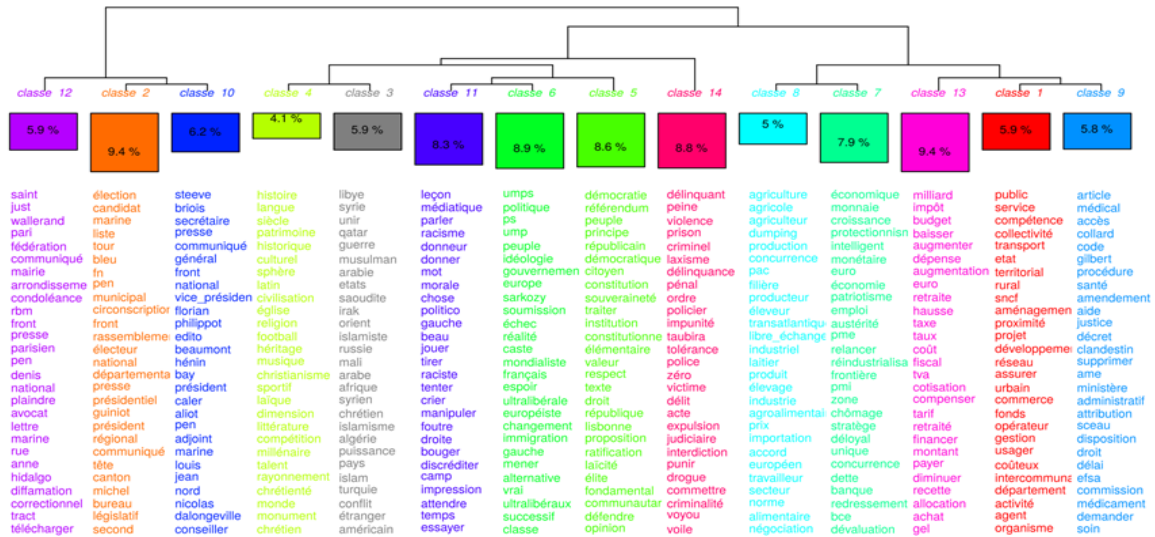


Dendrogramme 3 : Classification du site d'Europe Ecologie Les Verts

Cette analyse permet de distinguer, dans un premier temps, deux classes listant des noms de membres du parti et du vocabulaire relevant des réseaux socio-numériques (classe 2 et 1). Nous trouvons ensuite une classe commentant des résultats électoraux (classe 15), une classe qui porte sur les conférences internationales liées à l'écologie (classe 13), une classe sur le parlement européen (classe 8) et une classe sur le gouvernement (classe 3). La branche suivante du dendrogramme fait apparaître une classe sur le conflit syrien (classe 10) et une classe sur la justice qui semble plus particulièrement dénoncer des injustices (classe 9). Viennent ensuite une classe sur l'agriculture (classe 7) et une classe sur les énergies (classe 8). La dernière branche du dendrogramme présente une classe sur le logement (classe 14), une classe sur l'aménagement du territoire (classe 12), une classe sur la transition écologique au sens large (classe 11). Enfin, les deux dernières classes contiennent le vocabulaire économique : la classe 5 est caractérisée par le lexique lié au budget et à la fiscalité et la classe 4 par celui lié au système de protection sociale

2.3 Analyse du site du Front National

PIERRE RATINAUD, PASCAL MARCHAND

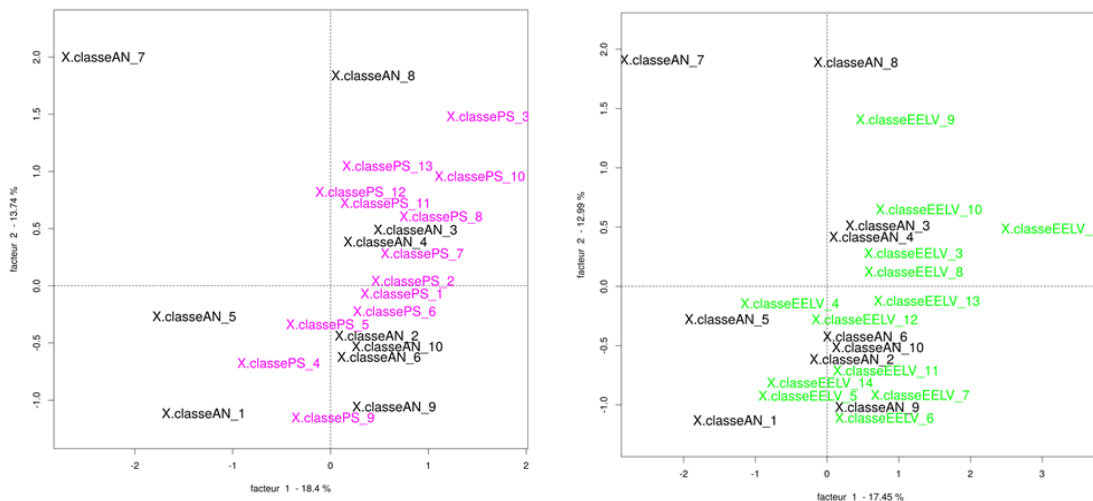


Dendrogramme 4 : Classification du site du Front National

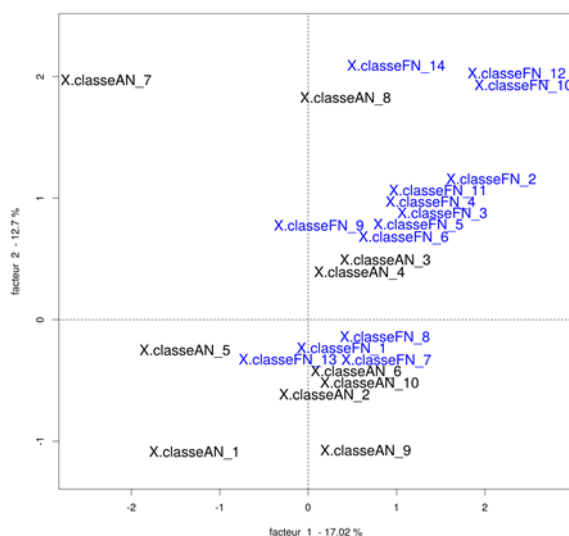
Comme dans les analyses précédentes, cette classification commence par différencier un groupe de classes dans lesquelles les noms propres sont surreprésentés : deux classes listent des acteurs du parti (classes 12 et 10) et une rend compte de résultats électoraux (classe 2). Le dendrogramme fait ensuite apparaître une classe sur l'histoire et la culture (classe 4), une classe sur le moyen orient et l'islamisme (classe 3), une classe sur les médias (classe 11), une classe sur l'« UMPS », une classe sur la démocratie et l'Europe (classe 5) et une classe sur la justice et la délinquance (classe 14). La dernière branche du dendrogramme présente une classe sur l'agriculture (classe 8), une classe sur l'économie (classe 7), une classe sur le budget et la fiscalité (classe 13), une classe sur les services publics (classe 1) et une classe sur la santé et la justice (classe 9).

3. Comparaison avec l'analyse factorielle des correspondances

La première analyse que nous proposons pour comparer les classifications obtenues sur les sites web des partis à celles de l'Assemblée Nationale est très classique en lexicométrie. Il s'agit d'une analyse factorielle des correspondances sur un tableau lexical complet (Lebart & Salem, 1994) qui contient chacune des 10 classes de discours issus du corpus sur l'Assemblée Nationale, auxquelles nous ajoutons les classes du corpus que nous souhaitons comparer. Pour cette analyse, nous avons conservé les formes pleines de fréquence supérieure ou égale à 5. Les graphiques 1, 2 et 3 rendent compte des résultats obtenus respectivement pour le Parti Socialiste, Europe Ecologie Les Verts et le Front National.



QUELQUES MÉTHODES POUR L'ÉTUDE DES RELATIONS ENTRE CLASSIFICATIONS DE CORPUS HÉTÉROGÈNES : APPLICATIONS AUX DÉBATS À L'ASSEMBLÉE NATIONALE ET AUX SITES WEB DE PARTIS POLITIQUES



Graphiques 1, 2 et 3 : plans factoriels des AFC sur les classes de l'Assemblée Nationale et respectivement du Parti Socialiste, d'Europe Ecologie Les Verts et du Front National.

La première remarque que nous pouvons faire sur ces résultats concerne la structuration de ces plans. Ici, la différence de taille entre les classes issues de l'Assemblée Nationale et les autres corpus est prégnante. Seules les premières semblent structurer le plan, car elles sont beaucoup plus « lourdes » que les autres. Nous pouvons donc constater que la position relative de ces classes ne change pas entre les analyses, comme si nous avions utilisé les classes des sites web en variables supplémentaires. Il apparaît également que les classes issues de l'analyse du site du Front National semblent sortir du champ créé par les classes de l'Assemblée Nationale.

4. Comparaison avec la distance de Labbé

L'indice de distance intertextuelle de Labbé (Labbé & Monière, 2000 ; Labbé & Labbé, 2003) est l'un des indices disponibles dans la littérature pour estimer dans quelle mesure deux textes sont proches ou différents du point de vue du lexique qui les compose. La question des distances intertextuelles est un champ fécond de la lexicométrie (Luong, 2003) et d'autres indices peuvent être mobilisés. L'utilisation de ce critère sera pour nous l'occasion de le tester dans un contexte très particulier (sur des classes de segments de texte et non des textes) et de l'implémenter en langage R pour le rendre disponible dans les outils se basant sur ce logiciel (IRaMuTeQ et TXM par exemple). Nous avons par ailleurs pleinement conscience que notre utilisation de cet indice s'éloigne souvent des préconisations des auteurs. Nous soulignerons ces écarts. Rappelons rapidement les principes de cette mesure. La distance entre les textes A et B, notée $D_{(a,b)}$ est obtenue par le calcul suivant :

$$D_{(a,b)} = \frac{\sum_{V \in A, V \in B} |F_{ia} - E_{ia(u)}|}{N_a + N'_b}$$

Nous supposons ici que le texte B a une taille supérieure au texte A. Dans ce calcul, les fréquences du lexique du texte B sont ramenées à ce qu'elles seraient si l'étendue du texte B était équivalente à celle du texte A (N_a). La fréquence attendue d'une forme i de B dans le texte A, notée $E_{ia(u)}$ dans la formule, est alors égale à la fréquence de la forme i dans B (F_{ib}) multiplié par le ratio de taille entre A et B :

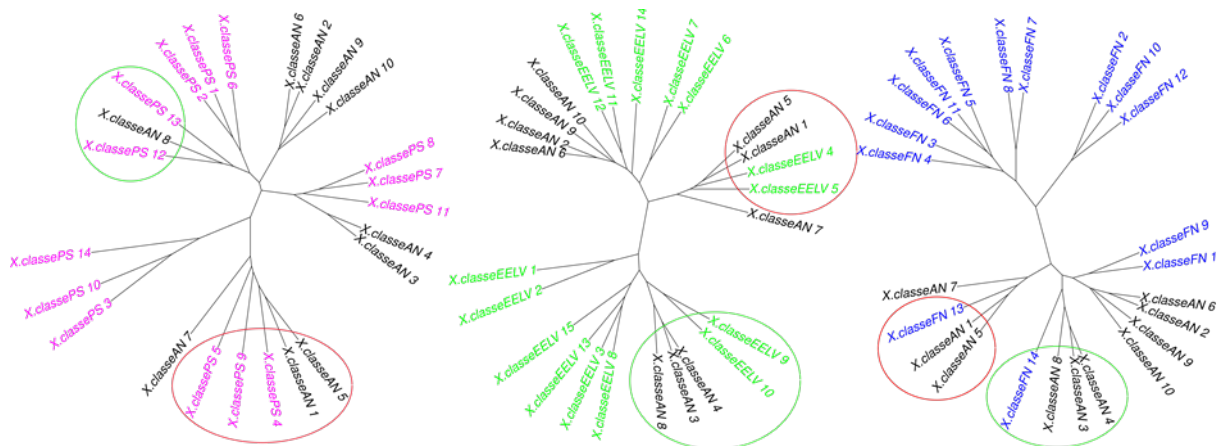
$$E_{ia(u)} = F_{ib} * U_{(a,b)} \text{ avec } U_{(a,b)} = U_a / U_b$$

On note N'_b l'étendu du texte B' , qui est la réduction du lexique de B en fonction de la taille de A . L'indice apprécie alors la distance entre la fréquence des formes dans A (F_{ia}) et celle dans B' . Ce calcul est appliqué à toutes les formes des textes aux exceptions suivantes :

- On ne considère qu'une seule fois les formes communes aux deux textes.
- On applique le calcul à l'ensemble des formes de A , mais seules les formes de B ayant une fréquence attendue dans A supérieure ou égale à 1 seront considérées.

Soulignons enfin que les auteurs déconseillent l'emploi de cet indice lorsque la différence de taille entre A et B dépasse un facteur 10 (Labbé & Labbé, 2003). Ce critère n'est pas respecté ici, le ratio de taille entre le corpus de l'Assemblée Nationale et le plus gros des corpus de site web, qui se retrouve à peu de choses près dans le rapport des tailles de classe, est de l'ordre de 50. Par ailleurs, nous n'appliquerons pas ce calcul à toutes les formes des classes, mais uniquement aux formes pleines, qui sont les seules à avoir structuré la classification. Enfin, nous avons limité le calcul aux formes présentant, dans l'ensemble des corpus², une fréquence supérieure ou égale à 5.

Les résultats que nous présentons sont, pour chaque comparaison, un graphique arboré d'une classification (méthode de Ward) sur la matrice des distances entre les classes.



Graphiques 4, 5 et 6 : représentation arborée d'une classification (méthode de Ward) sur les matrices de distances (indice de Labbé) entre les classes de l'Assemblée Nationale et respectivement, les classes du Parti Socialiste, d'Europe Écologie Les Verts et du Front National.

Cette analyse semble confirmer la première. Les distances entre les classes provenant du Front National et celles de l'Assemblée Nationale semblent supérieures à celles constatées pour les deux autres corpus. Il est intéressant de noter qu'à l'exception de la classe 14, toutes les classes du FN proches des classes de l'AN (classes 9, 1 et 13) sont celles qui présentent une surreprésentation des segments provenant du projet. On peut alors faire l'hypothèse que le projet du FN est rédigé avec un lexique plus proche du lexique parlementaire que ses autres productions. L'analyse permet également de mettre en évidence des thématiques communes : autour de la classe de l'AN sur la famille au sens large (entouré en vert) qui est proche des classes sur l'éducation (classe 12) et sur la santé (classe 13) pour le Parti Socialiste, de la

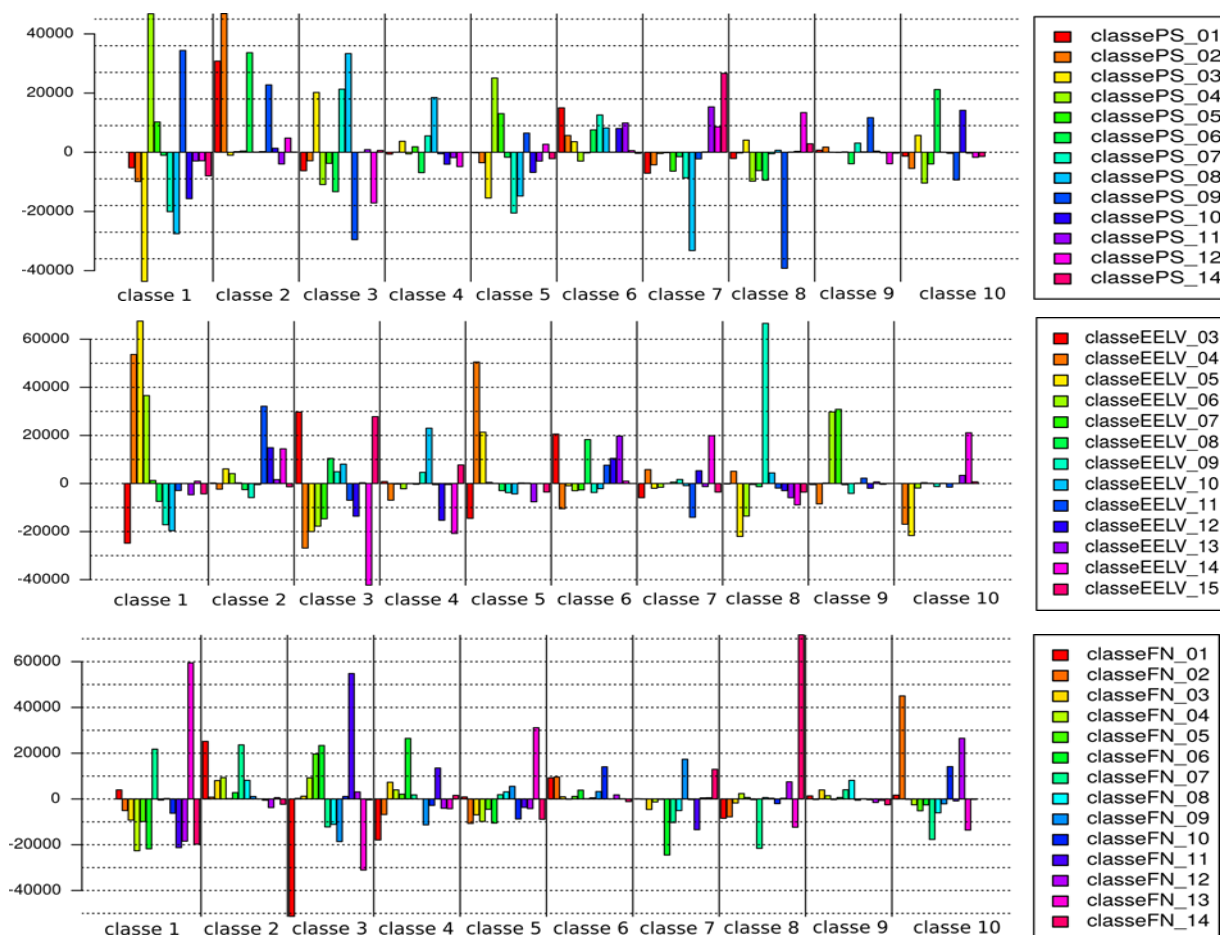
² Pour participer à cette analyse, une forme devait avoir une fréquence de 5 dans les corpus constitués des classes de l'assemblée nationale et des classes du corpus que l'on veut comparer.

QUELQUES MÉTHODES POUR L'ÉTUDE DES RELATIONS ENTRE CLASSIFICATIONS DE CORPUS HÉTÉROGÈNES : APPLICATIONS AUX DÉBATS À L'ASSEMBLÉE NATIONALE ET AUX SITES WEB DE PARTIS POLITIQUES

classe sur la dénonciation des injustices et sur la crise Syrienne pour le EELV (classe 9 et 10) et de la classe sur la délinquance pour le FN (classe 14). Une même proximité est visible pour les classes abordant les questions budgétaires et de protection sociale (entouré en rouge).

5. Comparaison avec les TGen

Nous reprenons ici une stratégie proposée par Marty, Marchand & Ratinaud (2012). Il s'agit de projeter un ensemble de lexique regroupé dans un TGen sur les profils de la classification cible. Nous avons extrait de chacun des profils des classes des corpus que nous souhaitons comparer, la liste de toutes les formes présentant un χ^2 de liaison aux classes d'au moins 15,2 (ce qui correspond à seuil de significativité $< 0,0001$ pour un ddl de 1). Ces TGen ont été nommés avec le nom de la classe qu'ils représentent. Ils sont ensuite projetés sur la classification de l'Assemblée Nationale. Les χ^2 de liaison des TGen aux classes cibles sont calculés sur la même logique que dans les profils : ils marquent la sur ou la sous-représentation des segments de texte contenant au moins une fois l'une des formes du TGen dans la classe.



Graphiques 7, 8 et 9 : projection des TGen des classes du Parti Socialiste, d'Europe Ecologie Les Verts et du Front National sur les classes de l'Assemblée Nationale.

Nous retrouvons dans cette analyse des éléments d'interprétation déjà suggérés par la précédente. La classe de l'AN sur la famille (classe 8) est de nouveau marquée par les classes du PS liées à l'éducation et à la santé (classe 12 et 13), par la classe d'EELV liée à la

dénonciation d'injustice (classe 9) et par la classe du FN sur la délinquance (classe 14). Les classes de l'AN traitant des questions économiques, budgétaires et de la protection sociale (classes 2, 1 et 5) sont toutes représentées par des thématiques sur les sites web des partis. Enfin, la classe de l'AN sur l'énergie et l'écologie (classe 9) est représentée par 2 classes de site d'EELV (classes 6 et 7) alors qu'elle est absente de l'analyse du FN.

6. Conclusion

Nous noterons avec Brunet (2003) que malgré la différence des approches proposées et l'hétérogénéité des corpus comparés, « la convergence est au rendez-vous » (p. 60). Ces analyses permettent, nous semble-t-il, de confirmer les impressions que laisse le parcours des profils des classes. Nous les avons testées sur un cas « critique » et nous pensons qu'elles pourront alors en grande partie répondre aux problématiques rencontrées par les chercheurs sur des corpus qui présenteraient a priori moins de différences que ceux que nous utilisons ici. Une dernière difficulté demeure pourtant : même si ce travail a permis la mise en place de quelques routines, l'utilisation concrète de ces méthodes reste délicate. Il appartient alors aux développeurs d'outils de lexicométrie de mettre en place les processus pour simplifier et automatiser ces comparaisons. Elles devront par ailleurs être complétées par des outils permettant de spécifier le lexique responsable des proximités entre thématiques, ce qui est tout à fait envisageable pour les distances de Labbé et pour les TGenS.

7. Bibliographie

- Brunet, E. (2003). Peut-on mesurer la distance entre deux textes ?, *Corpus*, (2) : 47-70.
- Labbé, D., & Monière, D. (2000). La connexion intertextuelle. Application au discours gouvernemental québécois. In M. Rajman & J.-C. Chappelier (Eds.), *Actes des 5èmes Journées Internationales d'Analyse statistique des Données Textuelles*, Lausanne : EPLF, 85-94.
- Labbé, C., & Labbé, D. (2003). La distance intertextuelle. *Corpus*, (2) : 95-117.
- Lamalle C., & Salem A. (2002). Types généralisés et topographie textuelle dans l'analyse quantitative des corpus textuels. *Actes des 6e Journées d'Analyse des Données Textuelles*, St Malo.
- Lebart, L., & Salem, A. (1994). *Statistiques textuelles*. Paris, France: Dunod.
- Luong, X. (2003). La distance intertextuelle. *Corpus*, 2.
- Marty, E., Marchand, P., & Ratinaud, P. (2013). Les médias et l'opinion: éléments théoriques et méthodologiques pour une analyse du débat sur l'identité nationale. *Bulletin de Méthodologie Sociologique*, 117(1) : 46-60.
- Pincemin, B. (2012). Hétérogénéité des corpus et textométrie. *Langages*. 187 : 13-26.
- Ratinaud, P., & Marchand, P. (2012). Application de la méthode ALCESTE à de "gros" corpus et stabilité des "mondes lexicaux" : analyse du "CableGate" avec IRaMuTeQ. In *Actes des 11eme Journées internationales d'Analyse statistique des Données Textuelles*, 835-844.
- Ratinaud, P., & Marchand, P. (2015). Des mondes lexicaux aux représentations sociales. Une première approche des thématiques dans les débats à l'Assemblée nationale (1998-2014). *Mots. Les Langages Du Politique*, 2015(108) : 57-77.
- Reinert, M. (1983). Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte. *Les cahiers de l'analyse des données*, VIII(2) : 187-198.
- Reinert, M. (1990). ALCESTE : Une méthodologie d'analyse des données textuelles et une application : Aurélia de Gérard de Nerval. *Bulletin de Méthodologie Sociologique*. 26 : 24-54.