



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE  
LAUSANNE

## PROJET DE SEMESTRE - RAPPORT

PROFESSEUR SUPERVISEUR : JÉRÔME BAUDRY

PROFESSEUR Co-SUPERVISEUR : ANTOINE BOSELUT

SUPERVISEUR : SIMON DUMAS PRIMBAULT

---

# LA CONSTRUCTION DE LA GRANDEUR SAVANTE

## XVIII<sup>ÈME</sup>-XIX<sup>ÈME</sup> SIÈCLE

---

LABORATOIRE D'HISTOIRE DES SCIENCES ET DES TECHNIQUES -  
LHST

---

CLÉMENCE BARSI (SYSTÈMES DE COMMUNICATION)

---

2022 Semestre de Printemps - 10 Juin, 2022

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Constitution du corpus . . . . .	4
1.2	Plan . . . . .	4
<b>2</b>	<b>Méthode</b>	<b>6</b>
2.1	Distant Reading . . . . .	6
2.1.1	Le dataset . . . . .	6
2.1.2	Exploration . . . . .	7
2.1.3	Création de nos mesures . . . . .	7
2.1.4	Contenu des éloges . . . . .	10
2.2	Close Reading . . . . .	16
2.2.1	Littérature . . . . .	16
2.2.2	Lecture de deux éloges comparés . . . . .	16
<b>3</b>	<b>Résultats</b>	<b>18</b>
3.1	Distant Reading . . . . .	18
3.1.1	Style des éloges . . . . .	18
3.1.2	Évolution du contenu des éloges . . . . .	20
3.1.3	Évolution de la structure des éloges . . . . .	22
3.2	Close Reading . . . . .	27
3.2.1	Littérature . . . . .	27
3.2.2	Lecture des deux éloges . . . . .	29
<b>4</b>	<b>Discussion</b>	<b>31</b>
<b>5</b>	<b>Conclusion</b>	<b>35</b>
<b>A</b>	<b>Statistiques sur les éloges</b>	<b>37</b>
<b>B</b>	<b>Contenu des éloges</b>	<b>40</b>

# Chapitre 1

## Introduction

### De quelle façon la figure du savant construite à travers les éloges diffère avant et après la Révolution ?

L'Académie Royale des sciences est fondée en 1666 par Colbert sous le règne de Louis XIV pour réunir des savants qu'il a lui-même choisi afin de développer les sciences et de conseiller le pouvoir[19]. Cela est aussi un moyen pour la monarchie absolue de centraliser les savoirs, de les encadrer et de les mettre à son service. Petit à petit, l'Académie remplace les salons qui se tenaient chez des particuliers où avaient l'habitude de se retrouver les savants pour échanger des idées [6]. L'Académie offre ainsi un lieu officiel pour se réunir lors de séances régulières ainsi que de participer à la publication annuelle de l'Académie pour les mémoires les plus exceptionnels présentés lors des réunions. Cela permettait aux savants de faire connaître leur travail en province et à l'étranger.

En effet, parallèlement au développement de l'Académie de Paris, la Société Royale de Londres, l'Académie de Berlin, l'Institut des Sciences et des Arts de Bologne et l'Académie des Sciences de Pétersbourg se créent et se développent tout comme les académies de province comme celle de Montpellier. La communication entre savants fut ainsi renforcée, mais également par la création des journaux scientifiques comme *Journal des Savants* et par les voyages faits par les savants contribuant au rayonnement de l'Académie [6].

L'Académie est le haut lieu de l'excellence des sciences françaises au dix-huitième siècle. En effet selon Lagrange, l'Académie des sciences est "Le premier tribunal de l'Europe pour les sciences". De plus, elle regroupe des savants de disciplines variées. En 1785, l'Académie comprend 8 classes différentes : les classes de physique générale, histoire naturelle et minéralogie, géométrie, astronomie, mécanique, anatomie, chimie et botanique.

La Convention ferme l'Académie le 8 août 1793 pendant deux ans mais la rouvre en 1795. Elle prend alors le nom d'Institut National des sciences et des arts regroupant les anciennes académies scientifiques, littéraires et artistiques qui étaient indépendantes sous l'Ancien Régime. Les premiers à rejoindre l'Institut nouvellement créé sont les académiciens subsistants de l'Académie royale. Ceux-ci étant âgés, une nouvelle classe les remplace rapidement qu'on appelle la première classe professionnelle de scientifiques sortis des écoles. Les nouveaux membres sont choisis par l'Académie elle-même qui ne considère que ceux qui ont publié la meilleure recherche.

Les Académiciens avant la Révolution sont des personnes choisies par le Roi, grâce à la Révolution l'Académie peut choisir elle-même ses membres. L'accès à l'éducation facilité par la création d'écoles comme l'École polytechnique, permet à des membres de la classe moyenne d'entrer à l'Académie. Il subsiste cependant des traces de l'Ancien Régime avec la création de la classe d'*Académiciens libres* où se trouvent des ducs, classe créée en 1816 sous la Restauration.

De plus, venir d'une famille aisée ou éduquée facilite l'accès aux sciences et par extension à l'Académie.

Tous les savants aspirent à la rejoindre un jour. L'Académie façonne donc les carrières des scientifiques au dix-neuvième siècle, aussi bien les membres que les aspirants à devenir membre[4].

La figure du savant que cherche à renvoyer l'Académie est particulièrement mesurable dans les éloges que celle-ci publie chaque année. Les secrétaires perpétuels de l'Académie lisent les éloges des académiciens morts lors des séances publiques ayant lieu deux fois par an avant la Révolution puis une fois par an après. Ils choisissent librement parmi les membres décédés ceux dont ils font l'éloge. Parfois ces derniers sont décédés depuis plusieurs années. C'est l'une des rares fois où le public est autorisé à l'Académie. Ces éloges lus aux séances publiques ont pour audience le peuple français mais aussi les cercles du pouvoir et la communauté scientifique étrangère. Après la Révolution, ces réunions attirent l'élite politique et intellectuelle de la France. Ces éloges sont également imprimés dans la publication annuelle de l'Académie : *Histoire et Mémoires de l'Académie Royale des Sciences à Paris*.

Avant la Révolution il n'y a qu'un secrétaire perpétuel, puis deux après, un pour les mathématiques et un pour la physique. Ils sont en charge d'aller aux réunions de l'Académie et de prendre en note tout ce qui y a été discuté. Ils doivent également publier deux ouvrages par an : un qui fait le compte rendu de ce qu'il s'est passé à l'Académie cette année-là puis le deuxième, cité précédemment qui contient donc les éloges. Les éloges sont un exercice rhétorique car étant lus aux séances publiques de l'Académie, ils nécessitaient tout l'art de bien parler de leurs orateurs. C'est pourquoi les secrétaires perpétuels étaient également choisis en fonction de leur habileté littéraire. Les éloges sont retranscrits après la séance publique afin d'être publiés. Il y a donc un changement de forme qui est cette mise à l'écrit. Ce projet se sert de ces retranscriptions.

Fontenelle, secrétaire perpétuel de 1697 à 1740, est le premier secrétaire à rompre avec la tradition de l'éloge qui maximise la louange et minimise les éléments biographiques. Il suit la tradition d'Aristote exprimant que l'éloge doit insister sur les actes nobles du sujet, les exacerber mais doit surtout être un résumé de la vie de la personne dont on fait l'éloge. Fontenelle donne l'exemple pour tous les secrétaires à venir en transformant l'éloge d'un panégyrique à une biographie semi-critique[15]. C'est ce que d'Alembert qualifie d'éloge historique, sa première caractéristique étant qu'il doit respecter la vérité. Cependant cette vérité peut être nuancée et parfois même tue, car le but premier d'un éloge reste de faire la louange d'une certaine personne pour qu'elle puisse servir de modèle à ses successeurs. Ce sont des sources biographiques très fiables car elles sont recherchées méticuleusement par ses auteurs mais ce sont aussi des présentations héroïques de la vie des sciences mettant à l'honneur les modèles de vertu et de sagesse que sont ces savants. On célèbre le parcours semé d'embûches que ces scientifiques ont dû parcourir afin de poursuivre une carrière scientifique. Les vertus des savants sont accrues afin de minimiser mais pas cacher les défauts de ces scientifiques vertueux dont sont faits les éloges.

Les secrétaires présentent non seulement l'académicien décédé sous un jour favorable mais aussi la science afin de générer de l'intérêt du public pour elle.

À travers ces vies contées dans les éloges, l'image de l'académicien des sciences se construit. Cette image va au-delà des individus et créer une identité collective [2]. Elle est à l'origine d'une persona savante que l'Académie contribue à propager à travers les éloges [18]. Un persona est à l'intersection entre l'individuel et le collectif. C'est une identité culturelle qui façonne l'individu dans son corps et son esprit mais créé aussi un collectif doté d'une physionomie partagée et reconnaissable [20]. Et l'Académie, à travers les éloges, attribue à un homme de sciences des

valeurs épistémiques et morales après sa mort contribuant à la fabrique de la grandeur savante [18].

En lisant de la littérature, l'idée est venue d'étudier l'impact de la Révolution française sur la figure des savants transmise à travers ces éloges. Cette Révolution a chamboulé le paysage politique de l'époque. De plus la Convention ferme l'Académie car elle la considère comme anti-démocratique et élitiste mais la rouvre 2 ans plus tard sous une autre forme [13]. Cela implique qu'il y a nécessairement eu un impact de la Révolution sur l'Académie et donc sur la figure du savant transmise dans ces éloges.

## 1.1 Constitution du corpus

Les éloges ont été recensés et récupérés depuis diverses sources : les Histoires de l'Académie mais aussi les œuvres complètes de certains secrétaires ou bien d'autres éditions d'éloges par Jérôme Baudry. Ensuite les éloges ont été passés à travers de l'Optical Character Recognition pour avoir un format lisible par une machine et les erreurs ont été manuellement corrigées par Simon Dumas Primbault.

N'existant pas de liste exhaustive des éloges prononcés, il a fallu consolider le corpus progressivement grâce à des listes partielles. 270 éloges ont été réunis allant de 1741 à 1943 de 33 auteurs différents.

Il y a eu également des problèmes d'encodage dus aux caractères spéciaux qu'il a fallu identifier tel que les accents, le 'É' étant encodé par  $\Delta$ . Les marques de nouvelles lignes telles que '\n' ou encore d'alinéa '\t' ont été enlevées. L'orthographe ayant également évolué depuis le dix-huitième siècle, le choix a été fait d'adapter tous les mots identifiés comme ayant changé à l'orthographe pratiquée de nos jours. Cela concerne principalement des mots finissant en "nts" tel que "documents" qui pouvait être orthographié comme "documens". Certains éloges étaient également mal enregistrés comme des notices à la fin d'autres qu'il a fallu retrouver ou étaient mal datés. D'autres éloges contenaient aussi des informations supplémentaires pas prononcées lors de l'éloge. Celles-ci étaient contenues dans des crochets qui ont été séparées de l'éloge et mis dans une autre colonne.

Enfin il a fallu transformer les fichiers en txt contenant les éloges, en un fichier csv contenant l'auteur de l'éloge, le savant qu'il concerne, sa date d'oraison et enfin l'éloge lui-même.

Pour isoler l'effet de la Révolution sur la figure du savant, nous avons choisi de garder les éloges composés entre 1773 et 1832. La date de 1773 a été choisie pour avoir un corpus uniforme avant la fermeture de l'Académie en 1793, tous les éloges étant écrits par le marquis de Condorcet. 1832 a été décidée pour voir l'effet de la Révolution sur un plus long terme et sous plusieurs régimes (la Première République, le Premier Empire, la Restauration et la monarchie de Juillet). 11 auteurs différents ont écrit les éloges entre 1795 et 1832. Cela amène à un total de 144 éloges, 79 avant 1793 et 65 après.

L'éloge le plus court est celui de Le Cat prononcé par Condorcet en 1773 et fait 128 mots. Le plus long quant à lui compte 22'992 mots, et est aussi rédigé par Condorcet à propos de L'Hôpital en 1777.

## 1.2 Plan

Afin d'étudier cette question, trois axes ont été choisis. Le premier axe se concentre sur des statistiques générales sur l'éloge. Le deuxième sur l'évolution du contenu de l'éloge et enfin le dernier sur l'évolution de la structure de l'éloge.

Tout le code est disponible sur ce lien : [https://github.com/LHST-EPFL/eloge\\_academie](https://github.com/LHST-EPFL/eloge_academie)

# Chapitre 2

## Méthode

### 2.1 Distant Reading

#### 2.1.1 Le dataset

Avant de répondre à la problématique, nous avons effectué quelques mesures sur le corpus que nous avions.

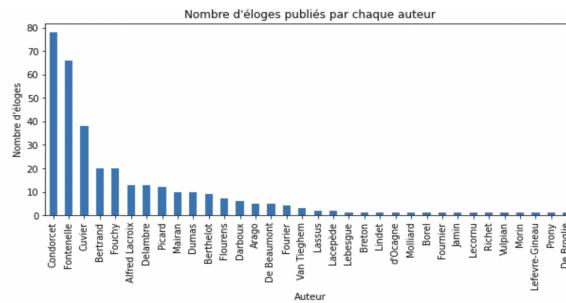


FIGURE 2.1 – Nombre d'éloges publiés par chaque secrétaire

Dans cette figure 3.1, nous pouvons observer qui sont les plus grandes sources d'éloges. Les deux plus gros fournisseurs sont Condorcet et Fontenelle, deux secrétaires de l'Académie du dix-huitième siècle.

Afin de comparer cette mesure avec la longévité des secrétaires, nous avons récupéré et créé un datafram avec les différentes dates de service des secrétaires [19]. Dans la figure suivante 2.2, nous pouvons voir le nombre d'années où les différents savants ont été secrétaire de l'Académie.

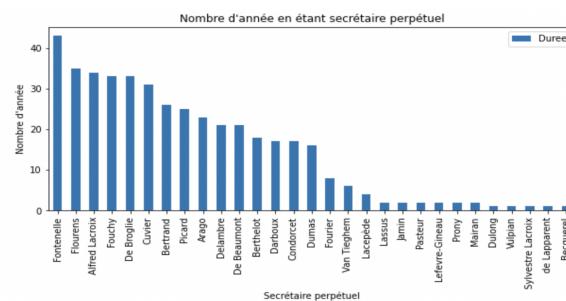


FIGURE 2.2 – Nombre d'années de chaque secrétaire

Finalement pour recouper les métriques précédentes et savoir qui a produit le plus d'éloges pour son nombre d'année au poste de secrétaire, nous avons divisé le nombre d'éloges publiés par le nombre d'années en tant que secrétaire. Cela nous donne la moyenne d'éloges publiés par an par secrétaire. Les résultats sont visibles dans la figure 2.3 ci-dessous :

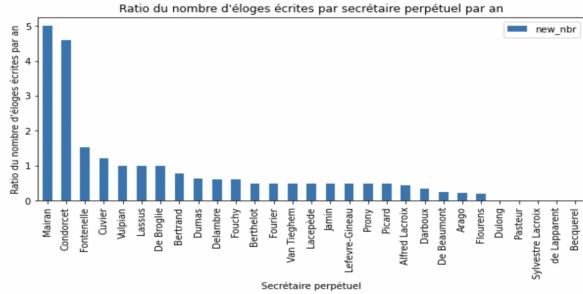


FIGURE 2.3 – Moyenne d'éloges publiés chaque année par secrétaire

Nous pouvons voir que les plus gros producteurs d'éloges par année sont Mairan et Condorcet, loin devant les autres. Ils sont tous les deux secrétaires au dix-huitième siècle.

### 2.1.2 Exploration

La première phase du projet a été une phase d'exploration, afin de tester le corpus et d'établir une problématique. Le seul corpus complet au début du projet étant celui de Condorcet, nous avons commencé par ce corpus. Nous avons essayé différents outils déjà entraînés afin de voir s'ils fonctionnaient sur le corpus. Cependant nous nous sommes vite rendu compte que la plupart de ces outils ne sont disponibles que pour l'anglais, comme le Part-Of-Speech tagger de NLTK. Nous avons aussi testé la librairie Textblob [11] pour étudier le sentiment du discours. Mais cette librairie est entraînée sur des revues de films anglaises et les résultats ne sont donc pas utilisables. Nous avons également essayé de trouver les adverbes modaux dans les textes pour tenter de voir l'assertion du discours, mais ces outils n'existant qu'en anglais, nous avons décidé de créer nos propres outils de mesure pour faire ressortir certains indices et orienter la lecture historique.

### 2.1.3 Crédation de nos mesures

Peu d'outils étant disponibles pour le français et l'éloge étant un type de rhétorique très particulier, nous avons choisi de définir nous-mêmes nos propres outils pour orienter la lecture historique et comparer les éloges pré et post-Révolution.

De plus nous ne souhaitions pas créer de gros outils qui produisent des résultats directement mais plutôt une collection de petits outils qui produisent un faisceau d'indices permettant de guider l'analyse textuelle et historique, ainsi que le close reading.

### Style des éloges

La première partie des mesures créées concernent directement le style des éloges.

Nous avons commencé par mesurer le nombre d'éloges écrits par an, pour voir s'il y avait une évolution. Si avant ou après la Révolution plus de savants avaient un éloge écrit à leur propos et donc que plus de savants différents étaient érigés en exemple.

Une deuxième mesure que nous avons utilisée est la longueur des éloges, voir s'ils se sont allongés ou raccourcis suite à la Révolution.

Une troisième mesure concerne la stylométrie des éloges. La longueur des mots employés par un auteur est une indication du style de l'auteur, nous avons donc voulu voir si la longueur moyenne évoluait [9].

Enfin la dernière mesure choisie sur le style des éloges est la longueur des phrases si elles sont plus longues ou au contraire plus courtes.

Nous voulions initialement nous pencher sur la stylistique des éloges à travers les figures de style. Nous avons dû nous raviser par manque d'outils.

Pour les prochaines mesures nous nous sommes servis de Part-Of-Speech. Le Part-Of-Speech consiste à identifier la nature des mots présents dans un texte. En NLTK, le POS tagger n'existe pas pour le français. Nous nous sommes donc servis de TreeTagger [16]. C'est un module python développé par l'université de Stuttgart, il offre également un outil de lemmatisation en français.

## Première personne

Pour commencer, nous avons étudié la proportion des pronoms à la première personne utilisée au cours des éloges sur tous les pronoms utilisés pour voir si l'importance du secrétaire a évolué au cours du temps. Cela donnerait une indication sur la subjectivité du discours. Cela marque des observations faites par l'orateur, cela reconnaît au narrateur des pouvoirs d'observation et ce n'est pas un relèvement passif des événements [23]. Nous pouvons supposer qu'avec la Révolution, il y a un changement dans l'importance de l'opinion du secrétaire dans l'éloge.

## Conjugaison

**Proportion des verbes au passé :** Nous avons également étudié la proportion des verbes au passé au cours du texte. Cela caractérise l'effet appelé *Project-Back* [23]. Cet effet génère des moments passés dans la tête du lecteur. Ce sont des éloges donc cela paraît évident que la proportion de verbe au passé soit haute, mais avec la Révolution nous pouvons faire la supposition qu'il y ait une évolution. Nous pouvons penser que le passé serait moins évoqué car synonyme des valeurs de l'Ancien Régime.

**Proportion des verbes au conditionnel et à l'impératif :** Pour voir l'assertion du discours, nous avons aussi relevé la proportion de verbe au conditionnel et à l'impératif par rapport à tous les verbes. Le conditionnel marque l'éventualité alors que l'impératif l'ordre. Nous avons pensé que la Révolution aurait pu faire évoluer ses proportions car elle est à l'origine d'un système plus démocratique en France.

## Adverbe des éloges

Et finalement, pour obtenir d'autres indices sur l'assertion, la temporalité et le sentiment transmis à travers ces éloges nous nous sommes servis des adverbes. Pour avoir une liste gérable, nous avons seulement gardé les adverbes ou bigrams d'adverbes (deux adverbes à la suite l'un de l'autre) cités plus de 5 fois à travers tous les éloges. Cela a créé une liste de 416 adverbes. Ces adverbes ont ensuite été manuellement placés dans l'une des catégories s'ils rentraient dedans : assertion, temporalité ou expérience. La classification des adverbes est montrée dans l'image ci-dessous 2.4.

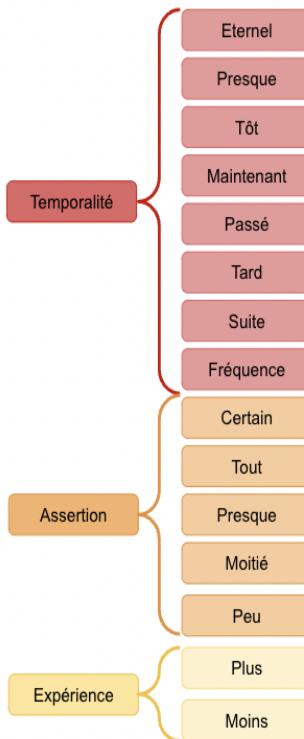


FIGURE 2.4 – Classification des adverbes

Certains adverbes rentraient dans plusieurs catégories tel que "souvent" qui peut être une qualification d'assertion ou de temporalité. Nous avons décidé de laisser ce type d'adverbes dans les deux catégories concernées.

**Temporalité :** Les adverbes classifiés dans la *Temporalité* représente 84 adverbes. Cette catégorie englobe tous les adverbes qui ont un lien avec le temps. Dans cette catégorie on retrouve 8 sous-types :

- *éternel* : les adverbes caractérisant une éternité ou une infaillibilité tel que "définitivement"
- *presque* : ce degré regroupe des adverbes qui sont presque dans la catégorie *éternel* mais pas complètement tel que "presque toujours" ou "presque jamais"
- *tôt* : cette catégorie contient des adverbes qui ont une sensation de rapidité, de précipitation, d'un événement qui va se produire très bientôt. S'y trouvent des adverbes tel que "subitement", "si vite" ou encore "presque aussitôt"
- *maintenant* : cette catégorie représente une simultanéité, un événement qui se produit actuellement. On y retrouve "aujourd'hui", "désormais"
- *passé* : ce regroupement contient les adverbes reliés à des événements antérieurs, y sont présents "auparavant", "anciennement"
- *tard* : cette catégorie ne contient que deux adverbes "tard" et "lentement", elle représente tout ce qui tarde.
- *suite* : les adverbes de ce type sont des adverbes où une suite logique à ce qui a été énoncé précédemment est impliquée comme "puis" ou "ensuite"
- *fréquence* : adverbes qui montrent une régularité tels que "journellement" ou "annuellement"

**Assertion :** La catégorie *Assertion* regroupe tous les adverbes relatifs au degré d'affirmation du discours. C'est-à-dire si cet adverbe rend la phrase très assurée ou au contraire peu. Dans cette même catégorie, les adverbes ont été réséparés en cinq catégories : *certain*, *presque*, *moitié*, *peu*, *tout*. Chacune de ces catégories reflète un différent degré d'assertion du discours. 64 adverbes sont rentrés dans cette catégorie.

Une catégorie comme *certain* contient des adverbes tel qu'"absolument" ou "certainement" dénotant de l'assurance. La catégorie *presque*, quant à elle regroupe des adverbes du type "probablement", "généralement" marquant une certitude quasi complète. La catégorie *tout* caractérise un englobement total comme "complètement", "totalement". Enfin les dernières catégories *moitié* et *peu* dénotent une assertion amoindrie. *Moitié* comprend des adverbes tel que "peut-être" où il y a plus de certitude que dans des adverbes comme "quelquefois" qui est quant à lui dans la catégorie *peu*.

**Expérience :** Finalement, la catégorie *Expérience* regroupe les adverbes relatifs à la positivité ou négativité du discours. Ce sont les adverbes qui communiquent un sentiment positif ou négatif. Ils ont donc été classifiés entre ces deux catégories. 70 adverbes se trouvent dans cette catégorie. Par exemple dans la catégorie *plus*, on retrouve "beaucoup mieux" et dans *moins*, il y a plutôt des adverbes du type de "cruellement".

## Utilisation des noms propres

Comme vu dans l'introduction, la coopération entre savants et les voyages augmentent au dix-huitième siècle. Pour percevoir cela, les noms propres regroupant les noms de lieux et de personnes, nous avons tenté de voir s'il y avait eu une évolution au cours du temps. Pour cela nous avons taggé les noms propres avec TreeTagger et examiner l'évolution de leur emploi.

### 2.1.4 Contenu des éloges

Nous terminons ici l'analyse statistique de l'éloge et passons à une analyse thématique des éloges. Pour faire cette analyse thématique nous nous servons de topic modeling. C'est une méthode d'apprentissage non supervisée qui permet d'analyser un ensemble de documents. Elle regroupe les documents qui se ressemblent entre eux et en ressort les mots les plus représentatifs.

Le topic modeling a d'abord été testé sur le corpus de Condorcet. Dans un premier temps, nous avons essayé de faire du topic modeling avec comme document l'entièreté de l'éloge. Donc essayer de regrouper les éloges en cluster. Nous avons commencé par essayer avec le Latent Dirichlet Allocation de la librairie Gensim. Le LDA est une des méthodes les plus populaires actuellement pour faire du topic modeling. C'est une démarche probabiliste. Le nombre de topic à générer est fixé à l'avance, la phase d'initialisation consiste à attribuer un thème à chaque mot de chaque document. Ensuite pour apprendre la distribution, chaque mot de chaque document est étudié, on fait l'assumption que tous les topics attribués sont corrects sauf pour le mot étudié. Puis on calcule deux quantités pour chaque topic. La première est la probabilité que le document  $d$  soit assigné au topic  $t$ ,  $P(t|d)$ . Et la deuxième que le topic  $t$  soit assigné au mot  $w$ ,  $P(w|t)$ . Le nouveau topic  $t$  est choisi avec la probabilité  $P(t|d) \cdot P(w|t)$ . Cette étape est répétée de nombreuses fois et finit par se stabiliser. Plusieurs topics sont attribués à un même document. Les mots associés à chaque topic sont obtenus en comptant les mots qui y sont associés dans le corpus.

En essayant d'optimiser la cohérence en itérant sur le nombre de topics, le topic modeling n'était toujours pas concluant. La cohérence est une mesure qui permet d'évaluer un modèle en calculant le degré de similarité sémantique entre les mots à hauts scores d'un topic. Nous sommes donc ensuite passés à une autre technique.

La Non-Negative Matrix Factorization en utilisant le module de sklearn a ensuite été essayée mais ne créait pas des topics très interprétables non plus [3]. Cette méthode est populaire car elle arrive à extraire des facteurs facilement interprétables. Comme nous pouvons le voir dans l'image 2.5 ci-dessous, l'input de cette méthode est  $V$ , la document-term matrice. Cette matrice est ensuite factorisée en deux matrices  $W$  et  $H$  non-négatives qui contiennent les représentations des topics. Cette méthode vise à minimiser l'erreur  $\|V - W \cdot H\|_F$ ,  $W \geq 0$  et  $H \geq 0$  en utilisant la norme Frobenius. La norme de Frobenius correspond à  $\|A\|_F = \sqrt{\sum_{1 \leq i \leq m} \sum_{1 \leq j \leq n} |A_{ij}|^2}$ .

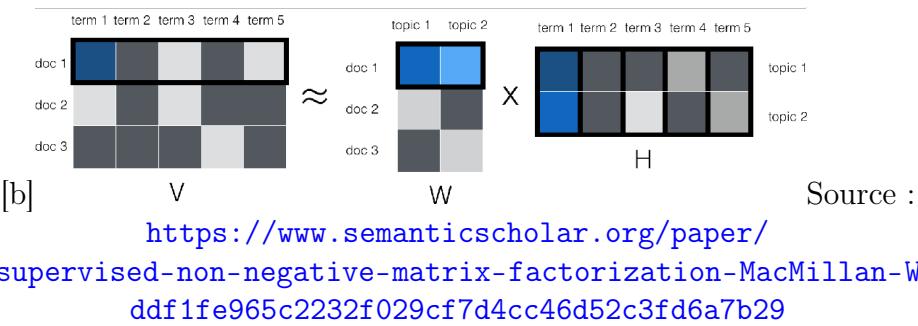


FIGURE 2.5 – Fonctionnement de la Non-Negative Matrix Factorization

Nous sommes donc passés à BERTopic qui a immédiatement eu des résultats beaucoup plus intéressants et sommes donc restés dessus quand nous sommes passés au grand corpus. Pour créer des topics plus précis, nous avons divisé les éloges en phrases et fait du topic modeling directement sur les phrases. Les éloges étant des textes souvent longs, nous n'obtenions que des topics très généraux.

BERTopic est un modèle développé en 2020 pour faire du topic modeling[7]. La première étape de BERTopic consiste à générer des embeddings pour les documents. Nous avons testé différentes sortes d'embeddings, celui de BERTopic par défaut en français, celui de CamemBERT, de FlauBERT et de SentenceTransformer. Nous avons choisi de tester CamemBERT car il est entraîné sur un très gros corpus français [12]. Nous avons également testé FlauBERT, il est quant à lui entraîné sur un corpus hétérogène français [10]. Et finalement SentenceTransformer [17] car c'est un des modèles qui fonctionnent le mieux pour créer des embeddings. Celui qui fonctionnait le mieux pour représenter les données était ce dernier. Ce dernier permet d'encoder une phrase directement dans un espace à 512 dimensions. Une fois que l'embedding est fait, BERTopic utilise l'algorithme UMAP pour réduire la dimensionnalité. Celui-ci permet de conserver la structure globale et locale des embeddings. Et finalement pour créer les clusters, il utilise l'algorithme de HDBSCAN sur les embeddings réduits. La dernière étape consiste à extraire les mots les plus représentatifs d'un cluster en utilisant TF-IDF. TF-IDF permet d'extraire les mots fréquents dans un cluster mais rare dans les autres. L'image ci-dessous représente la pipeline de BERTopic décrite 2.6 :

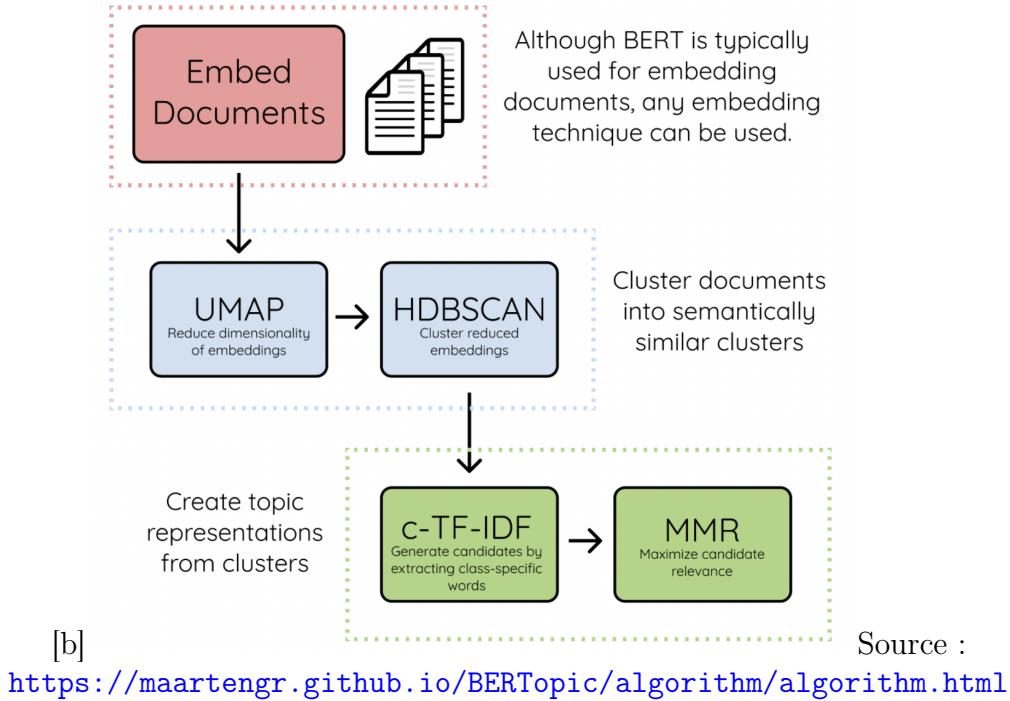


FIGURE 2.6 – Pipeline de BERTopic

BERTopic permet d'atteindre des résultats bien supérieurs aux techniques précédentes car il combine l'utilisation de Transformers, UMAP et HDBSCAN qui sont tous trois des méthodes créant d'excellents résultats.

La taille minimum des topics a été recherchée manuellement afin d'obtenir des topics représentatifs (les différentes valeurs testées étaient 10, 15, 18, 20, 25, 30). La valeur qui a créé la meilleure séparation des topics a été 20. Nous avons mis le paramètre `nr_topics` à 'auto'. Cela permet de fusionner automatiquement les topics les plus similaires. Nous avons mis `calculate_probabilities` à 'True' pour récupérer les probabilités de tous les topics par document. Et enfin `n_gram_range` à (1,2) pour générer des bigrams. Cela a permis par exemple de contre-carrer la nuance créée dans un topic sur l'histoire qui regroupait des phrases sur les historiens et sur la biologie avec le terme "histoire naturelle".

## Pre-Processing

Le pre-processing a eu lieu en deux temps.

Dans un premier temps, il a fallu découper le corpus en phrases. Pour cela les "M." qualifiant les "Monsieur" ont été remplacé par "M", pour faciliter la division en phrases. Les phrases ont été séparées par les délimiteurs : '.', '?', '!'. Nous avons considéré aussi de diviser les phrases avec les ';', mais cela créait une granularité trop importante au niveau des topics produits plus tard lors du topic modeling.

Pour utiliser BERTopic, il est recommandé de ne pas faire de pre-processing sur les données car le modèle se sert de tout le contexte d'un document. Cependant dans notre cas, sans pre-processing les topics obtenus n'étaient pas interprétables. Les éloges contiennent beaucoup de bruit. Les stop-words ont été enlevés en utilisant la librairie spacy et en enlevant les mots additionnels suivants : 'plus', 'ce', 'cette', 'encore', 'alors', 'qu', "qu'", car ils polluaient la création des topics. Spacy a été choisi plutôt que NLTK car son set de stop-words est plus grand que ce dernier. Les chiffres et le reste de la ponctuation ont également été enlevés. Les

mots ont été lemmatisés en utilisant le lemmatiseur TreeTagger pour français. Nous avons testé le lemmatiseur du model *fr\_core\_news\_sm* de spacy ainsi que le *SnowballStemmer* de NLTK en français. Mais ils produisaient des résultats très inférieurs à ceux du TreeTagger. Finalement pour enlever le bruit additionnel créé lors de la génération des topics, seuls les noms communs, les verbes et les adjectifs, tous de plus de deux lettres ont été gardés pour le topic modeling.

Nous nous sommes également rendu compte qu'après le pre-processing, il subsistait des phrases qui marquaient les bas de pages des éloges telles que "(Page 133) Condorcet". Ces phrases ont été enlevées.

## Evolution du contenu

L'évolution des topics au cours du temps a été observée en utilisant les visualisations de BERTopic. Deux différentes normalisations ont été utilisées pour la visualisation, la première a été la normalisation par le nombre d'éloges publiés chaque année afin de tester l'évolution de la proportion de ce topic par éloge. Enfin une deuxième normalisation a été utilisée : par le nombre de phrases à travers tous les éloges cette année-là, pour voir si le nombre précédent est biaisé par l'augmentation de la longueur des éloges. Nous avons choisi de faire ces deux normalisations car même si la proportion d'un topic a augmenté parce que l'éloge a été rallongé cela reste intéressant à observer.

## Evolution de la structure des éloges

**Hiérarchisation des topics :** Finalement pour étudier la structure des éloges et son évolution, les topics ont été classés manuellement dans une hiérarchie faite par nous-mêmes en suivant l'exemple de Jo Guldin dans l'article "Parliament's Debates about Infrastructure : An Exercise in Using Dynamic Topic Models to Synthesize Historical Change" [8]. Celle-ci est à trois étages. Le premier étage est divisé en *postérité, société, sciences, voyage, personne et théorie*.

La catégorie *postérité* contient des topics reliés à ce que le savant laisse derrière lui. On retrouve dans cette catégorie les topics suivants : "16\_imprimer\_écrire\_manuscrit édition", "22\_langue\_latin\_livre\_traduction" et "15\_mémoire\_souvenir\_oublier\_rappeler". Le topic 22 figure aussi dans cette catégorie car il est relié aux écrits des savants, comme le montre cette phrase classée dans ce topic : "Ces ouvrages sont principalement un traité et un dictionnaire de physique dont il a vu plusieurs éditions, et qui ont été traduits en plusieurs langues".

Les catégories *religion* et *politique* ont toutes deux été regroupées dans la super-catégorie *société* car toutes les deux relatives à la société actuelle.

Dans la catégorie *religion*, nous trouvons "32\_catholique\_clergé\_protester\_religion" et "21\_religion\_religieux\_chrétien\_dieu". Nous retrouvons dans ces topics des phrases telles que : "Le clergé romain aimait mieux être toléré et protégé par le gouvernement anglais, que de voir s'établir une liberté d'opinions toujours si effrayante pour des hommes accoutumés à dominer les esprits". Cette phrase est clairement liée à la société de l'époque justifiant que cette catégorie soit reliée à celle de la *politique*.

Dans la catégorie *politique*, nous retrouvons des topics comme "35\_im-pôt\_payer\_peuple\_indirect". Dans ce même topic, il y a des phrases comme "Un acte du parlement assujettit donc les colonies américaines à l'impôt du timbre et à quelques taxes sur les denrées". Cette catégorie regroupe les topics qui parlent donc du contexte politique

de l'époque et où nous pouvons donc y retrouver des phrases sur les mesures imposées aux colonies de l'époque.

Une autre super-catégorie est la catégorie des *sciences*. Cette super-catégorie est elle-même divisée en 8 catégories : *astronomie*, *biologie*, *chimie*, *médecine*, *physique*, *géologie*, *mathématiques* et *philosophie*. Cette super-catégorie contient donc tout ce qui est lié à une discipline en particulier. Chacune de ces sous-catégories contient les topics reliés à cette discipline. Par exemple dans la catégorie *mathématiques*, on retrouve deux topics : "19\_géomètre\_géométrie\_méthode\_problème" et "27\_mathématique\_mathématicien\_calcul\_étude\_mathématique".

Un autre regroupement est celui appelé *méthode*. Dans ce dernier, nous trouvons tout ce qui est lié à la pratique ou à la méthodologie de la science, des topics tels que "23\_expérience\_expérience faire\_expérience direct\_répéter\_expérience". Une phrase tirée de ce topic est "Il confirme, par des expériences, ce phénomène que Mariotte et Leibniz avaient déjà observé".

La catégorie *voyage* ne contient qu'un seul topic : "43\_voyage\_voyager\_voyageur\_tour" et des phrases telles que : "Cependant, si de Saussure n'eût porté à ses voyages que ces dispositions vagues et s'il n'en eût recueilli que ces impressions générales, nous n'aurions probablement point à faire ici son éloge". Dans ce topic, sont présents tout ce qui a trait avec le voyage donc, à l'extérieur ou non de la France.

Finalement, nous avons une dernière super-catégorie. Il s'agit de la catégorie *personne*. Dans cette même catégorie, nous retrouvons deux sous-catégories : *vie* et *figure-type*. Ces deux catégories ont été regroupées car l'une représente la vie du savant c'est-à-dire sa famille, sa mort, son éducation, etc... L'autre caractérise la représentation du savant, de cette figure de savant ou encore les titres qu'il a. Dans la catégorie *vie*, sont contenus des topics comme "12 Jeune\_jeunesse\_jeune homme\_homme". Dans *figure-type*, sont présents par exemple "2\_science\_progrès\_savant\_homme" ou encore "41\_cardinal\_cardinal\_luynes\_luynes\_vergne".

Cette hiérarchie est visible dans la figure 2.7 ci-dessous. Pour accéder à la version complète et interactive [cliquez ici](#).

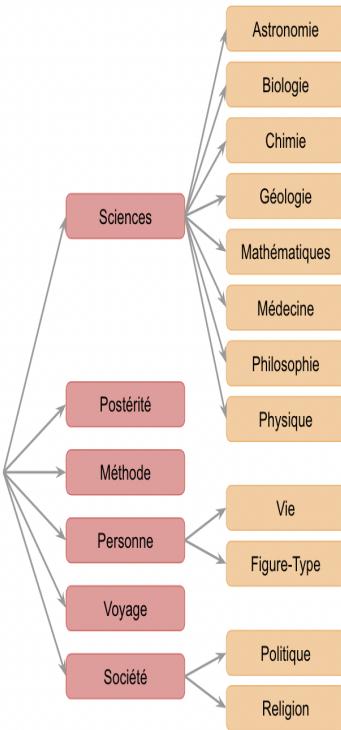


FIGURE 2.7 – Hiérarchie des topics

Pour justifier le choix de ces catégories, pour chaque topic, nous avons extrait les 10 phrases qui avaient le plus de chance d'être associées à ce topic. S'il y en avait plus de 10, 10 étaient choisies au hasard entre les plus probables.

Certains topics, tels que le topic 31 sur la force, étaient ambiguës. En effet certaines phrases rentrant dans ce topic étaient en rapport avec la physique d'autres avec les qualités du savant et pouvaient donc rentrer dans deux catégories : *vie* et *physique*. Nous avons ensuite tenté de voir si le deuxième topic ou troisième topic les plus probables des phrases permettraient de contrebalancer cette nuance. Cependant cela n'a pas été concluant. Nous avons donc choisi de placer ce topic dans la catégorie à laquelle le plus de phrases appartenaient. Il a donc été mis dans la sous-catégorie *vie* de la catégorie *personne*.

**Structure des éloges :** Les éloges ont ensuite été reconstruits avec les phrases, leur topic associé, leur sous-catégorie associée et catégorie associée.

Il a fallu trouver un moyen de pouvoir visualiser ces données. Au début, nous avons pensé colorer chaque phrase en fonction du topic associé ou de sa catégorie associée. Mais dans ce cas il aurait fallu regarder chaque éloge et essayer de voir s'il y avait des tendances ou des regroupements. Nous avons donc divisé les éloges en 10. Les 10 premiers pour-cent de l'éloge, les 10 suivants et ainsi de suite. Pour visualiser les résultats, nous avons ensuite fait des stacked bar pour les différents regroupements de topics avant et après la Révolution. Ces résultats ont été normés de deux manières différentes pour voir deux tendances différentes. La première est une normalisation verticale. C'est-à-dire que pour chaque topic, le nombre de fois où ce topic a été utilisé dans cette partie a été divisé par le nombre de topics de cette catégorie utilisés cette année. Par exemple si la catégorie astronomie est à 20% dans les 10-20% du texte par rapport aux catégories de sciences avant la Révolution, cela veut dire que de toutes les fois où

les sciences étaient mentionnées dans les 10-20% des éloges avant la Révolution, 20% des fois cela concernait l'astronomie.

La deuxième normalisation est horizontale. Nous divisons le nombre de fois où un topic spécifique a été mentionné dans cette partie du texte par le nombre total de fois où ce topic a été utilisé au cours des éloges. Si nous reprenons l'exemple précédent, dans le cas de cette normalisation, cela veut dire que 20% des fois où l'astronomie est mentionnée à travers les éloges avant la Révolution, elle l'est dans les 10-20% du texte.

## 2.2 Close Reading

Pour compléter le projet, nous avons terminé par du close reading pour recouper nos résultats et faire le lien avec le distant reading.

Dans un premier temps, afin d'orienter notre lecture grâce au faisceau d'indices constitués. Pour cela nous avons choisi deux éloges représentatifs des sous-corpus d'avant et après la Révolution. Cela permet de guider la lecture et de mettre à l'épreuve des résultats quantitatifs qui sont parfois peu signifiants ou instables.

Dans un deuxième temps, l'approche quantitative travaille sur des bribes de texte, lemma-tisées et sorties de leur contexte. Pour avoir une vue générale de ce qu'est un éloge et de sa fonction dans la construction de la grandeur savante il est nécessaire de se plonger dans les éloges.

Dans un troisième temps, cela permet de compléter l'approche quantitative par une interprétation qualitative car l'éloge est un genre si singulier que les algorithmes ne peuvent souvent en saisir la complexité.

### 2.2.1 Littérature

Afin de comparer nos indices avec ce qui avait déjà été dit sur les éloges nous sommes allés chercher dans la littérature. Cela a aussi permis de compléter ou d'expliquer nos résultats par le contexte historique.

Nous avons pu trouver des commentaires sur la structure rhétorique de l'éloge chez Paul[15] avant la Révolution. La structure que nous avons essayé de mettre à jour est thématique, elles sont donc complémentaires.

L'article d'Outram[13] a permis de comprendre l'image transmise à travers les éloges de Cuvier et d'avoir une idée du discours qu'il souhaitait transmettre après la Révolution.

Croslan[4] évoque le contexte historique de l'Académie après la Révolution, dans quel contexte les éloges sont lus ainsi que qui sont les académiciens après la Révolution, à quelle classe sociale ils appartiennent, quel est leur train de vie...

### 2.2.2 Lecture de deux éloges comparés

Le projet s'est conclu par une lecture comparée de deux éloges choisis l'un avant la Révolution, l'autre après. Les éloges choisis ont été ceux de Praslin écrit par Condorcet en 1785 et Volta écrit par Arago en 1831. Celui de Praslin a été choisi par rapport au topic généré "paix\_guerre\_vouloir paix\_protester" qui était très présent dans ce texte. Celui de Volta a été pris pour les topics de l'électricité très présents dans son éloge.

Le but de cette étude est de voir comment nos indices s'expriment ou de voir comment ce qu'ils essayent de mettre à jour peut se révéler différemment. Par exemple pour voir comment l'étude des adverbes qui avait pour but d'estimer l'assertion d'un texte est peut-être exprimée

d'une autre manière. Voir également ce que nos indices ont pu manquer et d'observer les limites du distant reading.

# Chapitre 3

## Résultats

### 3.1 Distant Reading

#### 3.1.1 Style des éloges

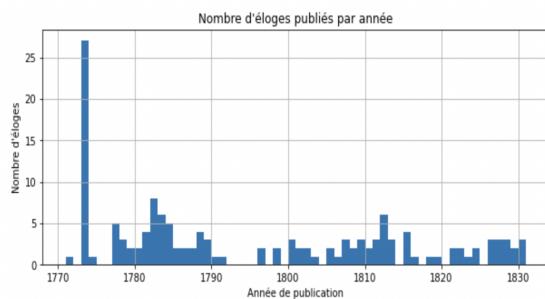


FIGURE 3.1 – Évolution du nombre d'éloges publiés par an

La figure 3.1 montre l'évolution du nombre d'éloges publiés par an. Le nombre d'éloges publiés par an n'a pas beaucoup évolué avant et après la Révolution. Nous pouvons voir une seule longue période sans éloge, il s'agit de 1792 à 1796. En dehors de cela, il n'y a jamais plus d'une année de suite sans éloge.

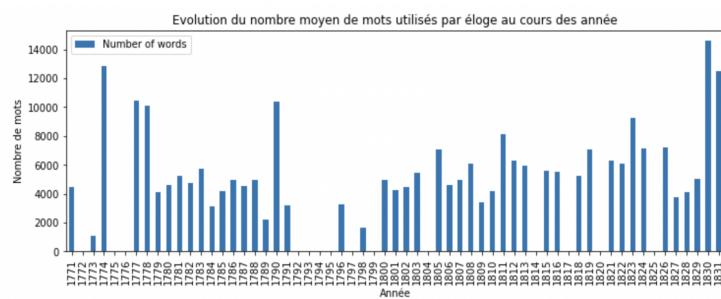


FIGURE 3.2 – Évolution de la longueur des éloges

L'évolution de la longueur moyenne des éloges par an est visible dans la figure 3.2. Les éloges semblent être légèrement plus longs après la Révolution à l'exception de certaines années comme 1798 ou 1809. Nous pouvons voir une décroissance lente jusqu'à la Révolution avec un étonnant sursaut en 1790.

Les mesures concernant l'évolution de l'utilisation de la première personne, des verbes au passé et de l'utilisation du conditionnel et de l'impératif n'ont pas montré de différences significatives entre le avant et le après la Révolution. Les graphes sont disponibles dans l'appendice.

## Adverbes

Degré	Moyenne pre-Révolution	Moyenne post-Révolution	T-Test p-value
Peu	0.09	0.05	0.02
Suite	0.14	0.12	0.5
Presque	0.33	0.26	0.02
Maintenant	0.11	0.17	0.01
Éternel	0.27	0.27	0.88
Tôt	0.03	0.09	0.00
Passé	0.02	0.04	0.13
Tard	0.00	0.00	0.75
Fréquence	0.00	0.01	0.08

TABLE 3.1 – Moyenne pour la catégorie Temporalité

**Temporalité :** Dans la tableau 3.1, nous pouvons voir que la différence dans la moyenne de la catégorie *suite* a 50% de chance d'être due au hasard. Les catégories *tard*, *éternel* ne manifestent pas de différences. Les adverbes qualifiant le passé, le présent ou un évènement dans le futur proche sont plus présents après la Révolution. En revanche les adverbes qui marquent des évènements qui ne se produisent pas de manière infaillible sont plus présents avant la Révolution. La catégorie *fréquence* est extrêmement peu présente avant ou après la Révolution.

Degré	Moyenne pre-Révolution	Moyenne post-Révolution	T-Test p-value
Plus	0.72	0.86	0.06
Moins	0.28	0.14	0.06

TABLE 3.2 – Moyenne pour la catégorie Expérience

**Expérience :** Concernant l'expérience, en proportion il y a plus d'adverbes négatifs avant la Révolution avec un niveau de confiance de 94% comme nous pouvons le voir dans le tableau 3.2. Il y a une inversion complète en 1790 juste après l'éclatement de la Révolution où il n'y a que des adverbes négatifs, cela est visible dans la figure 3.3.

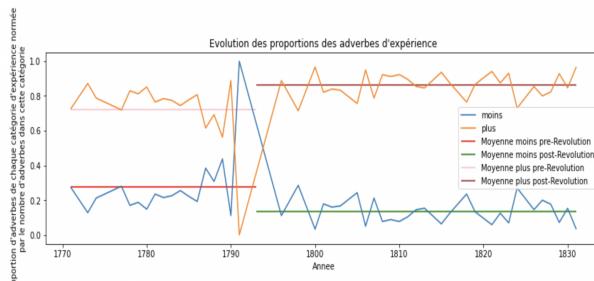


FIGURE 3.3 – Évolution des adverbes d'expérience

Degré	Moyenne pre-Révolution	Moyenne post-Révolution	T-Test p-value
Tout	0.11	0.19	0.00
Certain	0.43	0.49	0.11
Presque	0.24	0.19	0.04
Moitié	0.14	0.09	0.03
Peu	0.07	0.04	0.07

TABLE 3.3 – Moyenne pour la catégorie Assertion

**Assertion :** Il y a plus d'adverbes où l'assertion est forte après la Révolution, ce sont les adverbes des catégories *tout* et *certain*. La proportion d'adverbes des autres catégories où il y a un doute plus ou moins grand est plus grande à chaque fois avant la Révolution.

### 3.1.2 Évolution du contenu des éloges Politique

**Evolution des topics de politique**

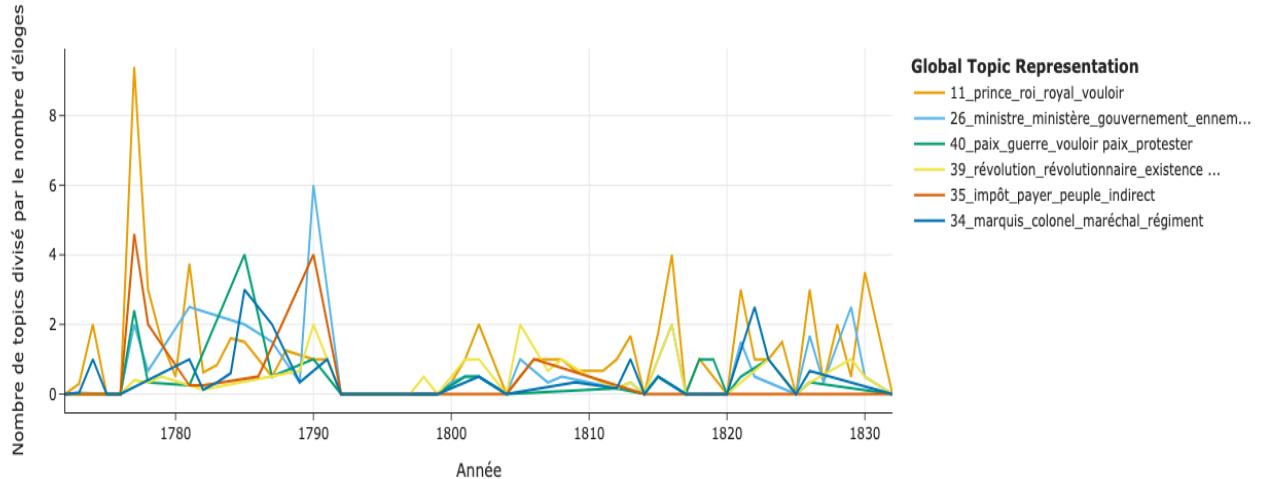


FIGURE 3.4 – Évolution des topics de politique au cours du temps normé par le nombre d'éloges cette année

Dans la figure précédente 3.4, l'évolution des topics de politique est observable. L'utilisation des topics concernant la politique a diminué au cours du temps. On voit deux pics en 1776 et en 1790. A partir de 1815, les mentions de politique redeviennent un peu plus fréquentes mais par pics.

# Religion

## Evolution des topics de religion

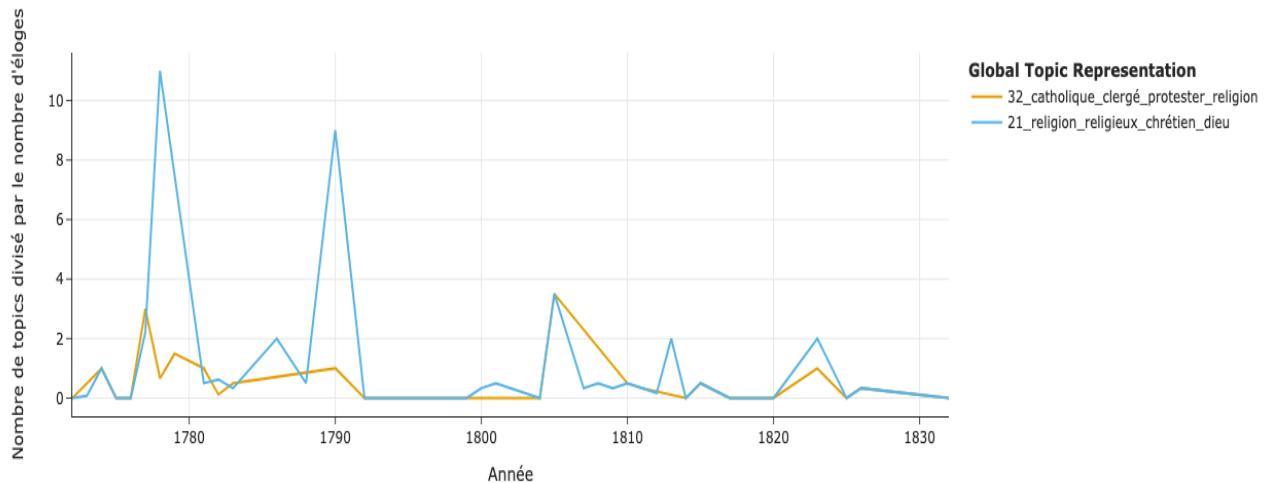


FIGURE 3.5 – Évolution des topics de religion au cours du temps normé par le nombre d'éloges cette année

Nous pouvons observer l'évolution de l'évocation de la religion dans le graphe 3.5. La religion connaît deux pics comme la politique en 1776 et 1790, en dehors de ceux-là elle est peu mentionnée et quasi plus après la Révolution.

## Méthode

Les topics relatifs à la méthode restent relativement constants au cours du temps. Le topic sur l'expérience reste constant jusqu'à 1830 où il connaît un pic important. Le graphe est disponible dans l'appendice.

## Postérité

Les topics sur la postérité ne connaissent pas beaucoup d'évolution au cours du temps.

## Sciences

- *Physique* : La physique est très peu mentionnée jusque 1825. Elle connaît un pic en 1830.
- *Philosophie* : La philosophie est beaucoup plus mentionnée avant la Révolution. Elle disparaît presque entièrement après celle-ci.
- *Mathématiques* : Ils ont un pic en 1773 et après leur évocation est constante au cours du temps.
- *Astronomie* : L'astronomie est beaucoup plus mentionnée après la Révolution. Elle connaît 6 pics qui sont probablement reliés à la mort d'académiciens spécialisés en astronomie.
- *Médecine* : La médecine est plus mentionnée avant la Révolution, elle connaît cependant un pic en 1828.

- *Géologie* : L'évocation de la géologie reste relativement constante au cours du temps et assez haute. Elle connaît deux gros pics en 1818 et 1823.
- *Biologie* : La biologie semble plus évoquée après la Révolution. En particulier la botanique qui connaît 4 pics en 1801, 1806, en 1815 et 1828. L'anatomie connaît un pic en 1803.
- *Chimie* : La chimie connaît quelques pics au cours du temps mais les plus gros sont en 1825 et 1830.

Les différents graphes de cette catégorie sont présents dans l'appendice.

## Vie

Le topic 12 sur la jeunesse du savant est beaucoup plus mentionné après la Révolution.

## Voyage

Sur le topic du voyage, la Révolution n'a pas particulièrement eu d'impact.

Les visualisations pour ces deux dernières catégories sont disponibles dans l'appendice.

### 3.1.3 Évolution de la structure des éloges

#### Structure globale

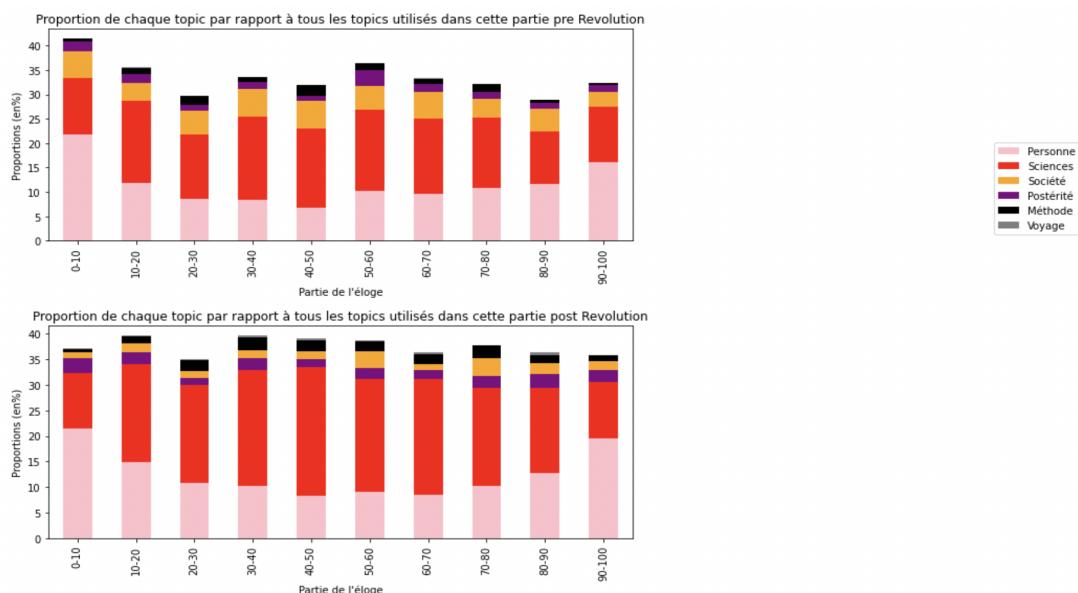


FIGURE 3.6 – Répartition de la structure global de l'éloge

Les 60% restants correspondent aux outliers, les phrases qui n'ont pas pu être placées dans des topics. Nous voyons dans la figure 3.6 que la science a une partie plus importante tout le long de l'éloge après la Révolution, en particulier vers le centre de l'éloge. Dans les deux cas les 10 premiers pour-cent accordent à peu près 20% à l'évocation de la personne. On voit également que ce qui se rapporte à la société est plus présent tout le long de l'éloge avant la Révolution et en proportion à peu près constante. La structure globale de l'éloge reste la même avant et après la Révolution. L'évocation de la personne est en U, étant plus mentionnée aux

extrémités de l'éloge et peu vers le centre. Pour la science, c'est le contraire où la science est de plus en plus mentionnée vers le centre de l'éloge.

La normalisation horizontale a confirmé ces résultats. Dans le graphe 3.7 ci-dessous nous voyons qu'effectivement le topic personne est plus utilisé dans les premiers et derniers pourcents du texte. Nous pouvons aussi voir que la catégorie science est plus utilisée vers le centre de l'éloge après la Révolution.

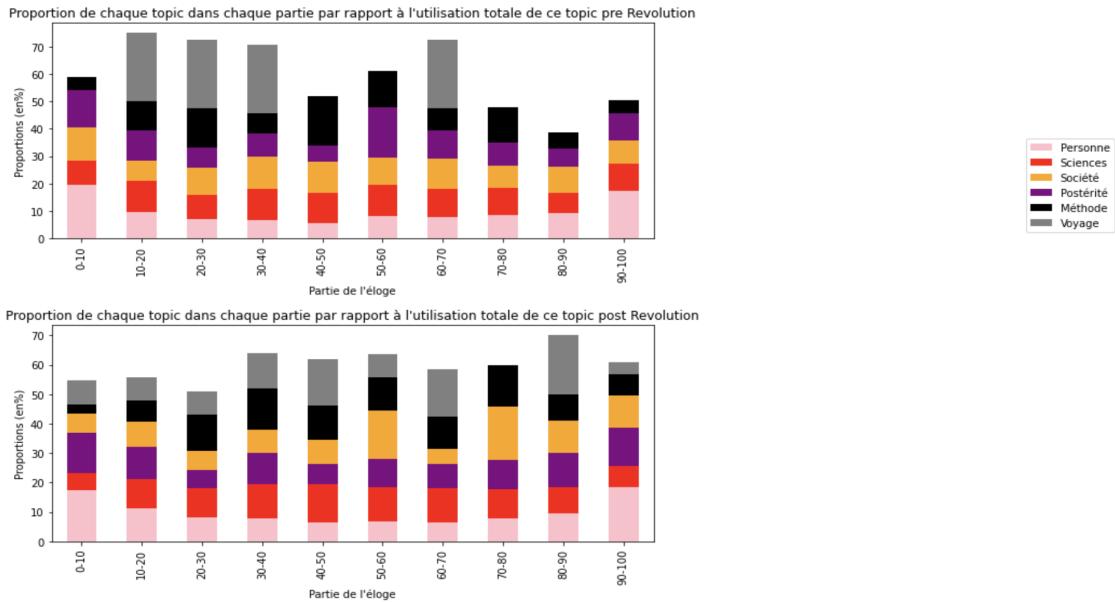


FIGURE 3.7 – Répartition de la structure global de l'éloge horizontalement

## Personne

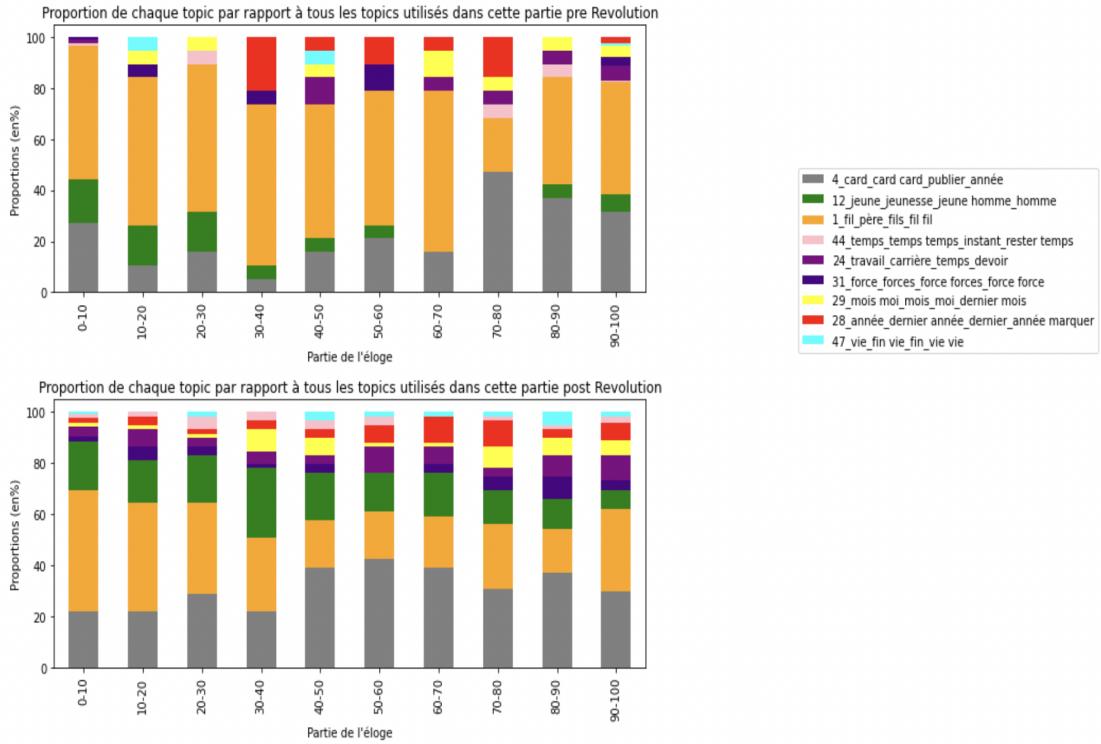


FIGURE 3.8 – Répartition des topics de vie

**Vie :** Dans la structure de l'éloge concernant la vie du savant visible dans la figure 3.8, nous pouvons voir une présence beaucoup plus importante tout le long de l'éloge du topic 1, il est présent après la Révolution mais en proportion plus petite. Après la Révolution, il consiste en maximum 50% mais tourne plutôt autour des 30%, alors qu'avant la Révolution il consiste en quasiment dans chaque partie du texte 50%. Après la Révolution, nous pouvons également voir que le topic 12 sur la jeunesse du scientifique est beaucoup plus présent tout le long du texte.

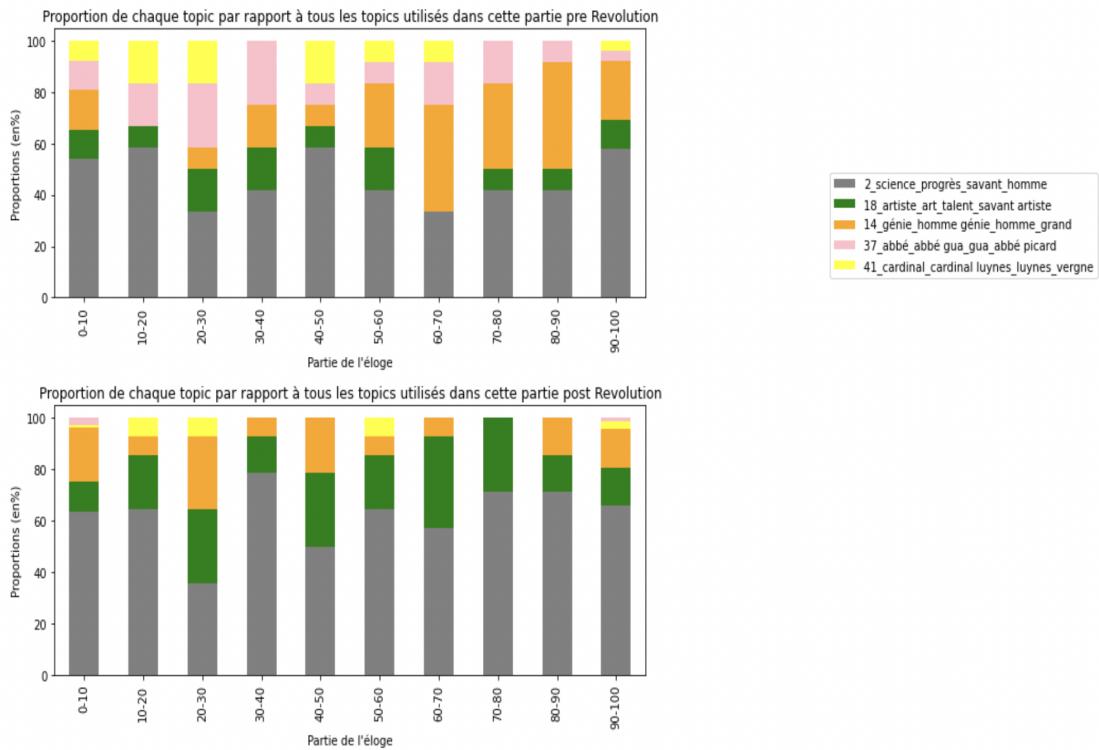


FIGURE 3.9 – Répartition des topics de Figure-Type

**Figure-Type :** Dans la figure 3.9, est observable la structure de la catégorie figure-type avant et après la Révolution. Nous pouvons voir que le topic 14 sur le génie est plus présent dans la dernière moitié du texte par rapport à après. Le topic 18 sur l'artiste, le savant et le talent est un peu plus présent en proportion après la Révolution et le topic 2 sur le progrès de la science est plus présent à partir des 30% de l'éloge après la Révolution. Finalement les topics 37 et 41 sur le cardinal de Luynes et l'abbé de Gua et Picard sont plus présents avant la Révolution tout le long de la structure de l'éloge.

## Postérité

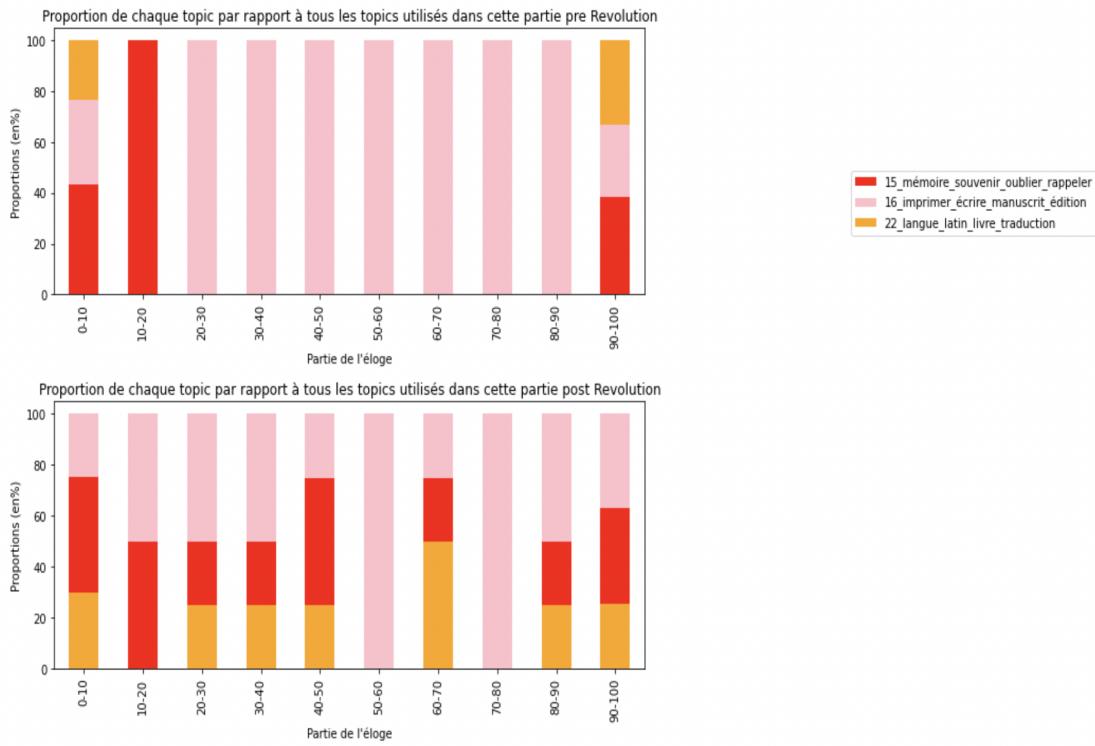


FIGURE 3.10 – Répartition des topics de postérité

Au niveau de la structure de l'éloge concernant la catégorie topic, nous observons que les 20 à 90% du texte sont exclusivement occupés par le topic *imprimer\_écrire\_manuscrit édition* avant la Révolution. Quant au 10-20% du texte, ils sont consacrés au topic *mémoire\_souvenir\_oublier\_rappeler*. La structure après la Révolution est différente dans le sens où les trois topics sur la postérité sont évoqués tout le long de l'éloge. Le topic sur la langue et la traduction est bien plus présent après la Révolution tout le long de l'éloge.

## Méthode

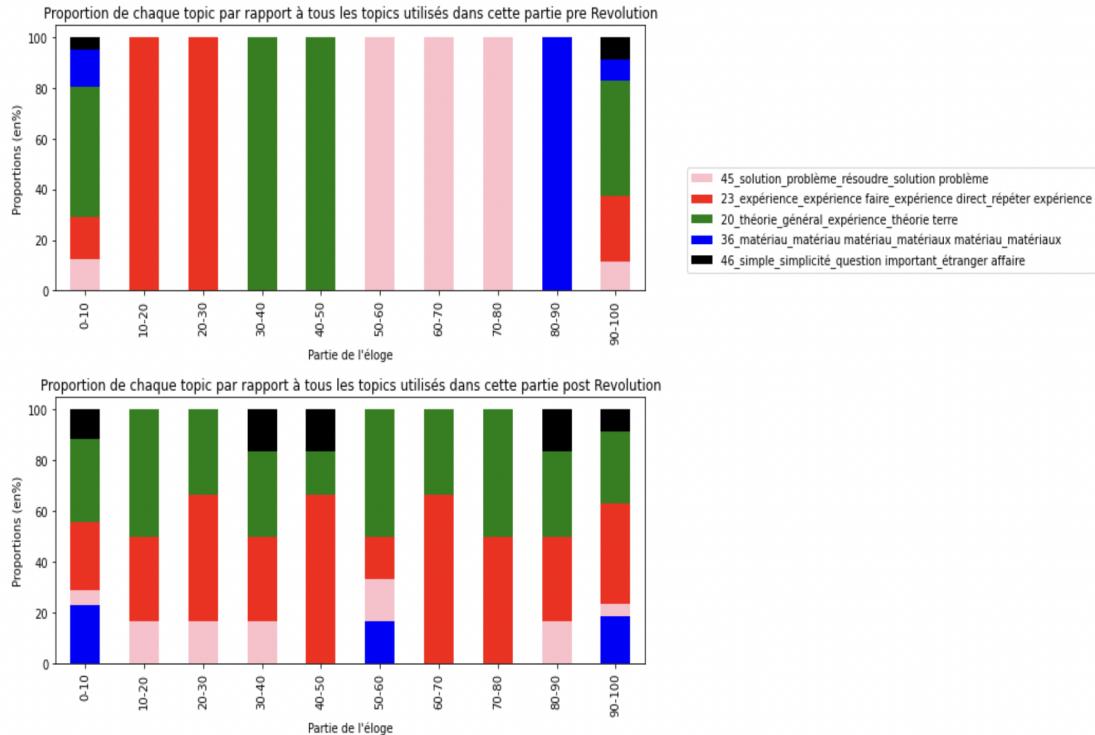


FIGURE 3.11 – Répartition des topics de méthode

Nous pouvons voir dans les topics de méthode dans la figure 3.11, qu'avant la Révolution il y a une progression dans l'évocation des topics. Les pour-cent de 10 à 30 sont constitués de l'expérience, les 20 pour-cent suivants par la théorie, les 30 d'après sur la solution du problème et enfin les avant-derniers 10 pour-cent sont sur les matériaux. Alors qu'après la Révolution c'est plus réparti, avec l'expérience et la théorie mentionné tout le long.

## Sciences

Dans la structure concernant l'importance de chaque discipline dans les différentes parties de l'éloge, peu de différences sont visibles. Dans chaque pourcentage du texte les différentes sciences sont représentées de manière à peu près égale.

## 3.2 Close Reading

### 3.2.1 Littérature

Le livre écrit par Charles Paul en 1980 prénommé *Science and Immortality* fait une étude des éloges écrits à l'Académie entre 1699 et 1790. Il décrit les éloges et comment ils ont créé l'image d'un scientifique désintéressé à la recherche de la vérité. Il étudie les procédés dont les secrétaires se servent pour transmettre cette image mais aussi rendre compte de l'activité scientifique de la France avant la Révolution.

Dans Paul[15], la structure rhétorique de l'éloge est expliquée. Elle rompt avec la structure traditionnelle de l'éloge qui est composée de 7 éléments : l'exordium, la proposition, la division, la narration, la confirmation, la réfutation et enfin la péroration. Au lieu de cela maintenant

l'éloge ressemble à une narration de la vie et du travail du savant mis à l'honneur. Le seul élément qui reste de l'éloge traditionnel est la péroraison. Dans cette péroraison les savants sont représentés comme des modèles de vertu et de sagesse dont il faudrait suivre l'exemple. Au lieu de l'exordium qui est supposé rendre le public intéressé à écouter la suite du discours, les éloges commencent abruptement par une déclaration sur la naissance ou les titres du savant. Ces éloges retracent la vie du savant mais accentuent les victoires afin de minimiser les défaites

Paul mentionne également que Condorcet fait régulièrement des remarques sur le contexte actuel qu'il soit politique, sur l'éducation ou sur les traditions qu'il estime raisonnables ou non, pertinentes ou démodées. Il alterne entre des évocations de la science sobre et des tirades contre la résistance des hommes contre les progrès de la science. Il se sert de ces éloges pour faire la propagande de la philosophie de la science et des Lumières.

Condorcet, contrairement à ses prédécesseurs, sous l'influence de la vague préromantique qui domine ses années en tant que secrétaire utilise des apostrophes, des hyperboles et des interjections. Il fait aussi preuve d'une sympathie renforcée pour les malheurs affrontés par les sujets des éloges.

Le deuxième livre dont nous nous sommes servis est écrit par Maurice Crosland en 1992 et appelé *Science under control-The French Academy of Sciences 1795-1914*. Ce livre essaye de comprendre l'Académie des sciences et ses procédures, comment elle interagit avec les différents régimes politiques du dix-neuvième siècle et façonne les carrières scientifiques de ce siècle. Crosland [4], permet de comprendre mieux le contexte de l'Académie et des éloges après la Révolution. Il permet aussi de comprendre qui sont les académiciens après la Révolution, la face publique et privée.

Ainsi après la fermeture de l'Académie à cause de la Révolution, la première classe d'Académiciens est constituée de ceux de l'ancienne Académie Royale qui ont survécu à la Révolution. Quand ces hommes sont morts, la classe qui les a remplacés est une nouvelle classe d'académiciens issus des écoles fondées après la Révolution, comme l'École Polytechnique, que l'on considère comme la première classe de professionnels de la science. Cependant des restes de l'Ancien Régime subsistent avec la création après la Restauration des Bourbons de la section des *Académiciens libres* afin de réintroduire les valeurs de l'Ancien Régime. Dans cette section nous pouvons y retrouver des ducs. Les Académiciens ne communiquaient que peu leurs vues sur la politique et la religion après la Révolution. Le monde des sciences soulignait plutôt les réussites et le talent que la naissance afin de le représenter comme une méritocratie. L'accès à l'éducation est plus facile après la Révolution, permettant qu'une grande partie des académiciens soient de la classe moyenne mais il subsiste qu'une situation confortable et un intérêt dans la famille pour les sciences aident à encourager les jeunes à faire des études. La science devient une carrière après la Révolution, ce qui était peu le cas avant. Des crédits scientifiques étaient préférés à des familles puissantes dans la science cependant cela reste plus aisément d'intégrer l'Académie en ayant des connexions à l'intérieur. Particulièrement à travers le parrainage et la relation de professeur élève.

Il est convoyé une image du savant qui pouvait vivre frugalement sur un revenu minimum, les savants de l'Académie percevaient un revenu de 1'500 francs. Tout cela pour éviter les opportunistes et présenter le savant comme un homme désintéressé.

Très peu d'académiciens après la Révolution prennent des positions publiques concernant la politique. Cependant il y a quelques exceptions comme Fourcroy qui a été élu à la Convention et faisait partie du club des jacobins ou Monge ouvertement Bonapartiste.

Le choix des savants dont est fait l'éloge est également une indication de la figure savante projetée par l'Académie. En effet le nombre d'académiciens morts chaque année étant supérieur au nombre d'éloges que les secrétaires pouvaient réciter chaque année, ils choisissaient le savant

qu'ils loueraient en fonction de ce qu'ils voulaient appuyer. C'est ce que fait par exemple Arago en optant à quatre reprises différentes de prononcer l'éloge d'académiciens qui ont contribué à la Révolution de 1789 et à la Première République. En effet pour Arago, un parti pris politique dans la vie d'un scientifique ne pouvait que plus l'accomplir dans sa carrière[4]. Montrant que le choix des savants mis à l'honneur est aussi une représentation de la grandeur savante transmise par l'Académie.

Les séances publiques où étaient proclamés les éloges étaient une des rares occasions où l'on voyait publiquement les académiciens. En effet, leurs réunions étaient souvent tenues à huis clos par crainte de se faire voler ses idées ou de se faire critiquer avant même qu'elles aient été publiées.

Jusqu'à 1802, les éloges devaient être assez courts car c'était une petite partie d'une séance consacrée à l'entièreté de l'Institut, donc également aux autres classes. Le matériel était donc compressé. Mais grâce à la réforme de l'Institut de 1803, les différentes classes pouvaient tenir leur propre réunion et cela a ainsi permis les éloges d'être plus longs et plus élaborés. De plus à partir de ce moment, les deux secrétaires sont élus de manière permanente en partie pour leurs habiletés littéraires. Après la Restauration en 1815, il y a un véritable effort d'intéresser le public dans les sciences. Les secrétaires étaient très conscients qu'il fallait représenter l'académicien mort de manière positive mais également la science aussi.

Dorinda Outram publie en 1978 une étude des éloges de Cuvier prénommée *The language of natural power : the "eloges" of Georges Cuvier and the public language of nineteenth century sciences*. Dans cette étude elle examine quelle représentation ressort des académiciens à travers les éloges de Cuvier. Cuvier est secrétaire de l'Académie au début du dix-neuvième siècle.

Outram [13] étudie les éloges de Cuvier, secrétaire de l'Académie de 1799 à 1801 et de 1803 à 1830. Elle évoque le langage semi-pastoral de ces éloges ainsi que la représentation du savant comme d'un philosophe apolitique et asocial. Le savant mène une vie loin des distractions, de la corruption, des intrigues politiques. Il mène une vie d'ermite, ignorant du monde et ne recherchant pas la gloire. L'accomplissement d'une vocation scientifique impliquait nécessairement le dépassement impitoyable et acharné d'énormes désavantages sociaux. Cuvier accuse la difficulté d'obtenir reconnaissance scientifique qui lance les scientifiques dans les eaux politiques. Le pouvoir du scientifique atteint la gloire en se servant du "clean power" qui est celui de la science. Nous pouvons retrouver la norme de désintéressement dans le langage pastoral, le besoin pour un scientifique de s'isoler souvent ruralement. Cependant il ne faut pas être extrême dans sa solitude car elle empêche la communication entre savants qui crée des failles dans certains travaux. Le discours du Cuvier représente également la science comme inviolable et sans valeur car elle opère dans un royaume différent de celui de la politique humaine.

### 3.2.2 Lecture des deux éloges

Finalement la lecture des deux éloges a permis de révéler certains indices.

Tout d'abord, nous pouvons noter que l'éloge de Praslin est bien plus court que celui de Volta. En effet le premier se condense en 8 pages, alors que le deuxième s'étale sur 17.

Une autre différence dans les deux éloges est la présence de l'opinion du secrétaire. En effet chez Condorcet plusieurs apartés telle que "Ces objets, je le sais, sont étrangers à ceux dont l'Académie s'occupe ; mais le titre d'homme, de citoyen, en imposant l'obligation commune de se dévouer au bien de la patrie, donne à tous, par une conséquence nécessaire, le droit d'avoir une opinion sur les intérêts publics, et ce droit est inséparable de celui de la dire" sont présentes. Son opinion est claire sur la société actuelle qu'il considère comme intéressée et peu ouverte au progrès. Il utilise également le terme "lumières" pour représenter les sciences et le savoir.

Chez Volta, on retrouve l'utilisation de la première personne du singulier mais ce n'est pas pour manifester l'opinion d'Arago mais pour faire des observations telles que : "Je regrette de ne pouvoir dire, avec une entière certitude, à qui appartient cette expérience capitale".

Nous retrouvons dans ces deux éloges la louange de qualités que Praslin et Volta auraient en commun. Ces qualités sont la modestie et l'amour de l'étude. Dans les deux éloges, est également louée l'importance que ces deux savants portent à l'éducation de leurs successeurs. Pour Praslin, c'est dans les réformes qu'il crée au ministère de la marine en y obligeant l'étude de la théorie. Chez Volta, la même chose est retrouvée dans ses qualités de professeur. Praslin est également célébré pour sa liberté de pensée, son opposition à certaines pratiques de son époque telle que la politique dans les colonies. Il est loué pour ses pensées avant-gardistes pour son époque. Condorcet met également en avant sa probité, sa franchise, sa clairvoyance et la façon dont il stimule la recherche en France pour égaler voire dépasser les autres pays. Cela est visible dans cette phrase issue de son éloge "en cherchant à répandre l'instruction, en employant leur pouvoir à favoriser les progrès des lumières, ont acquis des droits à notre reconnaissance". Finalement, il fait l'éloge de son assiduité à l'Académie pour continuer à s'améliorer dans les sciences. Volta est quant à lui mis en avant pour la vie simple qu'il mène ainsi que son génie.

Nous pouvons également voir des différences dans la structure des deux éloges. Celui écrit par Condorcet commence abruptement par les titres de Praslin ainsi que sa naissance. Pour Volta, il commence par une phrase introductory sur la recherche scientifique qu'il a menée sur la pile. Il se poursuit ensuite sur l'ensemble de la vie scientifique qu'il a mené de son éducation jusqu'à ses recherches et découvertes les plus importantes. Cela concerne la majeure partie de son éloge : 14 pages sur 17. Ensuite Arago retrace plus brièvement sa vie chronologiquement. Il évoque son mariage et sa famille, les différentes récompenses qu'il a reçues, les circonstances qui l'ont fait déménager et de ses qualités. Son éloge s'achève par l'information sur la personne qui va le remplacer à l'Académie. L'éloge de Praslin diffère car elle est complètement chronologique retracant les circonstances qui ont fait que Praslin a accédé à différents postes et ce qu'il a accompli à ces différents postes. Il se termine par une célébration des qualités de cet homme, nous retrouvons bien la péroration évoquée par Paul [15]. La Science est beaucoup plus mise à l'honneur dans l'éloge du dix-neuvième siècle mais cela est probablement dû au fait que Praslin soit un diplomate et pas vraiment un savant plutôt une personne s'intéressant aux sciences.

Dans les deux éloges, sont contextualisées les différentes actions des deux académiciens mais de manière plus exhaustive chez Volta. Mais une différence notable est que chez Volta, beaucoup de savants différents sont évoqués et comment ils ont contribué ou remis en question la recherche de Volta. Pour Praslin, seul un de ses contemporains est évoqué pour sa contribution.

Sur le sentiment qui ressort de ces éloges, celui publié par Condorcet ressort un peu plus négativement dans ses avis sur la politique actuelle. Les malheurs arrivant à Praslin sont tous tournés vers comment il les a affrontés, mais en accusant la société de l'époque. Alors que dans celui d'Arago, tout ce qui est négatif est tourné à l'avantage de Volta.

# Chapitre 4

## Discussion

Le nombre d'éloges publiés par an a peu évolué au cours du temps. Nous pouvons voir la fermeture de l'Académie entre 1793 et 1795, c'est la seule période de plusieurs années de suite où aucun éloge n'est publié. Lorsque l'Académie rouvre en 1795, elle passe à deux secrétaires perpétuels un pour les sciences physiques et l'autre pour les sciences mathématiques. Mais ces secrétaires changent régulièrement. Finalement en 1803, les secrétaires sont élus de manière permanente.

Après la Révolution, les éloges semblent plus longs, particulièrement à partir de 1800. Cela peut être dû à la réforme de l'Académie de 1803. Deux secrétaires sont nommés pour leurs qualités oratoires et littéraires [4]. De plus, on passe d'une séance publique commune à toutes les académies, à une séance publique par académie. Les éloges ont donc plus de temps alloués pour eux lors de ces séances et peuvent être plus élaborés. Cela explique sûrement cette légère hausse dans la longueur de ces éloges. Nous voyons aussi une lente décroissance avant la Révolution avec un sursaut en 1790 juste après la Révolution.

Les mesures qui n'ont pas fonctionné montrent les limites des outils actuellement disponibles. De plus les éloges sont un type de discours très particulier. Ils sont d'abord lus aux séances publiques puis retranscrits. C'est une forme de discours très emphatique, un style oral et par définition élogieux. Ils célèbrent la vie d'un homme décédé et sont donc très différents de la plupart des autres textes. Nous avons essayé par exemple de tester la subjectivité du discours avec l'utilisation des pronoms à la première personne du singulier. Nous n'avons pu distinguer aucune différence entre l'avant et l'après la Révolution. Cependant Condorcet était connu pour faire des apartés où il manifestait son opinion [15]. C'est aussi ce que nous avons pu constater dans l'éloge de Praslin où nous retrouvions plusieurs de ces digressions ce qui n'était pas le cas chez Volta. Il faudrait donc trouver une meilleure manière de relever ces apartés dans les éloges.

Une autre méthode dont nous avons pu constater les limites est celle utilisée pour tenter d'analyser l'évocation d'autres scientifiques ou lieux. Pour cela, nous avons essayé de relever la proportion de noms propres utilisés. Cela n'a pas fourni de résultats. Et en effet, dans les deux éloges lus beaucoup de noms propres étaient utilisés, bien que l'éloge de Volta contienne une quantité infiniment supérieure d'évocation de scientifiques différents et de lieux différents que celui de Praslin. Pour essayer d'améliorer cette mesure, du Named Entity Recognition pourrait être essayé et ainsi relever le nombre de lieux différents ou de personnes différentes évoqués au cours des éloges. De l'encodage en XML\_TEI pourrait également être utilisé.

Pour tenter de voir l'évocation du passé, nous avons tenté de combiner une recherche sur les verbes au passé et sur les adverbes qualifiant la temporalité. Le but était de voir si le passé était plus ou moins évoqué avec la Révolution, si elle avait eu un impact. En lisant les deux éloges, nous avons pu constater une évocation beaucoup plus précise du passé et d'un passé

bien plus lointain après la Révolution. Les résultats n'ont pas été concluants, la Révolution n'a pas l'air d'avoir eu d'impact. En revanche l'évocation du futur proche et du présent a montré une augmentation après la Révolution dans les adverbes. Cela suggère une volonté au-delà de juste faire l'éloge de la vie du savant plus importante après la Révolution.

Après la Révolution on a constaté une différence significative dans les adverbes de *Tout* et de *Certain* qui sont plus présents après la Révolution alors que tous les adverbes moins certains sont plus présents en proportion avant la Révolution. Le discours est étonnamment moins assertif avant la Révolution. Cela est peut être lié à la volonté de représenter la science comme une autorité supérieure aux politiques humaines [13]. L'autorité de la Science voit son rapport vis à vis de la société qui change, elle se détache des structures qui régissent la société humaine après la Révolution.

Avant la Révolution, le discours contient plus d'adverbes négatifs en comparaison à après. Cela peut correspondre aux tirades de Condorcet contre le régime et la société de l'époque qui se hissait contre le progrès des sciences. L'inversion en 1790 juste après la Révolution est intéressante à noter, tous les adverbes employés cette année-là sont négatifs. En lisant l'éloge de Praslin, nous avons pu retrouver un sentiment négatif. Alors que dans Volta, tout ce qui a pu arriver de négatif à Volta est immédiatement tourné de manière positive, comment il a réussi à surmonter cet obstacle plutôt que de faire un réquisitoire contre la société.

Une autre mesure que nous n'avons pas réussi à créer concerne les figures de style. Nous n'avons pas réussi à trouver ou à créer d'outils pour détecter l'ironie, le sarcasme ou encore l'hyperbole. À travers la littérature, nous avons pu comprendre que Condorcet avait souvent recours à des figures de style telles que l'hyperbole, l'apostrophe. Nous avons pu en retrouver certaines en lisant l'éloge de Praslin. Il s'agit d'un enjeu majeur pour comprendre ces textes aux qualités si singulières. Il y a très peu d'outils actuellement disponibles pour étudier le langage figuratif dans un texte et ils n'existent souvent que pour l'anglais ou alors il nécessite des corpus déjà annotés. Un article publié en 2018, a commencé à étudier la détection d'hyperbole sur un corpus anglais [22]. Un autre article datant de 2020 tente de détecter l'ironie et le sarcasme sur les réseaux sociaux [1].

Nous pouvons voir après la Révolution que la science prend une part plus importante dans l'éloge qu'avant la Révolution. Au contraire de la société qui occupe une partie plus importante avant. Cela montre la volonté de mettre à l'honneur la science indépendante de la politique actuelle et un pouvoir au-dessus et pur [13]. La personne quand à elle occupe la même proportion avant et après la Révolution. Nous pouvons voir qu'avant et après la Révolution, la partie où elle est la plus importante est dans les 10 premiers et 10 derniers pourcents. Nous observons la structure de la péroration dite chez Paul et observée dans les deux éloges, où à la fin de l'éloge les qualités de l'académicien décédé sont mises à l'honneur et où il est érigé en exemple. Nous pouvons également distinguer dans les structures globales des éloges les mêmes tendances. C'est-à-dire que dans les deux éloges, l'évocation de la personne semble être en forme de U, elle est plus discutée au début et à la fin de l'éloge et de moins en moins parlé vers le milieu. Et nous constatons la tendance inverse dans la science dans les deux cas où la science est moins discutée aux extrémités de l'éloge mais est de plus en plus discuté vers le milieu de l'éloge. Il y a une certaine continuité dans la structure thématique de l'éloge. La même quantité de topics non classables est présente avant et après l'éloge de même pour la proportion de topics sur la méthode et la postérité.

Dans l'évolution des topics politique, nous pouvons voir l'apotéosisation de la science après la Révolution dans la diminution de leur utilisation. Le pic observé en 1777 peut correspondre à deux choses : la Révolution Américaine qui a eu lieu un an plus tôt ou alors la première année où Condorcet rédige des éloges en tant que secrétaire officiel de l'Académie après avoir subi des

revers les années précédentes [14]. Et le pic de 1790 peut être relié à la Révolution qui débute en 1789. Finalement en 1815 a lieu la Restauration de la monarchie et des Bourbons en France. Le fait que ce soit des pics et pas une augmentation constante ou une constante est sûrement lié aux différents secrétaires qui ont différentes images du savant après la Révolution. Certains le souhaitent apolitique comme Cuvier[13] et d'autres considèrent que des opinions politiques ne font que parfaire leur carrière scientifique comme Arago [4].

La religion connaît les mêmes pics que la politique avant la Révolution et après n'est quasiment plus évoquée. Les deux pics peuvent être liés aux mêmes événements que pour la politique avec les bouleversements que la Révolution a engendré pour la religion. Ce résultat est en accord avec la volonté après la Révolution de moins mentionner les vues religieuses des savants [4]. Nous retrouvons également la volonté de représenter le scientifique comme indépendant des structures qui organisent les groupes humains [13].

Par ailleurs nous observons que dans la structure de l'éloge la manière d'évoquer la méthode est différente. Alors qu'avant la Révolution, on commence par évoquer l'expérience puis la théorie, on continue avec la solution et on termine par les matériaux, après la Révolution tout est mélangé. De plus l'expérience et la théorie prennent une partie importante de la proportion allouée à la méthode dans les éloges. Comme l'utilisation de ces topics n'a pas vraiment évolué au cours du temps à l'exception de l'expérience, cela veut juste dire que la structure dans la manière de représenter la pratique de la science a évolué avant et après la Révolution. Avant la Révolution, une chronologie logique de la recherche est distinguable. Cette continuité dans les topics de méthode peut être expliquée par la présence des mêmes académiciens avant et juste après la Révolution. En effet comme vu chez Crosland [4], les premiers académiciens sont les mêmes que juste avant la Révolution expliquant peut être une certaine continuité dans la manière de pratiquer la science.

Dans la catégorie postérité, les topics n'ont pas connu d'évolution notable au cours du temps. En revanche dans la structure, la différence est visible. Les 20% au 90% sont exclusivement occupés avant la Révolution par le topic sur l'impression et la rédaction de manuscrits. Alors qu'après la Révolution la structure est plus mélangée. Le topic sur la mémoire et le souvenir concerne les Mémoires que le savant écrit ainsi que le souvenir qu'il laisse derrière lui. Il y a donc une partie qui se superpose un petit peu avec le topic sur la rédaction de manuscrits mais il est plus orienté sur ce que le savant veut laisser à la postérité. Nous voyons aussi que le topic sur la langue est beaucoup plus présent tout le long de l'éloge après la Révolution ou avant il n'est mentionné que dans les premières ou les dernières phrases.

Concernant les sciences, nous pouvons constater une évolution des sciences représentées à l'Académie. En effet un pic une année est sûrement dû à la mort d'un académicien de cette discipline. Par exemple en 1831, le topic d'électricité connaît un pic fulgurant à cause de la mort de Volta. Ces évolutions montrent quelles disciplines sont mises à l'honneur à l'Académie avant et après la Révolution. Par exemple après la Révolution, nous voyons une diminution de l'utilisation des topics sur le philosophie ainsi que sur la médecine. Les topics de médecine peuvent représenter deux types de contenu. Un académicien ayant fait appel à un médecin comme dans la phrase : "Des palpitations, sur lesquelles un médecin ne pouvait se méprendre, lui annoncèrent son sort" ou alors un académicien étant médecin comme dans la phrase suivante : "M Bordenave était échevin de Paris, et aucun chirurgien avant lui n'avait été élevé à cette place". Cependant la plupart des phrases de la catégorie médecine concernent des académiciens médecins. Cela peut s'expliquer par la création d'une académie de médecine en 1820 qui écrit ses propres éloges [21]. Nous voyons donc après la Révolution, la disparition progressive de la philosophie. Au contraire de la médecine et de la philosophie, la physique les mathématiques, l'astronomie, la géologie, la biologie et la chimie connaissent des augmentations dans leur évo-

cation après la Révolution. Nous assistons à une diversification des sciences après la Révolution. Cela ne veut pas forcément dire qu'à l'Académie moins de médecins ou de philosophes y entraient, mais qu'il y a un changement dans les sciences que celle-ci veut représenter. Jusqu'à la fin du dix-huitième siècle, les philosophes sont les scientifiques de l'Ancien Régime et les différentes disciplines ne sont pas précisément définies. Après, on distingue des disciplines et donc des praticiens différents [5]. L'Académie se montre plus diverse dans les sciences qu'elle met à l'honneur. Cela peut être aussi lié au fait qu'avant la Révolution, un seul secrétaire s'occupait de faire tous les éloges alors qu'après il y en a un pour la physique et un pour les mathématiques. Cela fait du sens qu'ils choisissent de faire des éloges plutôt pour les scientifiques des disciplines qui les concernent que pour un philosophe.

L'augmentation de l'évocation de la jeunesse après la Révolution est aussi visible dans la lecture même des éloges. En effet dans celui de Volta, son éducation est évoquée ainsi que ses rencontres avec d'autres scientifiques et ses voyages. Le chemin qu'il a parcouru pour arriver à l'excellence scientifique est détaillé. En revanche chez Praslin, aucune évocation de sa formation n'est perçue à l'exception de ses débuts dans l'armée. Sa nature est brièvement évoquée l'inclinant à toujours apprendre et sa fréquentation assidue de l'Académie mais dans l'éloge, son génie est beaucoup plus ressenti comme inné que reproductible. Alors que Volta, bien que son don naturel pour les sciences soit évoqué, son parcours pour se développer intellectuellement est plus clair. L'accès à l'Académie après la Révolution semble plus représenté comme le fruit de beaucoup de travail. Avant la Révolution, la famille du savant est beaucoup plus présente dans l'éloge. Nous avons également pu retrouver chez Praslin dès la première phrase les titres qu'il a obtenu auprès du roi. Le savant est tout de suite représenté par sa famille alors que chez Volta cela vient en dernière position quand est déjà fait l'éloge de tous ses accomplissements scientifiques. On évoque qu'il a atteint dans la cour de Napoléon une place enviable mais cela est au second plan. Cela est probablement biaisé par le fait que Praslin soit un diplomate plutôt qu'un savant à proprement parlé mais il reste que l'Académie est plus représentée comme une méritocratie après la Révolution.

Nous pouvons voir qu'après la Révolution la famille du savant est beaucoup moins mentionnée le long de l'éloge. Le savant est représenté en fonction de ses accomplissements personnels plutôt que ce que sa naissance lui a apporté. La présence de la carrière est aussi plus élevée tout le long de l'éloge. Après la Révolution, le fil conducteur de l'éloge est la science et ce qui a permis au savant d'en arriver là. C'est ce qui était visible chez Volta, où tout son parcours scientifique était détaillé. Chez Praslin, ses apports à la France sont évoqués et ses qualités qui lui ont permis de réaliser sa carrière, mais comment il a pu y arriver est peu mentionné.

En proportion après la Révolution, le progrès de la science occupe une partie plus importante de la catégorie figure-type dans la deuxième moitié de l'éloge. Au contraire avant la Révolution, le génie de l'homme est plus mis en avant particulièrement vers la fin de l'éloge. Comme vu dans la lecture des éloges, cela montre qu'après la Révolution on insiste sur les progrès que la science peut apporter. La science est bien plus mise en avant après la Révolution, l'individualité de l'homme et de son génie est moins marquée qu'avant la Révolution. La qualification d'artiste et de savant est également plus présente après la Révolution vers la fin de l'éloge. Les topics sur l'abbé de Gua, l'abbé Picard et le cardinal de Luynes sont plus présents avant la Révolution. Condorcet rédige l'éloge de Luynes en 1788 et celui de l'abbé de Gua en 1786, justifiant leur présence plus grande avant la Révolution. Cependant le topic sur le cardinal reste un peu présent dans plusieurs parties de l'éloge après la Révolution. Ce topic sur le cardinal contient aussi quelques phrases en rapport avec la société religieuse de l'époque : "Le grand maître lui fit grâce ; mais cette grâce devait être confirmée par le pape". La religion est donc encore quelque peu présente après la Révolution le long de l'éloge.

# Chapitre 5

## Conclusion

Dans ce projet, nous avons pu développer des outils afin d'orienter la lecture historique des éloges de l'Académie des sciences. Cela a permis d'étudier l'évolution de la figure du savant transmise par l'Académie avant et après la Révolution.

Nous avons mis à jour en particulier une structure thématique qui complémente la structure rhétorique que nous avons pu retrouver dans la différente littérature. Dans cette structure thématique, nous avons pu voir une évolution de la figure du savant et de l'Académie des sciences. Après la Révolution, cette dernière se représente plus comme une méritocratie. Pour cela, elle représente plus le savant à travers son éducation et le chemin qu'il a parcouru dans sa carrière scientifique. La famille du scientifique est moins mentionnée et il est moins présenté en fonction de la place qu'il occupe dans la société. De plus le savant est moins représenté comme un génie que comme quelqu'un qui a travaillé pour parvenir à son niveau. Après la Révolution, nous pouvons également voir la volonté de l'Académie d'intéresser le peuple aux sciences et la professionnalisation des sciences. À travers cette analyse, nous avons pu assister à la diversification de la représentation des sciences après la Révolution et à la création de disciplines distinctes. La science a une part plus importante dans les éloges et est plus mise en avant par les progrès qu'elle apporte au monde et comment c'est une forme de pouvoir supérieure aux politiques humaines.

En étudiant l'évolution du contenu des éloges, nous avons pu observer l'apotilisation de la science souvent mentionnée dans la littérature. Nous avons aussi pu mettre à jour le peu de communication sur les religions de l'Académie après la Révolution. Avec la re-hausse de mentions politiques après la Restauration, nous pouvons aussi voir l'impact des différents régimes sur la représentation que l'Académie fait du savant.

Finalement, nous avons pu voir une continuité dans la manière d'aborder la science. Les différents contenus sur l'expérience ou la théorie restent évoqués en même quantité les éloges. En revanche la manière de retranscrire les éloges à évoluer avec une structure moins chronologique après la Révolution.

Nous avons aussi construit nous-mêmes des outils pour aider à la lecture. La classification des adverbes a permis de voir que le discours était moins assertif avant la Révolution et moins positif. L'évolution de la longueur des éloges a permis de rendre compte des réformes ayant lieu à l'Académie.

Cependant, tous les outils utilisés n'ont pas réussi à produire de résultats concluants. En effet, certains indices que nous avons créés, s'exprimaient autrement dans le texte que par le biais de nos mesures. Ça a été le cas de la subjectivité du discours que nous avons tenté de mesurer avec l'usage de la première personne. Cela montre que l'ordinateur aide à orienter la lecture historique mais ne suffit pas et qu'il faut le compléter avec une lecture des éloges. Dans

ce cas précis, l'ordinateur a sûrement eu des difficultés à analyser nos documents, car les éloges sont un type de rhétorique très particulier, un discours ensuite retranscrit pour faire la louange d'une personne décédée.

Nous avons également manqué d'outils pour certaines mesures en particulier l'étude du langage figuratif.

Pour parfaire notre compréhension de ses résultats, nous sommes allés chercher dans la littérature et avons fait une lecture comparée de deux éloges. Certains résultats ont pu être retrouvés, d'autres non, montrant bien les limites actuels du NLP en français et la difficulté de faire des analyses sur une rhétorique aussi particulière que l'éloge est.

## Future Work

Nous voyons donc à travers ce projet les limites de nos mesures et de l'ordinateur. Pour améliorer la mesure sur les noms propres qui avaient pour but d'identifier les scientifiques ou lieux différents exprimés dans le texte, deux options peuvent être considérées. La première est l'usage de Named Entity Recognition pour sortir les différentes personnes, organisation, lieux qui sortent du texte. La deuxième serait d'enrichir les textes avec de l'encodage en XML-TEI effectué à la main afin de reconnaître les lieux et les dates.

Pour améliorer l'analyse stylistique, il faudrait développer des moyens de reconnaître les différentes figures de style telles que l'ironie ou l'hyperbole. Pour cela, de l'analyse supervisée pourrait être essayée afin de créer un modèle pour reconnaître ces figures de style.

Trouver d'autres corpus pour recouper les résultats obtenus avec d'autres sources pourraient également être une piste à suivre.

## Acknowledgement

Je tiens à remercier Simon Dumas Primbault de m'avoir guidé et aidé tout le long du projet et Jérôme Baudry et Antoine Bosselut de l'avoir supervisé.

## Annexe A

### Statistiques sur les éloges

#### Style des éloges

Dans cette appendice, sont listés les graphes des mesures sur le style des éloges qui n'ont pas fourni de résultats.

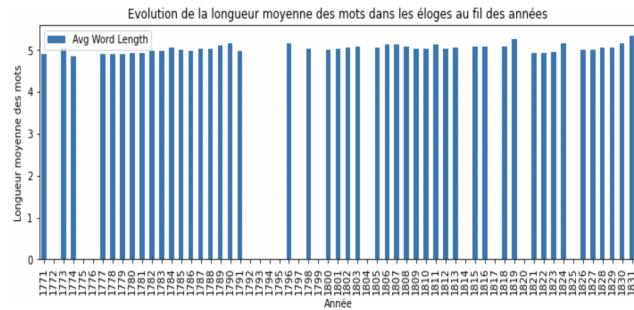


FIGURE A.1 – Évolution de la longueur moyenne des mots par éloges

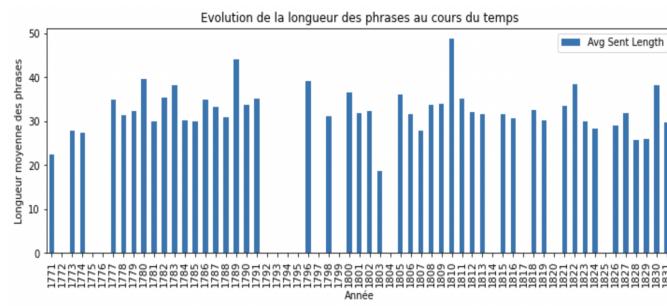


FIGURE A.2 – Évolution de la longueur moyenne des phrases dans les éloges au cours du temps

#### Part-Of-Speech

Dans cette partie de l'appendice, sont présents les graphes des mesures qui n'ont pas fourni de résultats dans le Part-Of-Speech.

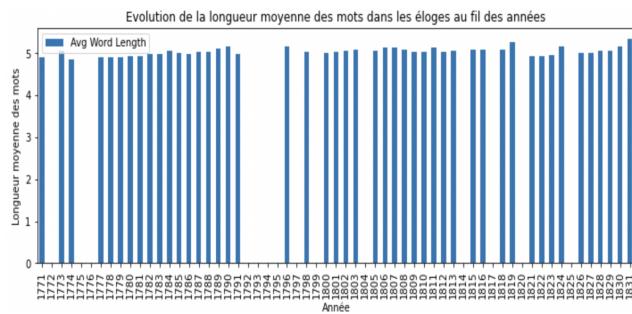


FIGURE A.3 – Évolution de la longueur moyenne des mots dans les éloges au cours du temps

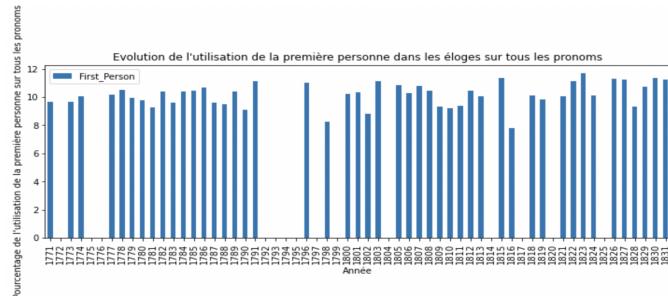


FIGURE A.4 – Évolution de l'utilisation de la première personne

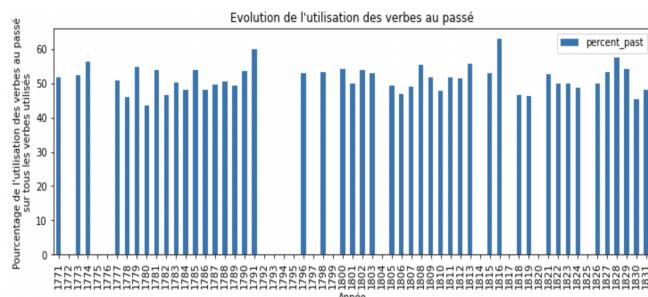


FIGURE A.5 – Évolution de l'utilisation du passé

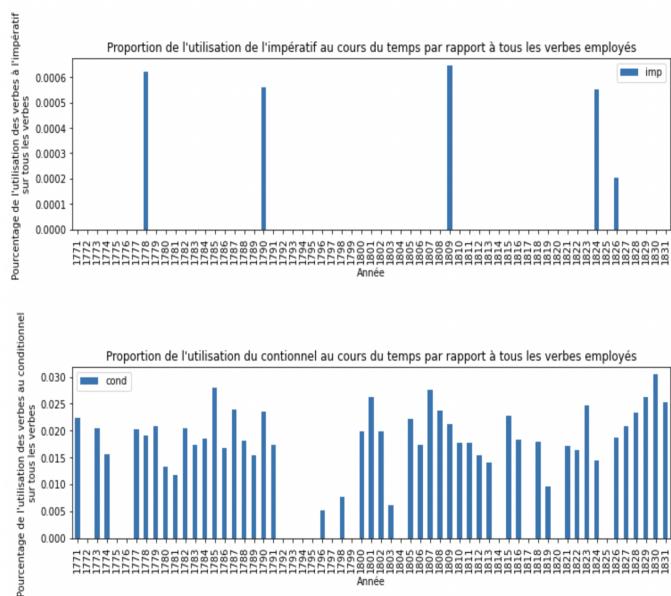


FIGURE A.6 – Évolution de l'utilisation du conditionnel et de l'impératif

# Annexe B

## Contenu des éloges

### Évolution des éloges

Cette section contient différents graphes sur l'évolution des topics au cours du temps.

#### Sciences

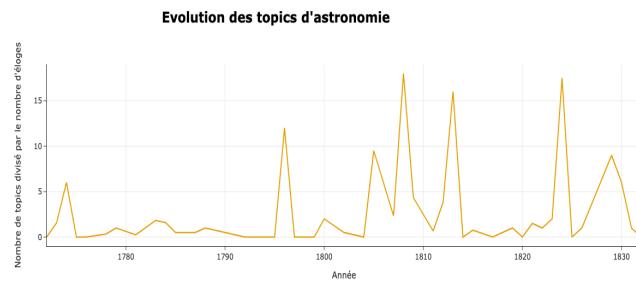


FIGURE B.1 – Évolution des topics d'astronomie

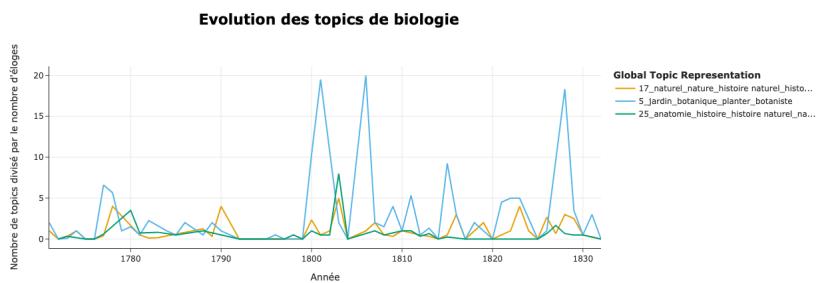


FIGURE B.2 – Évolution des topics de biologie

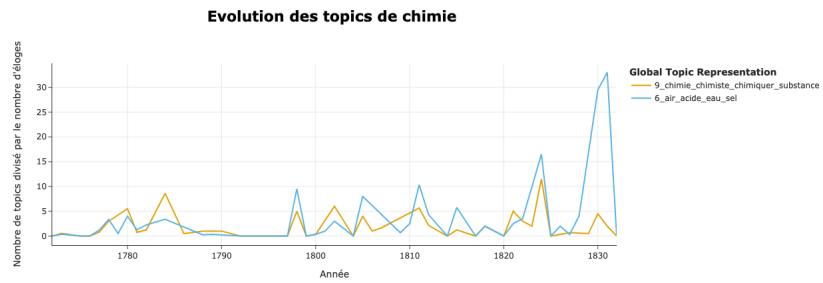


FIGURE B.3 – Évolution des topics de chimie

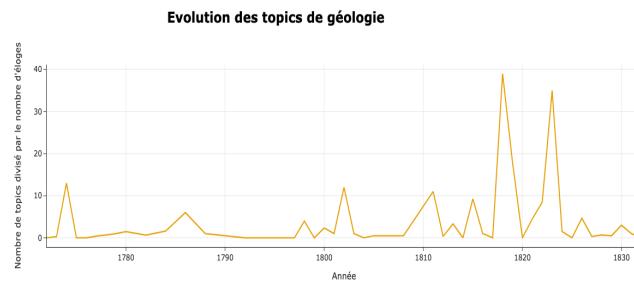


FIGURE B.4 – Évolution des topics de géologie

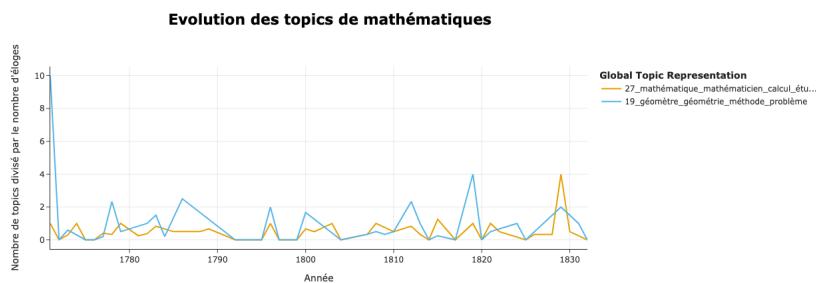


FIGURE B.5 – Évolution des topics de mathématiques



FIGURE B.6 – Évolution des topics de médecine



FIGURE B.7 – Évolution des topics de philosophie

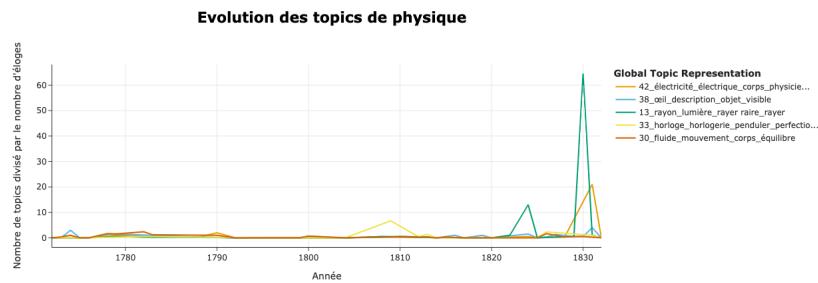


FIGURE B.8 – Évolution des topics de physique

## Vie

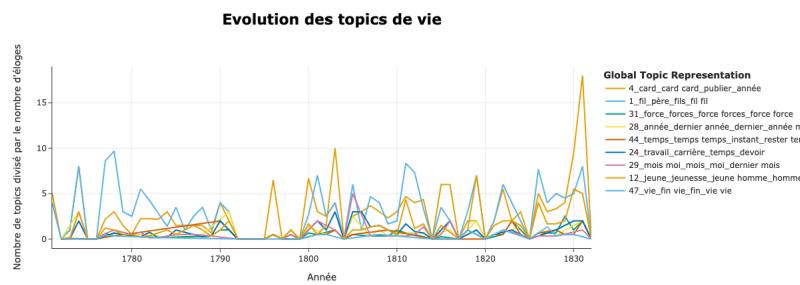


FIGURE B.9 – Évolution des topics de vie

## Méthode

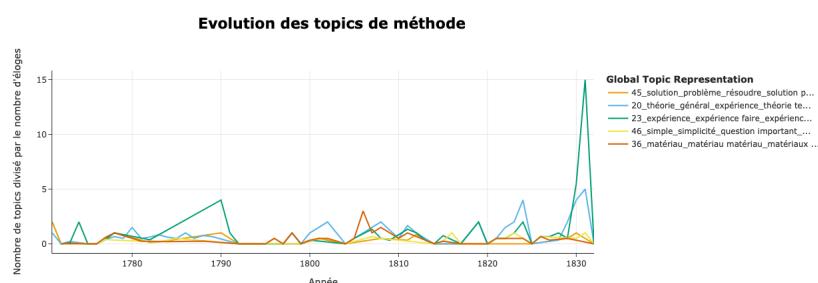


FIGURE B.10 – Évolution des topics de méthode

# Voyage



FIGURE B.11 – Évolution des topics de voyage

L'évolution de tous les topics normés par le nombre d'éloges publiés chaque année est disponible sur [ce lien](#).

# Structure des éloges

Cette section concerne l'évolution de la structure des éloges.

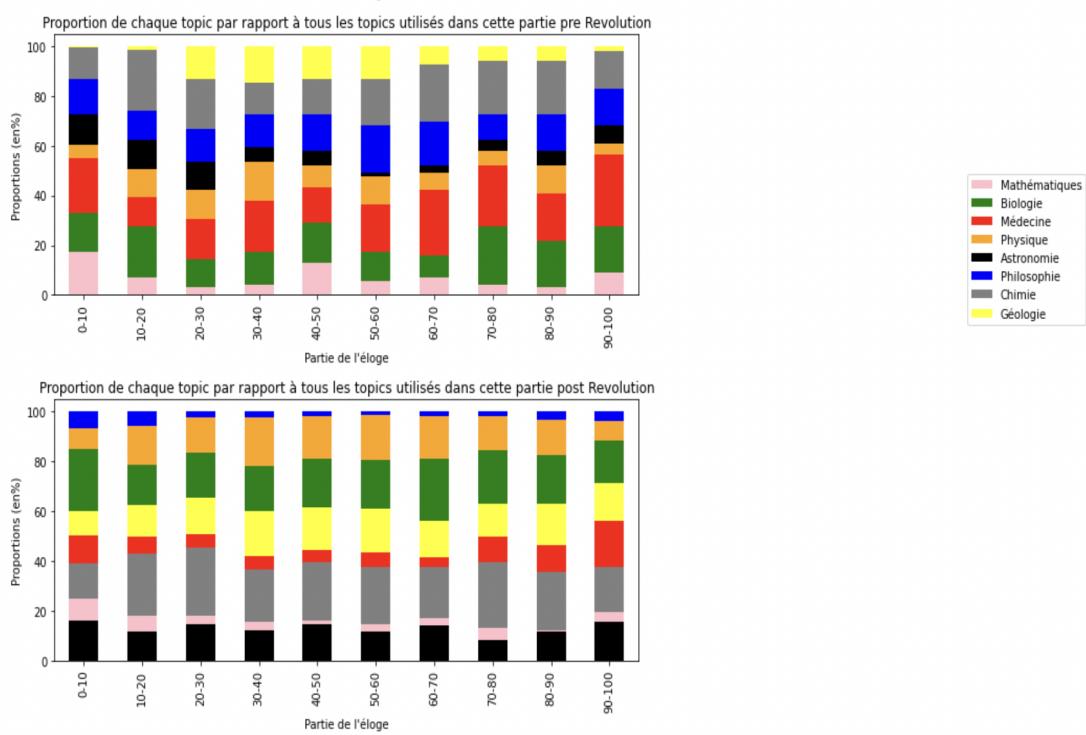


FIGURE B.12 – Évolution des topics de la structure de la catégorie science

# Références Partie Histoire

- [2] BRET. "Figures du savant, XVe-XVIIIe siècle". In : (2016).
- [4] CROSLAND. *Science under Control, The French Academy of Sciences 1795-1914*. 1992.
- [5] Stéphane Van DAMME. *À toutes voiles vers la vérité. Une autre histoire de la philosophie au temps des Lumières*. 2014.
- [6] DELORME. "La vie scientifique à l'époque de Fontenelle d'après les éloges des savants". In : (1937).
- [8] Jo GULDI. "Parliament's Debates about Infrastructure : An Exercise in Using Dynamic Topic Models to Synthesize Historical Change". In : 60.1 (January 2019). Sous la dir. de TECHNOLOGY et CULTURE, p. 1-33.
- [13] Dorinda OUTRAM. "The language of natural power : the "Eloges" of Georges Cuvier and the public language of nineteenth century science". In : (1978).
- [14] Irène PASSERON. "Grandjean de Fouchy, D'Alembert et Condorcet : Tracasseries et arrangements des secrétaires perpétuels". In : (2008).
- [15] PAUL. *Science and Immortality The Eloges of the Paris Academy of Sciences (1699-1791)*. 1980.
- [18] Arnaud SAINT-MARTIN. "Autorité et grandeur savantes à travers les éloges funèbres de l'Académie des sciences à la Belle Époque". In : (2012).
- [19] Académie des SCIENCES. *Histoire de l'Académie des sciences*. URL : <https://www.academie-sciences.fr/fr/Histoire-de-l-Academie-des-sciences/histoire-de-l-academie-des-sciences.html>.
- [20] Lorraine Daston H. Otto SIBUM. "Introduction : Scientific Personae and Their Histories. Science in Context". In : (2003).
- [21] "The self-made Mandarin : The eloges of the French Academy of medecine, 1824-47". In : (1998).
- [23] Alan James Hogarth Michael WITMORE. "Reflexive witnessing : Boyle, The Royal Society and Scientific style". In : (2003).

# Références Partie Code

- [1] Potamias R.A. Siolas G. Stafylopatis A. “A transformer-based approach to irony and sarcasm detection”. In : (2020). DOI : [10.1007/s00521-020-05102-3](https://doi.org/10.1007/s00521-020-05102-3). URL : <https://arxiv.org/abs/1911.10401>.
- [3] Lars BUITINCK et al. “API design for machine learning software : experiences from the scikit-learn project”. In : *ECML PKDD Workshop : Languages for Data Mining and Machine Learning*. 2013, p. 108-122.
- [7] Maarten GROOTENDORST. “BERTopic : Neural topic modeling with a class-based TF-IDF procedure”. In : *arXiv preprint arXiv:2203.05794* (2022).
- [9] François Dominic LARAMÉE. *Introduction à la stylométrie en Python*. 2018. URL : <https://programminghistorian.org/fr/lecons/introduction-a-la-stylometrie-avec-python>.
- [10] Hang LE et al. “FlauBERT : Unsupervised Language Model Pre-training for French”. In : *Proceedings of The 12th Language Resources and Evaluation Conference*. Marseille, France : European Language Resources Association, mai 2020, p. 2479-2490. URL : <https://www.aclweb.org/anthology/2020.lrec-1.302>.
- [11] Steven LORIA. *TextBlob : Simplified Text Processing*. 2020.
- [12] Louis MARTIN et al. “CamemBERT : a Tasty French Language Model”. In : *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020.
- [16] Laurent POINTAL. *TreeTagger Python Wrapper’s documentation !* 2018. URL : <https://treetaggerwrapper.readthedocs.io/en/latest/>.
- [17] Nils REIMERS et Iryna GUREVYCH. “Sentence-BERT : Sentence Embeddings using Siamese BERT-Networks”. In : *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, nov. 2019. URL : [http://arxiv.org/abs/1908.10084](https://arxiv.org/abs/1908.10084).
- [22] Enrica TROIANO et al. “A Computational Exploration of Exaggeration”. In : *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium : Association for Computational Linguistics, oct. 2018, p. 3296-3304. DOI : [10.18653/v1/D18-1367](https://doi.org/10.18653/v1/D18-1367). URL : <https://aclanthology.org/D18-1367>.