

Plongée dans le lexique du Conspirationnisme dans la presse nationale française

Clara Bordier¹, Matej Martinc², Marceau Hernandez¹, Senja Pollak², Gaël Lejeune¹

¹ CERES (Centre d'Expérimentation en méthodes numériques pour les recherches en SHS {name}.{surname}@sorbonne-universite.fr

² Jozef Stefan Institute - {name}.{surname}@ijs.si

Abstract

In this article, we present a diachronic study of terms relating to conspiracism in a corpus of the French general press (*Libération*, *Le Monde*, *Le Figaro*), from 2002 to 2019. Insofar as the press constructs shared social representations (Moirand, 2007), the aim here is to analyse the way in which conspiracy has become established in the media in relation to major events (attacks, elections, wars, etc.). Our aim is to gradually identify how newspapers position themselves in relation to what is labelled as 'false', in order to analyse the ways in which a legitimate and shared 'truth' is constructed. To this end, we use diachronic word embeddings (introduced by Hamilton et al. in 2016). These make it possible to highlight strong lexical trends from 2002 onwards that reflect an evolution in the discourse around these subjects. By cross-referencing this information with the political trends in the press sources, we are able to illustrate some particularly interesting semantic changes. We therefore propose to view diachronic word plunging as a way of enriching corpora in order to facilitate text mining with textometric tools, and we also present an automatic tool developed for this purpose.

Key words : Word Embeddings, Diachronic Embeddings, Semantic Shift, Press Articles, Discourse Analysis, Conspiracy Theories.

Résumé

Dans cet article, nous présentons une étude diachronique des lexèmes se rapportant au conspirationnisme dans un corpus de presse généraliste française (*Libération*, *Le Monde* et *Le Figaro*) avec des articles publiés entre 2002 à 2019. Dans la mesure où la presse construit des représentations sociales partagées (Moirand, 2007), l'objectif est ici d'analyser la manière dont le complotisme s'est installé dans l'espace médiatique par rapport à de grands événements (attentats, élections, guerres, etc.). Notre but est de repérer progressivement le positionnement des journaux face à ce qui est étiqueté comme "faux", afin d'analyser les modalités de la construction d'un "vrai" légitime et partagé. Dans cet objectif, nous exploitons des plongements de mots diachroniques (Diachronic Embeddings, introduits par Hamilton et al. en 2016). Ceux-ci permettent de faire ressortir des tendances lexicales fortes à partir de 2002 qui traduisent une évolution du discours autour de ces sujets. En croisant ces informations avec les tendances politiques des sources de presse, nous parvenons à illustrer des changements sémantiques particulièrement intéressants. Nous proposons ainsi de voir les plongements de mots diachroniques comme une manière d'enrichir les corpus afin de faciliter la fouille avec des outils textométriques et nous présentons également un outil automatique développé dans ce but.

Mots clés : Plongements Lexicaux, Plongements Diachroniques, Changement sémantique, Articles de Presse, Analyse de Discours, Théories du Complot.

1. Introduction

Les théories du complot sont un phénomène de société contemporain fortement représenté dans l'espace médiatique depuis plusieurs années, avec probablement un accroissement important à compter des attentats du 11 septembre 2001. Si ces théories font circuler de nombreux discours, farfelus et erronés, nombres de discours plus circulent également autour de celles-ci. On cherche régulièrement à savoir d'où elles viennent, comment elles circulent, qui y adhère, quels sont leurs effets, comment stopper leur propagation (Bonnet, Mercier, Siouffi, 2022). En revanche, peu d'études s'intéressent au rôle des médias généralistes dans la transformation de ce phénomène en véritable problème social. En effet, les médias généralistes, en construisant "des représentations sociales partagées du réel" (Moirand, 2007), participent à l'effort de définition et de dénomination du phénomène (Calabrese, 2018). La presse, bien qu'elle ne représente plus le type de consommation médiatique le plus généralisé, reste du moins la plus légitime à relater les faits, les bonnes informations. Si l'on souhaite comprendre comment les théories du complot sont venues à questionner les régimes de vérité dans la société, alors, examiner les discours de la presse nationale qui fleurissent sur la question permettent de comprendre à quel phénomène nous sommes confrontés.

La recherche présentée ici s'inscrit dans un travail plus large qui cherche à comprendre grâce à des méthodes à la fois quantitatives et qualitatives, comment la presse généraliste donne de l'autorité à son statut d'organe social légitime à dire le vrai. En définissant ce qui est considéré comme « faux », la presse met en capacité la désignation ou l'ancrage du « fait » journalistique comme relevant de la « vraie information ». En adoptant une approche constructiviste, nous partons du postulat selon lequel les théories du complot ne relèvent pas d'un problème psychologique (Cueille, 2020) ou cognitif (Bronner, 2013), individuel menant à la déviance, mais bien d'un problème systémique et contemporain relatif aux régimes de vérité qui se matérialisent dans les médias. Il ne s'agit pas de pointer du doigt la presse généraliste comme étant elle-même un maillon de diffusion des théories du complot en les mettant en lumière. Cependant, parce les journalistes, se sentent directement menacés dans leur autorité à dire le vrai (Bratich, 2020), ils participent à faire d'un phénomène anecdotique un problème de société à part entière.

Nous présentons dans cet article des explorations quantitatives qui permettent de repérer les évolutions diachroniques de la connotation des termes associés à la thématique des théories du complot. Nous souhaitons via cette méthode comprendre comment le discours autour des théories du complot a évolué dans le temps. Nous partons du postulat selon lequel les termes se rapportant au conspirationnisme ne constituent pas des événements médiatiques, mais s'incarnent comme des formules, des adjectifs, ou des tournures de phrases qui viennent qualifier ces événements. Ainsi, nous pouvons repérer, selon des périodes marquantes, comment les grands événements médiatiques (attentats, élections, guerres) ont transformé le discours sur le phénomène. L'avantage des méthodes lexicographiques diachroniques, réside dans le fait de pouvoir obtenir une vision synchronique des résultats, avant de plonger dans le détails des textes de manière qualitative. Ainsi, nous pouvons d'abord repérer des grandes périodes, avec des pics de publications, des thématiques saillantes et surtout, des transformations définitionnelles. Comment s'est développé le discours autour des théories du complot dans la presse généraliste française ? À quels événements renvoient les changements sémantiques au fil des années ? Quels sont les mots qui se transforment le plus ?

Dans la suite de cet article, nous reviendrons premièrement sur la constitution du corpus, la manière dont nous nous sommes appropriés les techniques de détection de changements sémantiques (Montariol et al. 2021) et l’apport de la méthode des plongements lexicaux pour explorer de manière quantitative notre objet. Ensuite, nous verrons les résultats que nous offrent ces méthodes, tout d’abord en observant quels sont les termes pour lesquels les changements sémantiques sont les plus marquants. Puis, nous détaillerons les évolutions sémantiques de quelques termes clés de notre étude. Enfin, nous récapitulerons les apports, les limites et les perspectives de cette approche par rapport à un travail d’analyse de presse.

2. Constitution des corpus et extraction des variations sémantiques

Notre corpus d’analyse est constitué d’articles de presse de trois titres de la presse nationale (*Le Monde*, *Le Figaro*, *Libération*) sur la période 2002-2019, récoltés sur la plateforme *Europresse* et mis en forme en XML grâce à l’outil *Europarser*¹. Nous nous sommes arrêtés avant la pandémie de Covid19, qui a permis à une multitude de théories du complot de fleurir, puisque nous souhaitons surtout comprendre l’évolution puis l’installation du problème dans le paysage médiatique français. Nous avons choisi des articles parus à partir de 2002², avec l’hypothèse d’un traitement médiatique accru sur la question, suite aux attentats du 11 septembre 2001. En France, la question des théories du complot connaît un intérêt grandissant dans les journaux vers 2002, suite à la publication du livre de Thierry Meyssan, “*L’effroyable imposture*”, dans lequel l’auteur accuse la CIA d’avoir fomenté les attentats. Avant, nous n’avons pas trouvé d’articles dans la presse IGP sur la question, du moins pas sur *Europresse*. Aussi, nous sommes conscients que le corpus ainsi constitué ne représente pas de manière exhaustive le traitement médiatique des théories du complot. Néanmoins, il permet d’avoir une vision assez large de la manière dont le problème s’est construit dans la presse française. Les mots clés d’entrée pour la collecte sont : « conspirationniste(s) », « conspirationnisme(s) », « complotiste(s) », « complotisme(s) » et la formule « théorie(s) du complot ». Au total, 1481 articles de presse forment ce corpus d’analyse séparé en trois périodes distinctes :

- 2002 à 2013: 459 articles
- 2014 à 2015: 429 articles
- 2016 à 2019: 593 articles

Ces périodes ont été choisies pour deux raisons : d’abord par souci de faisabilité statistique, il fallait disposer de suffisamment de données pour avoir des résultats pertinents et obtenir des sous-corpus de taille comparable. Ensuite, parce qu’un travail diachronique avait déjà été réalisé en analysant les pics de publication d’articles et il se trouve que la période entre 2002 et 2014 ne comporte pas tant d’articles sur la question des théories du complot. En revanche, à partir des attentats de 2015, nous avons remarqué une véritable prise en considération de la part de la presse de référence du problème complotiste en France. Il est donc pertinent de marquer la séparation des données à partir de ces dates, puisqu’elles sont constitutives de moment de rupture dans le traitement médiatique du phénomène. Aussi, la période qui suit ne sera plus autant axée sur les attentats que sur les élections présidentielles américaines, avec l’émergence de Trump comme acteur principal de la désinformation dans les démocraties libérales (Harjuniemi, 2022). Afin de vérifier si les changements sémantiques détectés étaient pertinents, il nous a fallu former un corpus de contraste. Pour constituer ce corpus de

¹<https://ceres.huma-num.fr/europarser/>

² Pour conserver une représentation équilibrée de chaque année, nous avons exclu l’année 2001

référence, nous avons donc sélectionné pour les trois mêmes journaux, les mille premiers articles parus chaque année de la période 2002-2019. Ce corpus de référence est destiné à établir si les changements sémantiques détectés sont spécifiques au sujet des théories du complot ou se retrouvent également dans la presse en général. Il est nettement plus fourni que notre corpus d'analyse (de l'ordre de 15 fois plus soit plus qu'un ordre de grandeur) de manière à faciliter l'élimination du bruit informationnel de la période. En ce sens nous nous éloignons de travaux récents en détection de tendance émergent (*Emerging Trend Detection*) où le corpus de contraste est constitué par paliers temporels (Dang et al. 2016, Liu et al. 2020, Nguyen et al. 2023).

Pour détecter les changements sémantiques dans le corpus, nous avons adapté à la langue française une méthode de détection du changement sémantique en diachronie qui exploite des plongements de mots (Montariol et al., 2021). Cette adaptation a principalement consisté à utiliser le modèle français de Spacy³ pour préparer le corpus (la méthode nécessitant lemmatisation et étiquetage morpho-syntaxique) et à utiliser Camembert (Martin et al. 2020) comme représentation de base⁴. Pour chaque nom présent dans les trois périodes du corpus d'analyse, nous calculons des plongements contextuels. Ces plongements sont ensuite regroupés grâce à l'algorithme k-means. Nous avons fixé à cinq le nombre de clusters puisque empiriquement, nous avons observé que c'était un bon compromis entre généralisation (avec moins de clusters les évolutions sémantiques étaient moins visibles) et spécialisation (pour de nombreux mots avoir plus de clusters était inutile faute d'un nombre d'occurrences suffisant).

Nous avons utilisé deux techniques pour analyser ces corpus. La première consiste à regarder dans chacun des deux corpus les dix mots qui subissent le plus de transformation, et les dix mots qui en subissent le moins. Nous avons également vérifié si les dix mots qui se transforment le plus dans le corpus d'analyse se retrouvent également dans le corpus de référence. Pour ces mots, nous avons observé la manière dont s'opèrent les changements sémantiques dans les deux corpus. Nous avons décidé d'y ajouter un regard plus qualitatif. En effet, à partir de méthodes connexes sur les mêmes corpus, nous avons pu repérer, notamment grâce à des clusters de mots, quels termes reviennent régulièrement et sont constitutifs d'un champ lexical en rapport avec les théories du complot dans la presse de référence. Nous avons ainsi observé la manière dont les clusters qui se rapportent à ces termes avec la méthode des plongements sémantiques évoluent dans le temps, une explication détaillée de la méthode est donnée dans (Martinc et al. 2024).

3. Analyse des évolutions sémantiques

Dans le tableau 1 (page suivante) nous présentons les dix mots qui présentent le plus de changements sémantiques entre les trois périodes décrites précédemment, dans notre corpus d'analyse et notre corpus de référence. Par soucis de faisabilité nous nous sommes limités à ces dix mots les plus significatifs.

Nous observons dans ce tableau une nette différence entre les mots qui se transforment le plus dans le corpus principal et le corpus de référence, puisqu'il n'y a pas de mots en commun. Le corpus présente donc bien à la fois un lexique à part entière avec des évolutions propres,

³<https://spacy.io/>

⁴ Le corpus de référence étant trop petit pour apprendre un modèle représentatif nous avons choisi l'adaptation (*fine tuning*).

détachées du lexique de l’agenda médiatique plus général. Nous pouvons ainsi noter la singularité des articles mobilisant des termes en rapport avec le complotisme, par rapport aux mots les plus changeant dans le paysage médiatique en général. Également, les termes pivots qui se transforment le plus au fil des années et qui nous semblent les plus pertinents à analyser par rapport à notre objet de recherche sont : “culte”, “revendication”, “finance”, “certitude” et “adhésion”. En effet, ces termes ont une charge symbolique, politique et sociale forte. “culte” renvoie à la religion, ou à une idéologie. “Revendication” a une portée politique et sociale, “finance” fait référence à un système social particulièrement critiqué, ici le capitalisme, et enfin “certitude” renvoie à la question de la vérité. Ce sont donc les transformations affiliées à ces termes que nous allons observer. Dans la série de figures que nous allons présenter nous montrerons pour différents lexèmes l’évolution des 5 clusters de termes associés.

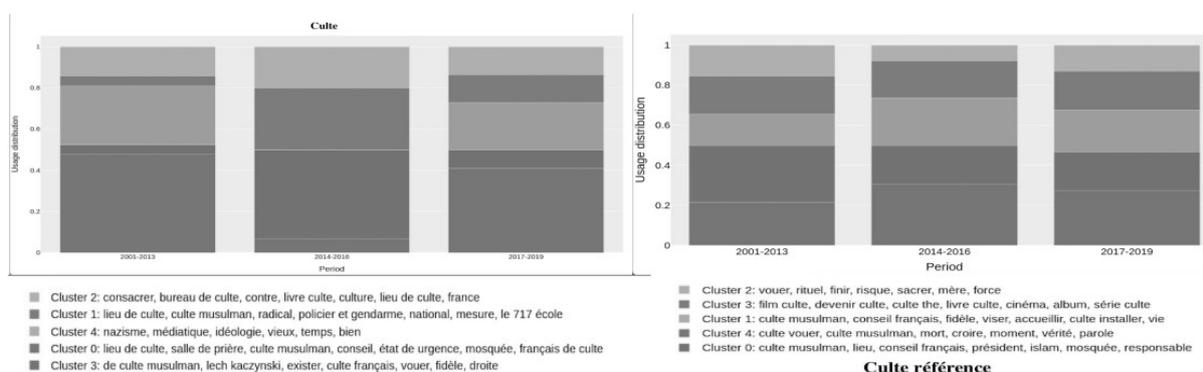


Figure 1: Évolution des termes associés à “Culte” dans les trois périodes de notre corpus d’analyse (à gauche) et dans notre corpus de référence (à droite).

10	approche	Euro
----	----------	------

Tableau 1: Mots pivots présentant le plus de changement sémantique entre 2002 et 2019

Sur la figure 1, nous voyons que le terme “culte” est majoritairement affilié aux pratiques religieuses musulmanes. Mais en y regardant de plus près, nous remarquons que chaque cluster à une particularité liée aux événements majeurs de la période. De 2001 à 2013, les mots qui se rapprochent le plus de “culte” font à la fois référence à la religion : “fidèle”, “de culte musulman”, mais aussi à un homme politique polonais : “lech kaczynski”, et à l’extrême droite. Aussi un autre cluster renvoie au “nazisme”, à “l’idéologie” et au “vieux”, “temps”. Il s’agit ici du lien historique entre les théories du complot et l’antisémitisme. De 2014 à 2016, l’accent est mis sur deux grandes institutions de la république française : “police et gendarme” et “717 école” ou encore “état d’urgence”, “conseil”. La thématique qui ressort de ce cluster est celle des attentats de 2015 et plus précisément les mesures anti-attentat mises en place par l’État pour lutter contre l’islam “radical”. En 2017 et 2019, la question des attentats et de la sécurité intérieure se tarit et nous retrouvons de nouveau le *cluster 4* et le *cluster 3* qui dominant largement. Dans le corpus de référence, “culte” est également corrélé à la religion musulmane. Il n’y a pas de spécificité entre le corpus d’analyse et de référence. La religion

musulmane est une thématique de prédilection pour la presse généraliste, surtout durant, et depuis, les attentats de 2015, néanmoins, nous ne retrouvons pas de mots en lien avec la droite ou l'extrême droite.

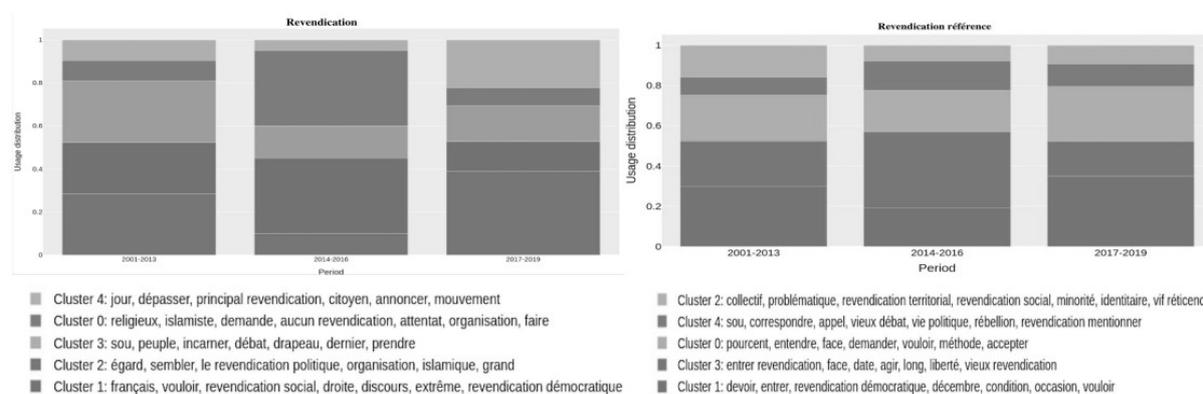


Figure 2: Évolution des termes associés à “Revendication” dans les trois périodes de notre corpus d’analyse (à gauche) et dans notre corpus de référence (à droite).

En ce qui concerne le terme “revendication”, la figure ci-dessous présente des clusters proches de revendications politiques liés aux termes “droite”, “extrême droite” ou “peuple”, encore une fois, de 2001 à 2013 et de 2017 à 2019. C’est lors de la période des attentats que nous retrouvons la question de l’islam et des attentats. Dans le corpus de référence en revanche, le terme “revendication” renvoie majoritairement à des questions sociales politiques : “minorités”, “rébellion”, “collectif”, sont des termes le plus souvent associées au mouvement sociaux étiquetés à gauche de l’échiquier politique. Les théories du complot semblent donc spécifiquement liées à des thématiques revendicatives émanant de courants politiques de la droite de l’échiquier politique voire globalement religieuses. Ces résultats font sens dans la mesure où les communautés politiques de droite ou d’extrême droite font ou ont fait circuler nombre de théories du complot (Giry, 2017). Nous pouvons même émettre l’hypothèse selon laquelle la thématique du “grand remplacement” est la plus adéquate pour comprendre l’évolution du terme “revendication” du corpus d’analyse, notamment parce que le terme “grand” est présent dans le *cluster 2*, majoritaire durant la période des attentats. Il est plus difficile de comprendre à quel événement renvoie le cluster 1, majoritaire durant la dernière période. En effet, les termes qui y sont affiliés pourraient renvoyer à la fois aux élections américaines de 2017, mais aussi au mouvement social des Gilets Jaunes en France. Pour préciser les résultats, un travail qualitatif, dans le texte, est alors nécessaire.

Le terme “finance” est intéressant car il change vraiment de signification au fil du temps. Dans la première partie du corpus, il est lié à la “finance internationale”, “finance mondiale” mais aussi à “juif”. On retrouve ici une des thématiques principales du conspirationnisme qui est celle de l’antisémitisme et de la finance soupçonnée de relever d’un système de domination lié aux communautés juives. Néanmoins, le terme apparaît seulement dans la première période de récolte. Ceci s’explique notamment par la transformation et la diversification des thématiques liées aux théories du complot. Si historiquement elles se forment en partie sur de l’antisémitisme, il s’agit tout de même de récits protéiformes qui évoluent avec le temps et les

grands événements politiques et sociaux. Ici par exemple, une période de crise économique vient construire un ennemi différent de celui d’hier : la finance mondiale, ou le libéralisme. En effet, durant la période de 2014 à 2015, ce sont des thématiques liées à des crises économiques qui prennent le dessus. Dans le corpus de référence, le terme finance est associé aux finances publiques, françaises, liées à l’école publique, privée et à la banlieue. Ce qui confirme bien la thématique de l’antisémitisme propre au conspirationnisme, c’est la prégnance du mot “juif” par rapport au corpus de référence. Aussi, le corpus d’analyse montre des termes corrélés à une finance mondiale privée et moins française et publique.

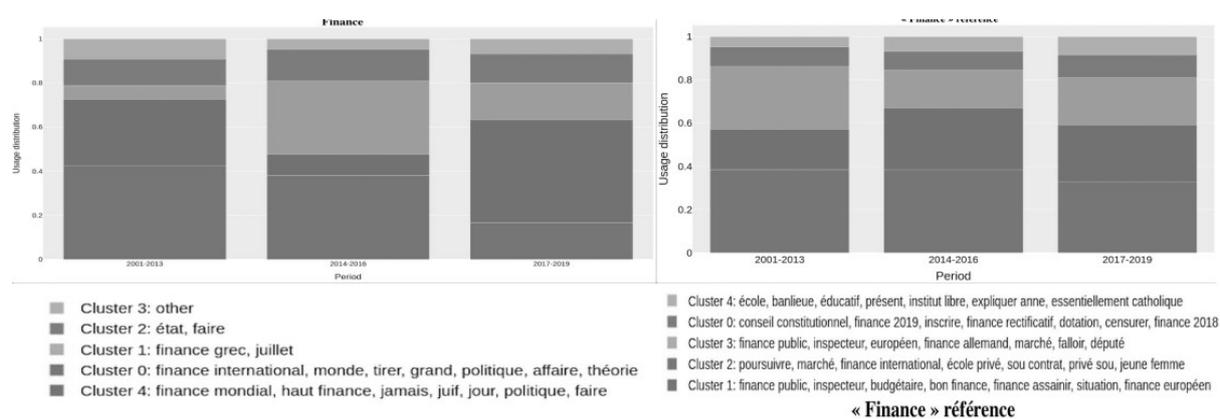


Figure 3: Évolution des termes associés à “Finance” dans les trois périodes de notre corpus d’analyse (à gauche) et dans notre corpus de référence (à droite).

Enfin pour le terme “certitude” (Figure 4), c’est bien la question de la vérité qui se pose, du moins dans le corpus d’analyse. De 2001 à 2013, les résultats montrent une certaine défiance vis-à-vis de la production de la certitude : “doute”, “fragilité”, “croire”. Mais à partir de 2014, ce sont des termes moins présents quoique tout de même dans les clusters : “liberté”, “faire”, “prendre”, “refuser”. À chaque fois la question de la certitude est liée à la politique, l’électorat et au jugement. Dans le corpus de référence on parle également d’un “manque” de certitude, ou d’une certitude “ébranlée”. Par ailleurs, dans le corpus de référence, le terme n’évolue pas tant au fil des années, les clusters de mots qui s’y rapportent restent les mêmes. Ce mot est particulièrement important pour cette recherche puisqu’il parle directement des théories du complot et du rapport au discours de vérité, en ce sens où la question se pose autour de ce que l’on sait d’immuable. La certitude c’est ce qui est fermement installée, c’est être sûre. Ici, le conspirationnisme ébranle cette certitude et par extension, ce que l’on décrit comme étant vrai. On notera que seuls 4 clusters apparaissent sur la visualisation car le cluster 0 est très petit (moins de 1 % des occurrences). ?

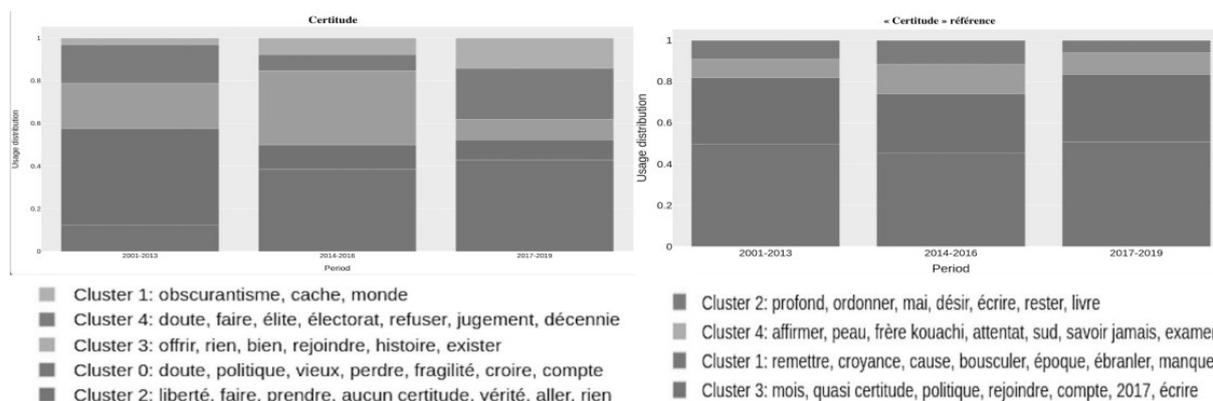


Figure 4: Évolution des termes associés à “Certitude” dans les trois périodes de notre corpus d’analyse (à gauche) et dans notre corpus de référence (à droite).

Malgré l’intérêt de cet outil pour la mise en visibilité de mots clés et leurs transformations sémantiques au cours du temps, il est parfois difficile de différencier les clusters les uns des autres. En effet, plusieurs clusters semblent à première vue proposer des transformations, mais les mots qui les constituent sont en fait des synonymes, des variations légères dans l’orthographe que la lemmatisation n’a pas permis de faire disparaître. On a aussi parfois le même terme dans différents clusters. Par exemple, dans les clusters obtenus pour le terme “certitude” (Figure 4), “doute” est à la fois dans le *cluster 4* et le *cluster 0* du corpus d’analyse, on peine donc parfois à trouver des termes véritablement spécifiques. Les différences, lorsqu’elles sont saillantes, se font parfois au mot près. Pour certains groupes de mots, il n’y en a même pas. Ces redites sont-elles liées au trop faible nombre d’articles, à un problème de paramétrage, ou tout simplement à des termes qui ne changent pas fondamentalement au fil des années ? Toujours est-il que cette méthode nécessite des analyses complémentaires, plus qualitatives, pour comprendre plus avant les enjeux des évolutions sémantiques propres aux théories du complot dans la presse généraliste.

4. Conclusion et perspectives

L’utilisation des plongements lexicaux diachroniques pour notre terrain d’enquête sur la construction du discours médiatique autour des théories du complot, nous a permis d’avoir une vision synchronique et diachronique des termes saillants. Plus que de nous offrir des découvertes inattendues, cette méthode nous a permis de confirmer des hypothèses soulevées par un travail préalable qualitatif de thématisation des articles. D’abord, le conspirationnisme n’est pas un événement médiatique en tant que tel mais bien une manière de qualifier ce qu’il se passe durant des grands événements. Les attentats, les guerres, la politique, la finance cristallisent des thématiques conspirationnistes. Les changements sémantiques se produisent surtout pendant la période des attentats de 2015, avec une focalisation quasi systématique sur l’islam. L’événement médiatique principal qui permet d’opérer ces changements est l’attentat, et l’usage des lemmes de “complot” désigne l’islam comme principal levier du conspirationnisme en France. Nous avons déjà souligné l’importance de cet événement aussi

par rapport au pics de publication d'articles sur la question. Les plongements nous permettent ici de mettre en lumière l'importance de thématiques complotistes corrélées à la religion musulmane pendant cette période. La méthode des plongements lexicaux permet à la fois de confirmer des hypothèses soulevées par une thématization manuelle des articles. Également, les motifs de l'antisémitisme et des discours d'extrême droite sont constitutifs des articles produits sur la question. Malgré tout, cette approche nécessite des apports plus qualitatifs, directement dans le texte pour comprendre comment le discours se construit et pourrait aussi être complétée par une analyse d'autres supports discursifs, car les thématiques saillantes relèvent ici du sens commun, propre au discours de presse généraliste. Les théories du complot sont tantôt reliées à l'antisémitisme, l'islam, la finance etc. Avoir un regard plus circulaire sur d'autres registres textuels comme la presse magazine orientée politiquement (*Valeurs Actuelles*, *Le Point*, *Marianne*) ou la presse régionale, permettrait de préciser ou de découvrir de nouvelles tendances, ou encore, un travail sur les réseaux sociaux pourrait aussi mettre en exergue des dépassements de ces thématiques ou non.

Remerciements

Nous remercions le Programme des Investissements d'Avenir (PIA3) qui finance le CERES et le Partenariat Hubert Curien (PHC) Proteus de collaboration scientifique Franco-Slovène.

Bibliographie

- Bonnet V. et Mercier A. et Siouffi G. (2022). Les circularités complotistes : lecture interdiscursive. *Mots. Les langages du politique*, 130, pp. 9-17.
- Bratich J. (2020). Civil Society Must Be Defended: Misinformation, Moral Panics, and Wars of Restoration. *Communication, Culture & Critique*.
- Bronner G. (2013). *La démocratie des crédules*. PUF. 360p.
- Calabrese L. (2018). Faut-il dire migrant ou réfugié ? Débat lexico-sémantique autour d'un problème public. *Langages*, Vol. 210, no. 2, pp. 105-124.
- Cueille J. (2020). *Le symptôme complotiste : aux marges de la culture hypermoderne*. Érès, Coll. Sociologie clinique, 280p.
- Dang Q. et GAO F. et ZHOU, Y. (2016) Early detection method for emerging topics based on dynamic bayesian networks in micro-blogging networks. *Expert Systems with Applications*, vol. 57, p. 285-295.
- Giry J. (2017). *Le complotisme 2.0, une étude de cas de vidéos recombinante : Alain Soral sauve Glenn et Tara dans The Walking Dead*. Quaderni, no 94, pp. 41-52.
- Hamilton, W. L. et Leskovec, J. et Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics p. 1489-1501.
- Harjuniemi T. (2022). Post-truth, fake news and the liberal 'regime of truth' – The double movement between Lippmann and Hayek. In *European journal of communication*, vol.37(3), pp. 270-283.
- Liu, B. et Han, F. X. et Niu, D. et Kong, L. et Lai, K. et Xu, Y. (2020). Story forest: Extracting events and telling stories from breaking news. *ACM Transactions on Knowledge Discovery from Data*

(TKDD), 14(3), 1-28.

Martin L. et Muller B. et Ortiz Suárez P. J. et Dupont Y. et Romary L. et Villemonte de La Clergerie E. et Seddah D. et Sagot B. (2020). CamemBERT: a tasty French language model. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, p. 7203-7219

Martinc M. et Laruncet A. et Bordier C. et Hernandez M. et Pollak S. et Lejeune G. (2024). The Evolution of Bias in French News Media: How Does Political Orientation Affect Semantic Change ?. Actes des 17es Journées internationales d'Analyse statistique des Données Textuelles (JADT), à paraître.

Moirand S. (2007). Les discours de la presse quotidienne. Observer, analyser, comprendre. [en ligne]. PUF. [Consulté le 10 février 2023]. <https://shs.hal.science/halshs-00551377>

Montariol S. et Martinc M. et Pivovarova L. (2021). Scalable and interpretable semantic change detection. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2021, pp. 4642-4652.

Nguyen, N. K. et Delahaut, T. et Boros, E. et Doucet, A., et Lejeune, G. (2023). Contextualizing Emerging Trends in Financial News Articles. *Proceedings of the Fourth Workshop on Financial Technology and Natural Language Processing (FinNLP)*, pages 1–9, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.