

# Faire figure d'autorité : l'analyse de réseaux appliquée au discours

## Authority Figures: Social Network Analysis applied to Literature

Marine Riguet<sup>1</sup>, Alaa Abi-Haidar<sup>2</sup>

<sup>1</sup> Labex OBVIL, Paris IV-Sorbonne, marineriguet@gmail.com

<sup>2</sup> LIP6, Université Pierre et Marie Curie, alaa.abi\_haidar@upmc.fr

**RÉSUMÉ.** Cet article s'attache à identifier les figures scientifiques qui font autorité dans la critique littéraire française de la seconde moitié du XIX<sup>e</sup> siècle. À l'aide d'une méthode automatisée de reconnaissance d'entités nommées et d'une analyse de réseaux empruntée aux sciences sociales, nous tenterons tout d'abord de déterminer la place et le rôle des savants cités par les critiques littéraires. Une analyse bibliométrique complétera notre étude de manière à pouvoir distinguer les simples allusions des références textuelles proprement dites. À travers ces différentes méthodes permises par les techniques numériques, il s'agit de proposer l'ébauche d'un réseau d'influences, au cœur du dialogue qui se noue alors entre discours littéraire et discours scientifique.

**ABSTRACT.** In this paper, we use an original method of unsupervised named entity recognition to automatically extract names of authors (critics, scientists, writers...) from hundreds of manuscripts from the second half of the 19th century. We construct citation networks that allow us to study influential figures, and to identify main actors in this domain. Moreover, we rank them to analyse the prestige and discuss the emergence of some notable collaborations.

**MOTS-CLÉS.** Reconnaissance d'entités nommées, analyse de réseaux, littérature du XIX<sup>e</sup> siècle, interdisciplinarité.

**KEYWORDS.** Digital humanities, named entity recognition, network analysis, interdisciplinary studies.

### 1. Introduction

La critique littéraire est par définition un discours de l'« autre ». Discours de la littérature, avant tout, métadiscours qui la place dans les marges des œuvres en qualité de commentaire critique, elle occupe ce statut ambigu pour lequel il s'est souvent vu relégué au rang de littérature « secondaire », voire de « genre mineur ». Critiquer, c'est au premier chef parler de ce qui parle ailleurs, et de la façon dont ça parle. Mais la critique littéraire est aussi le foyer d'échanges interdiscursifs, ce que favorise précisément son positionnement en marge. Elle se fait par excellence écriture « palimpseste » [GEN 82], superposant les marques d'autres discours et, selon la terminologie de Genette, d'une « transtextualité » comprise comme « tout ce qui l[a] met en relation, manifeste ou secrète, avec d'autres textes » [Ibid., 7]. En ce sens, elle a la particularité de se tisser sous la forme d'un vaste réseau reliant sujets et objets, « citants » et « cités », c'est-à-dire auteurs critiques, auteurs critiqués et figures d'autorité.

Nous situons l'émergence d'une critique littéraire comme discours autonome et légitimé dans la seconde moitié d'un XIX<sup>e</sup> siècle, au temps d'un positivisme triomphant. Thibaudet le remarque au commencement du siècle suivant,

la critique telle que nous la connaissons et la pratiquons est un produit du XIX<sup>e</sup> siècle. Avant le XIX<sup>e</sup>, il y a des critiques. Bayle, Fréron et Voltaire, Chapelain et d'Aubignac, Denys d'Halicarnasse et Quintilien sont des critiques. Mais il n'y a pas la critique. [THI 30, 7]

À partir de Sainte-Beuve surtout, qui appuie le premier la nécessité d'une critique littéraire « aussi exacte que possible », ayant non plus la prétention de juger selon son goût mais la tâche « d'observer sans relâche, d'étudier et de pénétrer les conditions des œuvres diversement remarquables et l'infinie variété des formes de talent » [S-B 67, 87], une ligne se dessine. Cherchant les fondements de sa légitimité en même temps qu'elle affirme son programme, la critique s'enracine dans l'épistémè du siècle par le recours à des instances qui lui sont extérieures et qui bénéficient déjà d'une

institutionnalisation, à l'instar de la philosophie ou des sciences exactes. En outre, son rôle à la fois pédagogique, culturel, politique et social, la situe par définition au carrefour des influences et des circulations interdisciplinaires, à l'articulation de discours exogènes dont elle se fait à son tour le vecteur ou la vulgate, et qu'elle active en son sein comme « champs de concomitances », selon la terminologie de Foucault :

il s'agit alors des énoncés qui concernent de tout autres domaines d'objets et qui appartiennent à des types de discours tout à fait différents ; mais qui prennent activité parmi les énoncés étudiés soit qu'ils servent de confirmation analogique, soit qu'ils servent de principe général et de prémisses acceptés pour un raisonnement, soit qu'ils servent de modèles qu'on peut transférer à d'autres contenus, soit qu'ils fonctionnent comme instance supérieure à laquelle il faut confronter et soumettre au moins certaines de ces propositions qu'on affirme.[FOU 69, 81]

Nous nous intéresserons ici à la façon dont la critique forme son discours à travers un réseau d'influences scientifiques, et de quelle manière elle réussit à se positionner par rapport aux références qu'elle convoque. À l'aide d'une méthode automatisée de reconnaissance d'entités nommées et d'une analyse bibliométrique, nous nous attacherons plus particulièrement à identifier les savants qui font autorité dans la critique littéraire française de la seconde moitié du XIX<sup>e</sup> siècle, de sorte de déterminer leur place au sein du discours, le poids qu'ils occupent, et l'interaction qu'ils motivent.

## 2. Extraction automatique des entités nommées<sup>1</sup>

Pôle important de la recherche en Humanités numérique, la reconnaissance d'entités nommées (REN) consiste à repérer informatiquement des objets textuels catégorisables tels que des noms de lieux, de personnes, d'organisations, des valeurs numériques, des titres, des dates, etc. À ce titre, elle offre des perspectives intéressantes pour identifier les autorités citées par la critique littéraire française de la seconde moitié du XIX<sup>e</sup> siècle. Notre intérêt s'est donc porté sur une méthode automatique permettant d'extraire l'ensemble des noms de personnes apparaissant dans notre corpus d'étude, sans qu'il soit besoin d'étiqueter chacune de ses entités à la main. UNERD<sup>2</sup> (*Unsupervised Named Entity Recognition and Disambiguation*), méthode non-supervisée de reconnaissance et de désambiguïsation d'entités nommées, sert bien ce dessein ; fonctionnant sans apprentissage préalable, elle s'appuie sur un dictionnaire afin de détecter les mentions certaines, et de désambiguïser les autres selon leur contexte d'emploi. Lors d'études préliminaires, cette méthode a déjà fait ses preuves en livrant des résultats avec une marge de bruit et d'erreurs moindre – sa performance, calculée par F-score<sup>3</sup>, a ainsi été estimée 20 % supérieure à la méthode DBpedia Spotlight<sup>4</sup> et 10 % supérieure à celle de BaLIE<sup>5</sup>.

Pour réaliser cette étude, un corpus numérisé de 242 ouvrages de critique littéraire française publiés entre 1850 et 1910 (environ 20 millions de mots) a été préalablement constitué à partir des bases déjà existantes d'ARTFL<sup>6</sup> (Université de Chicago) et du Labex OBVIL<sup>7</sup> (Paris IV-Sorbonne). Cette étape

<sup>1</sup> Nous remercions tout particulièrement Carmen Brando et Francesca Frontini pour leur aide précieuse dans cette étape du travail.

<sup>2</sup> Mosallem, Y. Abi-Haidar, A. and Ganascia J.-G., « Unsupervised Named Entity Recognition and Disambiguation: An Application to Old French Journals », in *Lecture Notes in Computer Science series*, vol. 8557, Proceedings of ICDM 2014, St. Petersburg, Russia, 2014 ; et Abi-Haidar A., Wild, O., and Ganascia J.G., « A Simple yet Efficient Method for Unsupervised Named Entity Recognition and Disambiguation », *WIRES Data Mining and Knowledge Discovery (DMKD)*, 2016.

<sup>3</sup> F-score est une mesure de performance qui combine la précision et le rappel, en proposant leur moyenne harmonique.

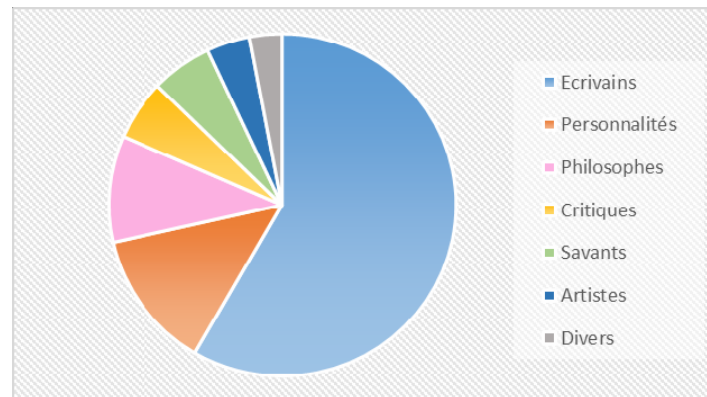
<sup>4</sup> Mendes, Pablo N., et al., « DBpedia spotlight: shedding light on the web of documents », *Proceedings of the 7th International Conference on Semantic Systems*, ACM, 2011.

<sup>5</sup> Nadeau, D., Turney, P., & Matwin, S., *Unsupervised named-entity recognition: Generating gazetteers and resolving ambiguity*, 2006.

<sup>6</sup> Lancé en 1982, le projet ARTFL de l'Université de Chicago propose une large base de données textuelle de littérature et de critique françaises. URL : <http://artfl-project.uchicago.edu/>

<sup>7</sup> L'Observatoire de la vie littéraire (OBVIL) est un laboratoire d'excellence de Paris IV-Sorbonne, créé en partenariat avec l'Université Pierre et Marie Curie, et dévolu à l'analyse des transformations des conditions de la création littéraire et de l'épistémologie critique

de numérisation ne se résume pas seulement à la compilation des textes bruts, en vue d'une exploitation dans leur format TXT, mais comprend un enrichissement et une structuration sémantique soignée grâce à l'encodage XML-TEI. L'extraction des noms de personnes du corpus par UNERD a permis d'obtenir la liste de plus d'un millier de mentions, qu'il a ensuite fallu trier manuellement ; dans le but d'affiner la pertinence des résultats et leur analyse, nous avons pris le parti de conserver exclusivement ceux qui revenaient plus de vingt fois dans l'ensemble du corpus. Les 594 noms finalement retenus révèlent la grande hétérogénéité des personnes réelles<sup>8</sup> citées par la critique, et que l'on peut distinguer en trois types : les écrivains anciens et contemporains qui sont objets du discours ; les autorités convoquées, c'est-à-dire les auteurs, philosophes et savants qui prennent part à la stratégie argumentative du critique ; et les figures du monde référent, les personnalités qui se rapportent à la réalité contemporaine ou historique. Leur catégorisation par domaine (écrivains, personnalités politiques ou historiques, philosophes, critiques, savants, artistes, et autres) rend leur répartition visible selon la figure suivante.



**Figure 1.** Proportion des personnes citées par domaine, selon leur diversité

Si, on le constate sans surprise, la part des écrivains cités est la plus importante, confirmant que la critique parle bien, et avant toute chose, de littérature, il n'en reste pas moins que la proportion des savants en présence, égale à celle des critiques, n'est pas négligeable. Plus précisément, 95 hommes de science (compris au sens large, sciences exactes, sociales et humaines) figurent parmi les noms convoqués au sein du corpus. Pour plus de précision, nous avons procédé à une seconde extraction de cette liste restreinte : les titres de civilité ont été enlevés, et les différentes orthographes d'un même nom, causées par les particules, les tirets, les accents ou les  $\alpha$ , ont été fusionnées entre elles. Les résultats permettent ainsi de cibler spécifiquement les références aux savants. Globalement, ceux-ci laissent apparaître quatre grands niveaux d'influence : les représentants de la science grecque antique (Pythagore, Thalès, Ptolémée, Euclide, Aristote) ; ceux qui ont, les premiers, affirmé une philosophie positive face à une lecture métaphysique du monde et posé les fondations de la science rationnelle (Bacon, Newton, Galilée, Copernic, Kepler) ; les pionniers de l'histoire naturelle et de la classification des connaissances à l'âge classique (Linné, Buffon, Jussieu) ; enfin, les maîtres de la science moderne, qui s'attachent désormais à étudier les organismes et à ordonner le monde à partir de la notion de vie (Lamarck, Cuvier, Berthelot, Geoffroy Saint-Hilaire, Magendie, Bichat, Bernard, Pasteur, Darwin, Spencer, Hæckel, Charcot, Flourens, etc.).

### 3. Pour une esquisse des réseaux d'influence

Suite à ces résultats préliminaires, il reste encore à déterminer avec précision qui cite qui. Car les critiques ne donnent pas tous une importance égale aux références scientifiques, et les différentes dynamiques qui se dégagent du réseau sont d'autant moins généralisables à l'ensemble de la

par le biais d'outils informatiques. Une édition électronique savante du corpus de critique littéraire française est consultable en ligne, URL : <http://obvil.paris-sorbonne.fr/corpus/critique/>

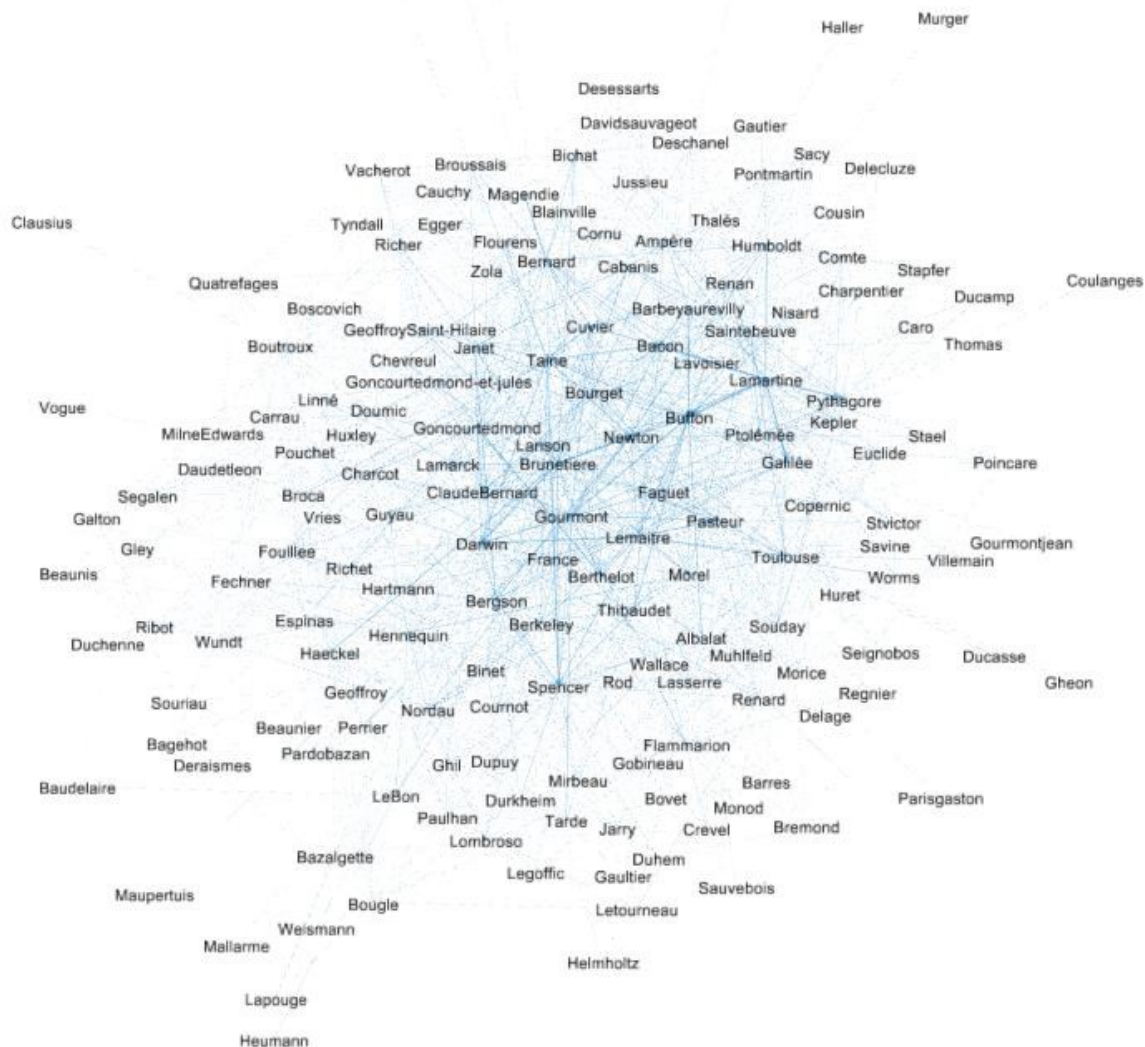
<sup>8</sup> Nous avons également choisi d'exclure ici les personnages fictifs.

production littéraire du XIX<sup>e</sup> siècle que le traitement dévolu aux savants est considérablement disparate. Rappelons-le, la critique de la seconde moitié du siècle débat ; et si les sciences exercent à l'époque une influence croissante, certains auteurs restent farouchement réfractaires à l'esprit scientifique du temps, à l'instar d'un Émile Montégut ou d'un Armand de Pontmartin. Difficile également de fixer globalement la place d'une figure exogène qui vient pour les uns servir d'argument d'autorité, pour d'autres de sujets d'étude, ou simplement de citation référentielle. Il nous faut donc situer la place que les savants occupent et la part d'influence à laquelle ils prétendent au sein du discours critique. Pour explorer plus avant l'agencement de ces citations nominales, nous éprouvons dès lors le besoin de cartographier nos résultats sur le modèle des réseaux sociaux. La construction automatique de réseaux, généralement pratiquée en sciences sociales, offre en effet la possibilité de représenter les liens entre plusieurs entités selon leur degré d'interaction et d'occurrence, de sorte de mettre à jour les dynamiques du réseau et de détecter ses principaux acteurs. Son usage vient surtout servir les cartographies du web, en retraçant les liens hypertextes, et l'analyse des réseaux sociaux (*social network analysis*) ; mais à ce titre, elle semble aussi ouvrir des pistes intéressantes pour saisir la manière dont un discours s'appareille à un autre, dont les diverses références textuelles sont convoquées, se connectent entre elles et interagissent avec les auteurs qui les citent<sup>9</sup>. Pour ce faire, nous avons donc exploité les résultats de notre seconde extraction via Gephi<sup>10</sup>, logiciel libre d'analyse et de visualisation de réseaux. Une première spatialisation (figure 2) rend visible l'organisation interne du réseau de références scientifiques au sein du corpus : les nœuds, à savoir les noms des critiques citant et des savants cités, sont répartis selon la densité de leurs liens.

---

<sup>9</sup> Voir notamment Liu, X., Bollen, J., Nelson, M. L., & Van de Sompel, H, « Co-authorship networks in the digital library research community », in *Information processing & management*, 41(6), 2005, p. 1462-1480 ; et Abi Haidar, A., Six, A., Ganascia, J. G., & Thomas-Vaslin, V., « The Artificial Immune Systems Domain: Identifying Progress and Main Contributors Using Publication and Co-Authorship Analyses », in *Advances in Artificial Life*, ECAL, vol. 12, 2013, p. 1206-1217.

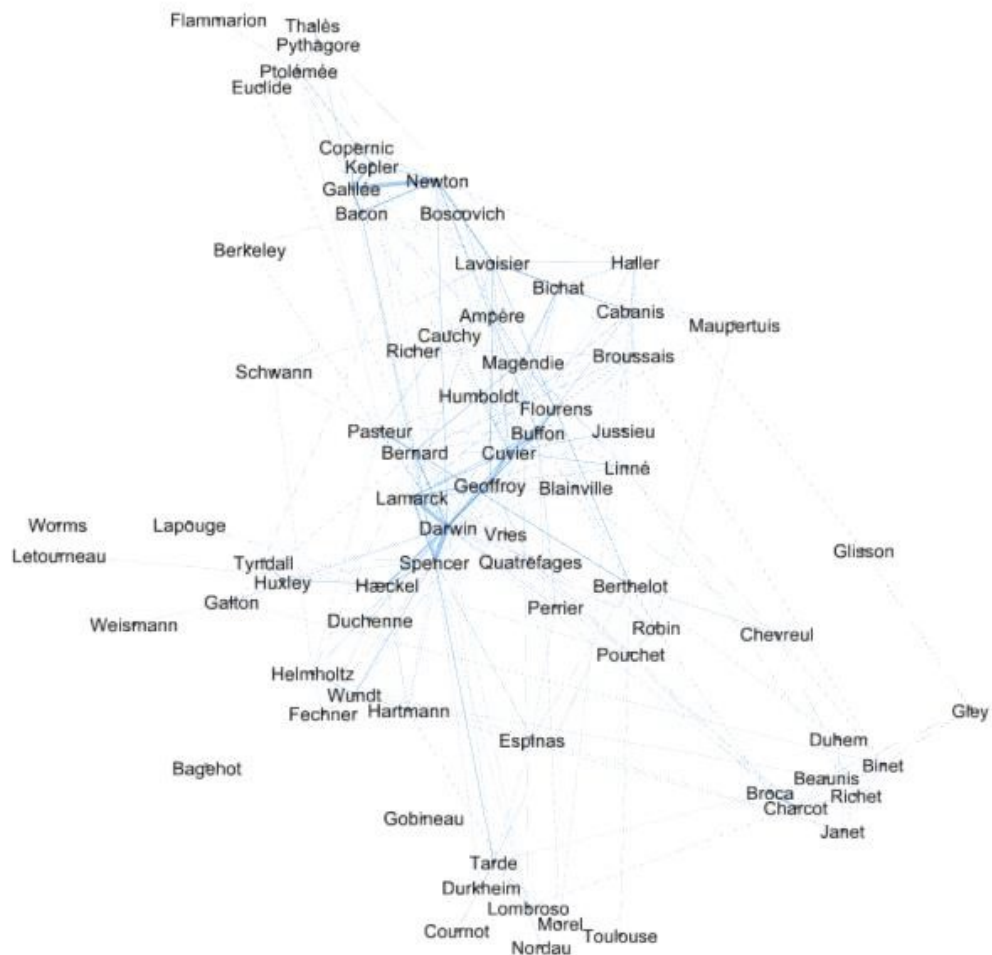
<sup>10</sup> Mathieu Bastian, « Gephi : An Open Source Software for Exploring and Manipulating Networks », in *AAAI Publications, Third International AAAI Conference on Weblogs and Social Media*, 2009. La première version de Gephi date de 2008, et sa version la plus récente est téléchargeable en ligne. URL : <https://launchpad.net/gephi/0.6/0.6alpha1>



**Figure 2.** Réseau global des références scientifiques (graphe dirigé de type « Force Atlas » par Gephi)

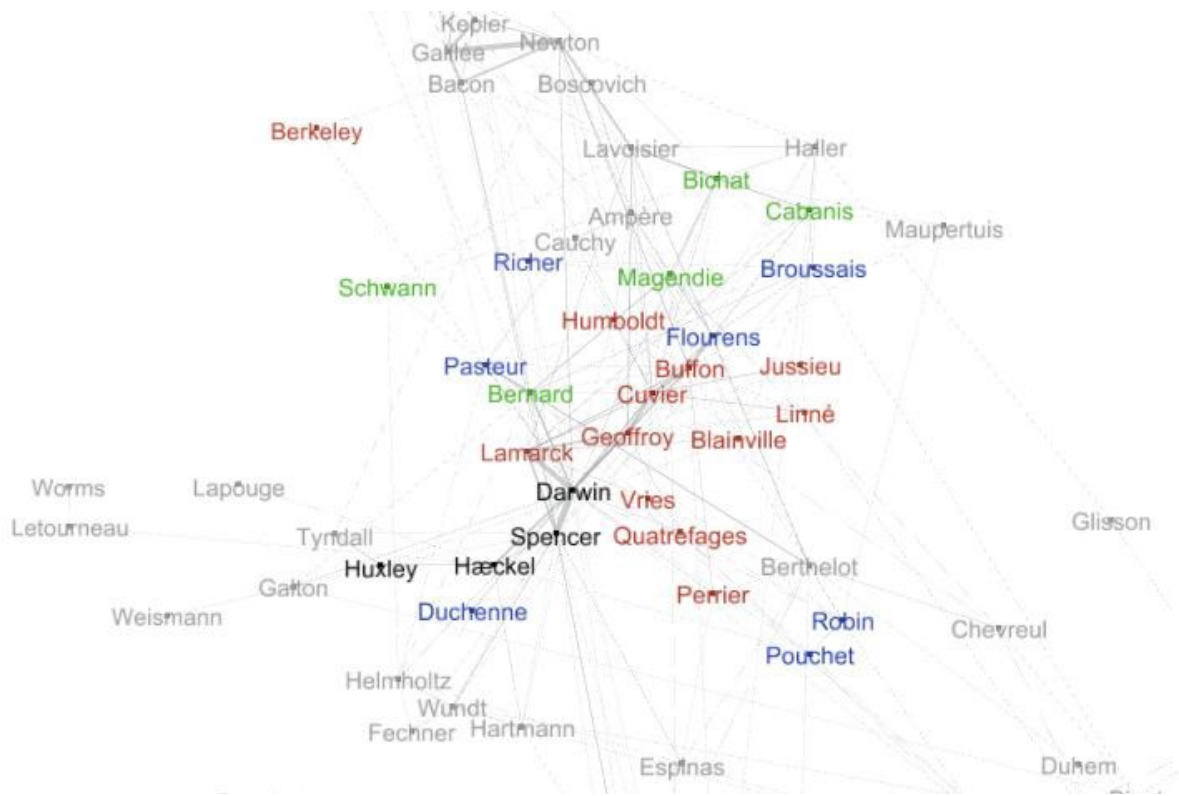
Autrement dit, la disposition des noms et l'épaisseur des liens font apparaître la place des différents acteurs, et notamment, au centre du réseau, celle des savants les plus influents et des critiques qui s'y réfèrent abondamment. Le graphe est qui plus est dirigé, c'est-à-dire qu'il garde le sens des liens entre deux nœuds : si Lamartine et Buffon sont connectés, c'est bien par un lien allant de Lamartine à Buffon et non l'inverse. S'observe ainsi une forte interaction entre des auteurs comme Hippolyte Taine, Jules Lemaitre, Anatole France, Ferdinand Brunetière, Gustave Lanson, les frères Goncourt, etc. et certains hommes de science particulièrement convoqués : Charles Darwin, Georges Buffon, Herbert Spencer, Isaac Newton, Georges Cuvier, Louis Pasteur, Marcellin Berthelot, Étienne Geoffroy Saint-Hilaire...

En outre, une spatialisation des savants co-occurents (figure 3), c'est-à-dire des savants cités dans une même phrase par un critique, clarifie cette répartition des références scientifiques autour d'un noyau d'influences plus appuyé.



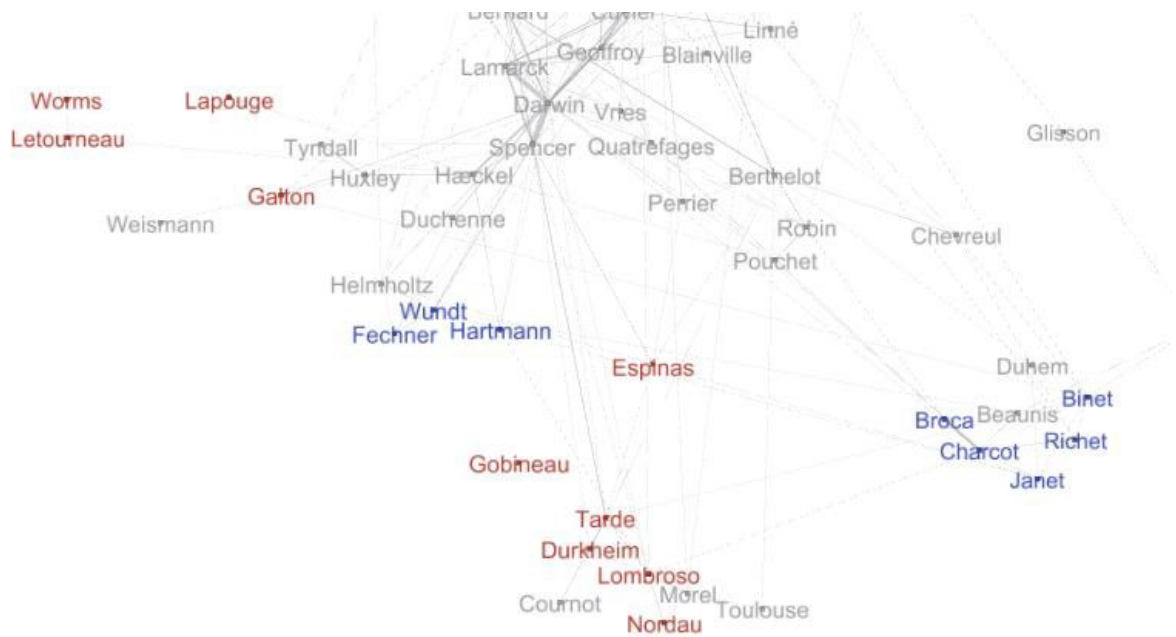
**Figure 3.** Graphe non-dirigé de savants alignés selon leurs co-occurrences (Gephi)

Le graphe est non-dirigé dans la mesure où les nœuds représentent exclusivement les savants cités ; si le sens des liens n'a donc ici aucune signification particulière, la position des nœuds, en revanche, traduit par alignement leurs situations de co-occurrences. Ce réseau met au jour un foyer central particulièrement dense (explicité par couleurs dans la figure 4 ci-dessous) : une constellation évolutionniste – Darwin, Spencer, Hæckel et Huxley (en noir) – relie les principaux physiologistes – Cabanis, Schwann, Magendie, Bichat, Claude Bernard (en vert) –, les naturalistes – Geoffroy Saint-Hilaire, Quatrefages, Humboldt, de Vries, Lamarck, Cuvier, Blainville, Buffon, Jussieu, Linné, Perrier, Berkeley (en rouge) –, et les médecins – Duchenne, Broussais, Pasteur, Pouchet, Robin, Flourens, Richer (en bleu). Un tel agglomérat témoigne d'une surreprésentation des sciences de la vie, parallèle à l'essor que celles-ci connaissent au cours du siècle.



**Figure 4.** Zoom sur les co-occurents au centre du réseau, colorés par domaine

Surtout, cette interconnexion référentielle dessine la façon dont la critique axe son discours autour d'un changement épistémologique : parler de science, c'est avant tout faire mention d'un « nouveau régime de vérité » [SEG 11], d'une nouvelle représentation du monde vivant et de l'homme. En plus de ce foyer central, le réseau de co-occurents rend visible d'autres constellations de références : d'une part, autour d'une étude interne de l'homme, avec l'école de psychologie allemande – Wundt, Fechner, Hartmann – et la neuro-psychologie – Broca, Charcot, Richet, Binet, Janet – (en bleu sur la ci-dessous) ; et, d'autre part, autour d'une étude générale de la nature humaine, avec la sociologie, la criminologie et l'anthropologie naissantes – Lombroso, Tarde, Espinas, Durkheim, Nordau, Galton, Worms, Vacher de Lapouge, Letourneau (en rouge sur la figure ci-dessous).

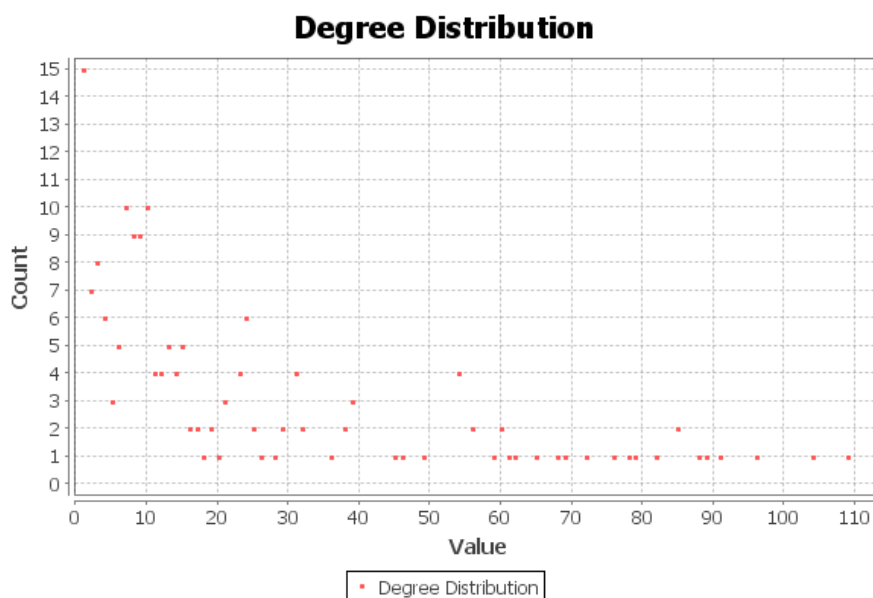


**Figure 4.** Zoom sur les constellations de co-occurents, colorés par domaine

Quoique de manière plus sporadique, on note aussi des mentions à la physique (Duhem, Lavoisier, Ampère, Tyndall, Helmholtz), à la chimie (Berthelot, Chevreul, Haller), aux mathématiques (Cournot, Cauchy), ou à la cosmologie (Maupertuis). Par ces citations multiples, la critique se place comme point de confluence interdisciplinaire, cœur d'une nouvelle dialectique où le vivant s'érige en « domaine de référence » [*Ibid.*] et où l'homme moderne, placé au centre, reconquiert sa place.

#### 4. Le savant, figure consensuelle ?

Dans la mesure où elle appartient à un domaine exogène, la figure du savant obéit à une construction particulière : la critique littéraire ne se contente pas de la convoquer, elle la met en scène comme représentation des idées du siècle, grande instance d'un discours érudit et institutionnalisé ; en somme, elle ne cite ni n'importe qui, ni n'importe comment. La distribution des références aux savants dans le corpus est, à ce titre, éloquente. Le graphe suivant (figure 5) représente la fréquence de distribution des degrés pondérés : autrement dit, il figure chaque nœud selon le nombre de liens entrants (X cité par Y) et sortant (X citant Y) qui le connecte aux autres nœuds du réseau. De cette façon, il rend visible la répartition des savants cités en fonction du nombre de critiques qui les citent, et révèle ici une distribution plutôt homogène des références scientifiques au sein du discours critique.



**Figure 5.** Distribution des degrés pondérés pour les savants cités (Gephi)

La partie gauche du graphe, en effet, montre que la majorité des nœuds se concentre sur des degrés assez faibles (d'une valeur comprise entre 0 et 20). Cette concentration des degrés témoigne qu'il est aussi peu de critiques citant une grande variété de savants, que de savants cités par un grand nombre d'auteurs. Quelques nœuds, beaucoup moins nombreux, ont en revanche une valeur particulièrement forte (entre 50 et 114), et illustrent la façon dont certains savants se distinguent par le nombre de références dont ils font l'objet et la diversité des critiques qui les citent. Il convient dès lors de s'interroger sur l'émergence de figures d'autorité scientifiques qui semblent s'imposer à travers une construction collective, si ce n'est consensuelle, au sein de la critique littéraire.

En analyse de réseaux sociaux, la notion de centralité est une caractéristique structurelle servant à déterminer les positions des nœuds les uns par rapport aux autres ; à cet égard, les entités identifiées comme « centrales » semblent jouer un rôle d'importance, ou exercer une influence plus forte sur le reste du réseau<sup>11</sup>. Différents calculs s'attachent à mesurer les indicateurs de centralité. Dans un article

<sup>11</sup> Voir notamment Stanley Wasserman and Katherine Faust, *Social network analysis: Methods and applications*, vol. 8, Cambridge university press, 1994.



qui fait désormais autorité [FRE 79], Linton C. Freeman en établit trois types que nous choisissons de reprendre à notre tour, afin d'identifier les autorités les plus actives au sein du réseau d'influence scientifique qui nous occupe.

#### 4.1. La centralité de degré

La centralité de degré (degree centrality) se calcule assez simplement sur le nombre de liens connectant un nœud à ceux qui l'entourent. Selon Freeman, cet indicateur est particulièrement significatif pour mesurer l'implication d'un acteur au sein d'un réseau donné. À partir du réseau de co-occurrences constitué précédemment, il nous fournit donc un classement des savants les plus co-occurents par les critiques du corpus.

Nom	Bernard	Saint-Hilaire	Newton	Buffon	Spencer	Darwin	Bacon	Cuvier	Pythagore	Galilée
Mesure	238	237	168	167	167	166	164	164	164	163

Nom	Berthelot	Lamarck	Le Bon	Ptolémée	Espinas	Janet	Kepler	Pasteur	Broca	Charcot
Mesure	162	162	162	162	161	160	157	157	156	156

**Tableau 1.** Tableau des 20 savants ayant la centralité de degré la plus forte

Avec une mesure de centralité de 238, Claude Bernard arrive en tête, central pour être à la fois une figure d'autorité en médecine et en physiologie ; mais sa position privilégiée au cœur du réseau de co-occurrences témoigne surtout d'un élargissement disciplinaire : la critique ne se contente plus de le convoquer dans le cadre strict de son champ d'expertise, elle le hisse plus généralement comme figure de savant, et l'affilie aussi bien à Pasteur qu'à Cuvier, Darwin ou Buffon, en tant qu'homme de science et de progrès. Ce phénomène de généralisation s'applique d'une manière similaire aux premiers savants du tableau, tels Geoffroy Saint-Hilaire (237), Newton (168), Buffon (167), Spencer (167) ou Darwin (166). On retrouve ainsi, presque sans surprise, les grands noms de l'histoire de sciences, ceux qui ont déjà été intronisés aux siècles précédents (Newton, Buffon, Bacon, Galilée, Pythagore, Ptolémée, Kepler...), et ceux qui animent l'actualité scientifique et intellectuelle de l'époque. En outre, il est intéressant de noter une relative uniformité des indices de degré, variant de 238 à 156 pour les vingt premiers résultats, qui trahit la façon souvent énumérative dont les critiques intègrent les savants à leur discours. Citons, parmi tant d'autres exemples, Paul Stapfer :

Des hommes tels que Buffon, Lamarck, Cuvier, Geoffroy Saint-Hilaire, Ampère, Claude Bernard, Darwin, Pasteur, d'une part ; tels, d'autre part, que Bopp, Niebuhr, Burnouf, Littré, Spencer, Max Muller, Renan, nous apparaissent aujourd'hui, non comme des génies isolés et capricieux, mais comme les coopérateurs d'une immense œuvre commune qui achemine l'homme à la conquête des choses et de la vérité...[STA 93]

Non seulement les savants co-occurents sont extraits de leurs domaines originels, mais ils sont dotés d'un pouvoir référentiel suffisamment fort pour se passer de toute contextualisation ; d'ailleurs, il est ici moins question de science au sens théorique, que de culture. Les savants, surtout les plus « centraux », ne figurent pas dans le discours critique comme hommes de science mais de savoir. Ils acquièrent une posture éminente parce qu'ils guident les peuples vers la connaissance et réalisent l'homme, si ce n'est « comme maître et possesseur de la nature » [DES 37], du moins dans la pleine conscience de ce qui l'entoure. Ils travaillent en somme avec les philosophes, qu'ils peuvent même être

amenés à orienter, dans le but d'éclairer les esprits ; et c'est à ce titre qu'ils procurent au critique littéraire un modèle d'autorité. Ce sont eux, ces « vrais savants » auxquels se réfère Brunetière :

[...] ceux dont les immortelles découvertes ont balancé ou compensé la stérile abondance de la littérature impériale et révolutionnaire, Laplace et Monge, Berthollet et Fourcroy, Chaptal, Cuvier, Lamarck, Geoffroy Saint-Hilaire...[BRU 98, 399]

#### 4.2. La centralité d'intermédiarité

Autre mesure, la centralité d'intermédiarité (*betweenness centrality*) se réfère à la position d'un nœud sur des « lieux de passages » centraux du réseau, c'est-à-dire sur les chemins les plus courts reliant tous les nœuds entre eux. Pour plus de clarté, elle peut être schématisée comme suit :



**Figure 6.** Schéma de nœuds à forte centralité d'intermédiarité dans un réseau<sup>12</sup>

En d'autres termes, ce calcul met à jour les autorités intermédiaires servant à relier les foyers ou communautés disparates ; elles ne sont pas forcément celles qui pèsent le plus ou qui sont le plus citées, mais elles tiennent en revanche une place centrale dans l'interconnexion du réseau. Pour nos savants co-occurents au sein du corpus critique, nous obtenons les résultats suivants :

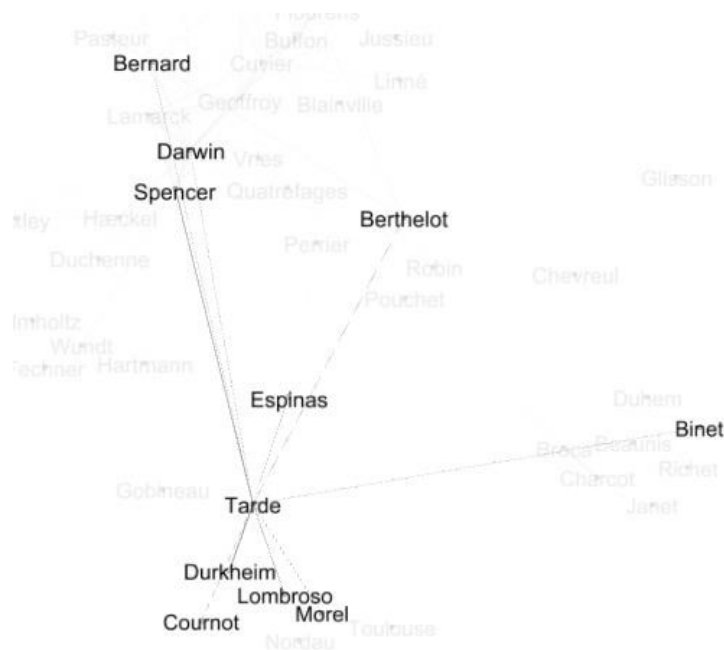
Nom	Buffon	Spencer	Tarde	Bernard	Newton	Hartmann	Espinas	Bacon	Darwin	Saint-Hilaire
Mesure	34.943	31.792	24.743	23.273	23.079	19.565	18.712	18.416	18.135	17.682

Nom	Pythagore	Ptolémée	Morel	Cuvier	Broca	Berkeley	Janet	Le Bon	Galilée	Perrier
Mesure	16.382	15.677	14.532	13.789	13.629	13.545	13.038	12.700	12.582	11.621

**Tableau 2.** Tableau des 20 savants ayant la centralité d'intermédiarité la plus forte

Ce classement est intéressant dans la mesure où il révèle des figures scientifiques qui n'apparaissent pas dans les résultats précédents, et qui ne sont pas prédominantes en termes d'occurrences. Gabriel Tarde, par exemple, est à l'évidence moins fréquemment convoqué par la critique littéraire que Darwin ; pour autant, il joue un rôle d'intermédiaire plus fort entre les différents agglomérats de co-occurrences (avec une centralité d'intermédiarité de 24,743, il est le troisième savant à la position d'intermédiarité la plus centrale).

<sup>12</sup> URL de référence : <http://master-iesc-angers.com/utilisation-du-logiciel-gephi-pour-lanalyse-cartographique>



**Figure 7.** Tarde et ses co-occurents

Nouveau « maître » de la « psychologie collective » [DOU 03] d'après Doumic, ses travaux sur la criminalité et le pouvoir des foules le lient aux psychologues, anthropologues et criminologues contemporains tels Ribot et Le Bon, avec qui il entretient des relations amicales, ou Lombroso, dont il conteste les théories. On l'attache également à Spencer et à Durkheim pour ses analyses sociologiques, dont ses fameux principes d'invention et d'imitation. « Philosophe ingénieux et amer », enfin, il est celui qui définit la vie comme « la poursuite de l'impossible à travers l'inutile » [GOU 25], et qui participe en ce sens à une réflexion du siècle amorcée par les évolutionnistes. Il tient pour ainsi dire un rôle pivot au cœur du réseau, en tant qu'instance catalysant une circulation d'idées particulièrement remarquable hors des frontières disciplinaires.

Un constat similaire peut être fait pour Eduard von Hartmann, qui a la sixième position d'intermédiarité la plus centrale (avec une mesure de 19.565). Connus surtout pour sa *Philosophie de l'inconscient*, publiée en 1868 et traduite en français en 1877, il popularise l'idée d'un élan psychique proche de l'instinct, dissocié des fonctions intellectuelles conscientes. De ce fait, affilié à Schopenhauer et Hegel sur le plan philosophique, il assure surtout le lien entre les psychologues qui s'intéressent aux mouvements de la conscience d'une part, et les physiologistes qui traitent des besoins organiques d'autre part ; sans oublier que sa nationalité allemande le rattache *de facto* à ses confrères de l'époque, comme illustration de l'influence intellectuelle germanique. Aussi lit-on chez Lanson :

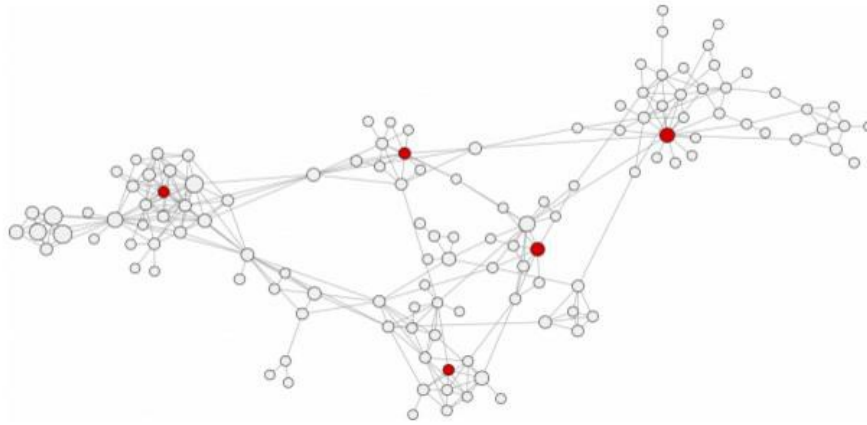
De l'Allemagne, nous avons connu surtout, de première ou de seconde main, le matérialisme scientifique de Büchner, l'évolutionnisme systématique de Hæckel ; le pessimisme de Schopenhauer nous a conquis ; et M. de Hartmann a mis pour un temps l'inconscient à la mode. [LAN 95]

Figures mobiles, ces savants révélés par la centralité d'intermédiarité assurent, au sein du discours critique, une fonction de mise en circulation interdisciplinaire, voire transdisciplinaire, du savoir.

#### 4.3. La centralité de proximité

Enfin, troisième type de mesure selon Freeman, la centralité de proximité (*closeness centrality*) traduit la distance moyenne entre une entité et le reste du réseau, en calculant la longueur moyenne des chemins les plus courts reliant un nœud donné à tous les autres nœuds. Elle sert à indiquer le degré

d'indépendance [BRASS 92], de proximité ou d'éloignement, d'une autorité avec son voisinage. Les nœuds ciblés peuvent être schématisés de la façon suivante :



**Figure 8.** Schéma des nœuds à forte centralité de proximité dans un réseau<sup>13</sup>

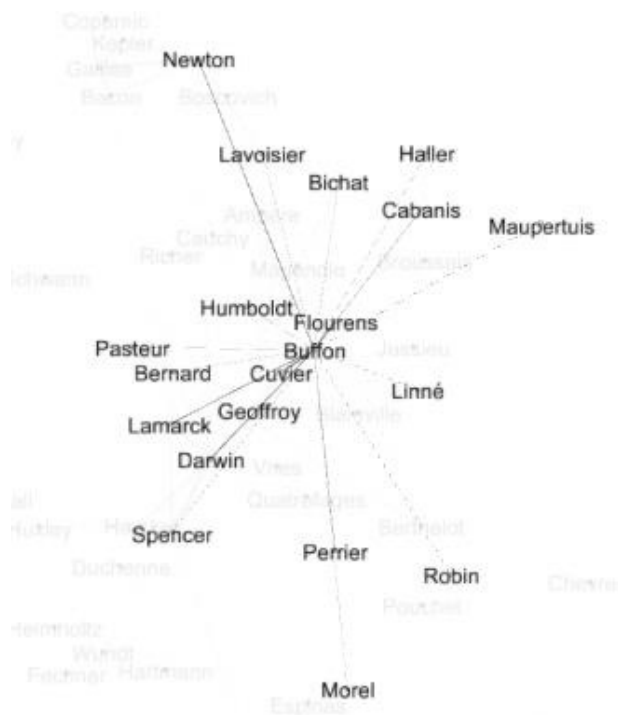
Cette mesure de centralité est peut-être pour nous la plus pertinente ; elle fait apparaître les savants qui, à proprement parler, font autorité dans l'ensemble du réseau.

Nom	Buffon	Newton	Bacon	Darwin	Spencer	Bernard	Cuvier	Galilée	Saint- Hilaire	Pythagore
Mesure	0.0120	0.0120	0.0119	0.0119	0.0119	0.0117	0.0117	0.0117	0.0117	0.0117
Nom	Espinas	Lamarck	Ptolémée	Berthelot	Janet	Kepler	Tarde	Berkeley	Pasteur	Perrier
Mesure	0.0116	0.0116	0.0116	0.0115	0.0115	0.0115	0.0115	0.0114	0.0114	0.0114

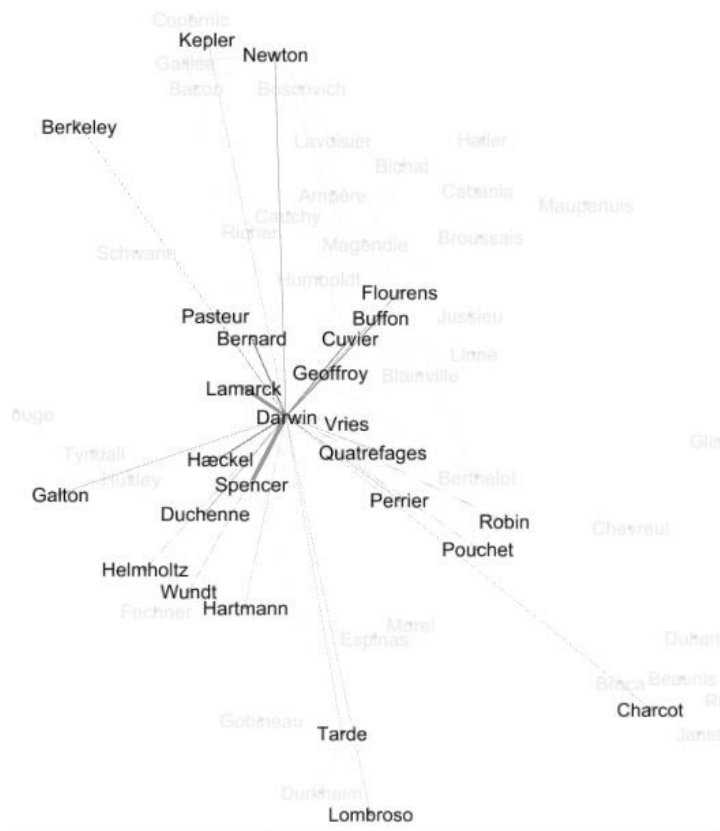
**Tableau 3.** Tableau des 20 savants ayant la centralité de proximité la plus forte

De façon cohérente, les résultats sont ici assez proches de ceux calculés par la centralité de degré. Les trois premiers nœuds à la centralité de proximité la plus forte, à savoir Buffon pour son modèle de classification des espèces, Newton pour son système universel, et Bacon pour sa méthode empirique, s'imposent uniformément comme grandes instances de la science rationnelle ; à leurs côtés, on peut également citer les astronomes Kepler et Galilée.

<sup>13</sup> URL de référence : <http://master-iesc-angers.com/utilisation-du-logiciel-gephi-pour-lanalyse-cartographique/>



**Figure 9.** Buffon et ses co-occurents



**Figure 10.** Darwin et ses co-occurents

À partir de ces grands noms s'établit l'histoire d'un « âge scientifique », un changement de paradigme ouvrant à la connaissance moderne du monde, ainsi que l'avance Renan dans son discours de réception à l'Académie française de 1879 :

Qui ne voit que Galilée, Descartes, Newton, Lavoisier, Laplace ont changé la base de la pensée humaine, en modifiant totalement l'idée de l'univers et de ses lois, en substituant aux

enfantines imaginations des âges non scientifiques la notion d'un ordre éternel, où le caprice, la volonté particulière, n'ont plus de part ? [REN 87]

Parmi les savants du XIX<sup>e</sup> siècle, Darwin et Spencer incarnent, de la biologie à la philosophie, la révolution évolutionniste qui estampe le siècle ; Claude Bernard et Pasteur, l'intronisation de la médecine au rang de science et l'expérimentation en laboratoire ; Cuvier, Lamarck, Geoffroy Saint-Hilaire, Berthelot et Perrier, la nouvelle représentation du monde vivant à laquelle s'attache la biologie naissante. En plus de ces références aux sciences exactes, la critique littéraire convoque certaines figures des sciences humaines et sociales qui commencent d'émerger, à l'instar d'Espinas et de Tarde en sociologie, ou de Janet en psychologie. Nombre d'autres noms, on l'a vu, circulent ainsi dans le discours critique. Mais il est intéressant de noter la façon dont se dessine d'ores et déjà parmi ces diverses mentions des figures prédominantes, affranchies de leur cadre originel et mises au service des idées du siècle, ou dont se discute à l'inverse la toute récente postérité.

## 5. Étude bibliométrique

Derrière ce réseau de références nominales, qui ne fait qu'ébaucher une première distribution des acteurs scientifiques dans la critique littéraire, les références bibliographiques offrent une deuxième couche d'analyse. En effet, les titres d'ouvrages scientifiques cités par la critique révèlent de façon plus précise l'emploi de ces références exogènes, la part des textes les plus réceptionnés, des théories influentes, et le rôle qu'ils tiennent pour l'un ou l'autre critique – un même savant pouvant venir appuyer des idées très distinctes selon l'ouvrage au travers duquel il est mentionné. La méthode bibliométrique nous est en ceci utile qu'elle sert au dénombrement des références bibliographiques au sein de corpus. Nous ne ferons pas preuve ici d'innovation. Consacrée assez généralement au référencement des travaux scientifiques, la bibliométrie date du début du XX<sup>e</sup> siècle ; la toute première étude, attribuée communément à Cole et Eales en 1917, s'attache alors à répertorier l'ensemble de la littérature traitant d'anatomie et publiée dans les années 1850-1860 [COL 17]. Cinquante ans plus tard, la bibliométrie prend un nouvel essor sous l'impulsion de sociologues américains, qui en font une méthode de prédilection afin d'évaluer l'activité de la recherche. En tête, les travaux d'Eugene Garfield [GA 79 ; GA 83] développent la statistique bibliométrique comme indice de la productivité et de la position stratégique des différents acteurs scientifiques. L. Miles Raising, en 1962, définit la bibliométrie en ces termes :

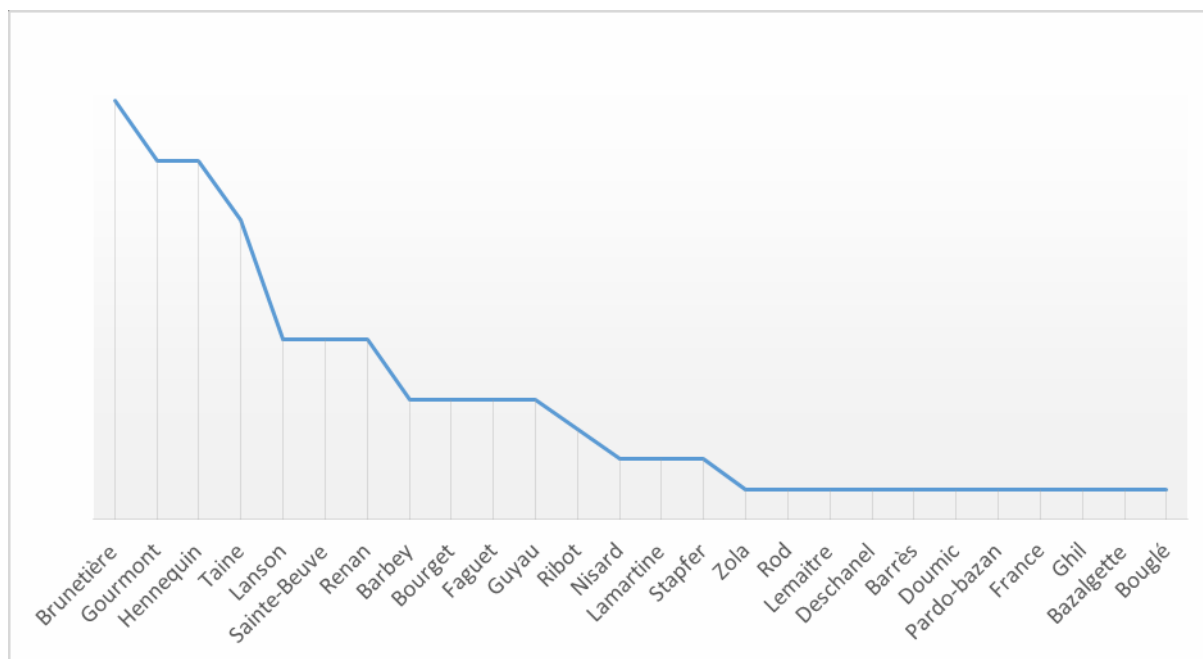
l'assemblage et l'interprétation de statistiques relatives aux livres et aux périodiques [...] pour démontrer des mouvements historiques, pour déterminer l'utilisation par la recherche nationale et universelle des livres et des journaux, et pour s'assurer dans de nombreuses situations locales de l'utilisation générale des livres et des journaux. [RAI 62]

Ces dernières années, de nombreux travaux se sont interrogés sur l'importance de la citation d'un point de vue non plus seulement quantitatif mais qualitatif, et notamment sémantique<sup>14</sup>. Mais si sa pratique a été élargie ces dernières décennies aux sciences humaines et sociales – et son usage largement discuté –, peu se sont jusqu'à présent interrogés sur sa potentielle pertinence pour l'analyse littéraire.

Eu égard à notre travail préalable d'encodage, l'extraction des ouvrages cités au sein du corpus critique a pu être réalisée automatiquement, à partir des éléments annotés à l'aide de la balise XML-TEI <title>, puis les résultats triés manuellement. Contrairement à la part relativement importante de savants nommés par la critique, les titres scientifiques paraissent peu nombreux face à la grande majorité d'ouvrages de littérature et de philosophie. Darwin, Buffon, Spencer, Bacon et Claude

<sup>14</sup> Voir par exemple H. D. White, « Citation Analysis and Discourse Analysis Revisited », in *Applied Linguistics*, 2004, p. 89-116 ; S. Teufel, A. Siddharthan, D. Tidhar, *An annotation scheme for citation function*, in *Proceedings of Sigdial-06*, Sydney, Australia, 2006 ; ou encore B. Jörg, *Towards the Nature of Citations*, German Research Center for Artificial Intelligence, Language Technology Lab, Saarbrücken, Germany, 2008.

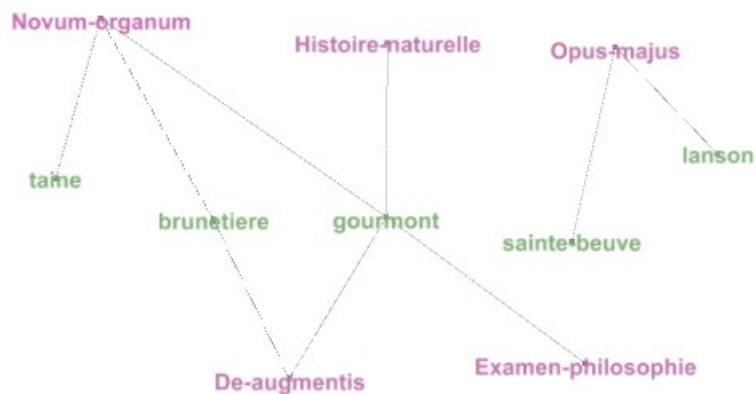
Bernard demeurent, dans l'ordre, les savants dont les titres sont les plus fréquemment cités ; d'autres, pourtant moins nommés par la critique littéraire, transparaissent à travers la popularité d'un de leurs essais, tel le *Cosmos* d'Humbolt, auquel se réfèrent abondamment Nisard, Brunetière, Lamartine et Renan, ou tel encore l'*Exposition du système du monde* de Laplace, que l'on retrouve notamment sous les plumes de Taine, Renan et Anatole France. Il est intéressant de noter la façon dont les quelques références bibliographiques dominantes s'imposent unanimement au point de faire consensus auprès de la critique : aussi émerge-t-il distinctement une liste de « canons » scientifiques, érigés par la critique sur le modèle des canons littéraires, c'est-à-dire une poignée d'ouvrages destinés à faire autorité dans le champ disciplinaire qui est le leur – du *Novum Organum* de Bacon à l'*Histoire de la création des êtres organisés d'après les lois naturelles* d'Haeckel, en passant par les *Principia Mathematica* de Newton, l'*Histoire naturelle* de Buffon, *De l'origine des espèces* de Darwin, les *Premiers principes* de Spencer ou l'*Introduction à l'étude de la médecine expérimentale* de Claude Bernard. Une telle exploration bibliométrique a également le mérite de mettre au jour les critiques qui semblent manier les références scientifiques avec le plus d'aisance, en ne se cantonnant pas à la simple vulgate mais en élargissant l'éventail de titres sur lesquels ils s'appuient. Nous retrouvons ici ceux qui intéresseront le plus notre étude ; soit les critiques qui, sans se contenter de se faire l'écho d'une certaine actualité scientifique, usent abondamment et sciemment des références scientifiques, jusqu'à les intégrer à leur propre discours. Ainsi que permet de le visualiser le graphe ci-dessous, Brunetière, Gourmont, Hennequin, Taine, Lanson, Sainte-Beuve, Renan, Barbey, Bourget, Faguet, Guyau, Ribot, Nisard, Lamartine, Stapfer, Zola, Rod, Lemaitre, Deschanel, Barrès, Doumic, Pardo-bazan, France, Ghil, Bazalgette, Bouglé :



**Figure 11.** Diversité de titres scientifiques cités par critique

Chez cette demi-douzaine de critiques, l'occurrence et la variété de références scientifiques ne paraissent guère anodines : il ne s'agit pas seulement de citer les savants comme instances d'autorité, historiques ou sociales, mais de reprendre, commenter, discuter leurs ouvrages et leurs théories, de sorte d'initier un véritable dialogue transdisciplinaire. Intégrée au travers de ce réseau bibliographique, la science se fait à la fois source et objet du discours critique.

La répartition des titres au sein du corpus en dit alors long sur la manière dont chaque savant a été réceptionné par la critique littéraire et dont ses idées ont pu être influentes. Examinons par exemple le cas de Bacon, cité en abondance :



**Figure 12.** Répartition des titres de Bacon dans le corpus critique

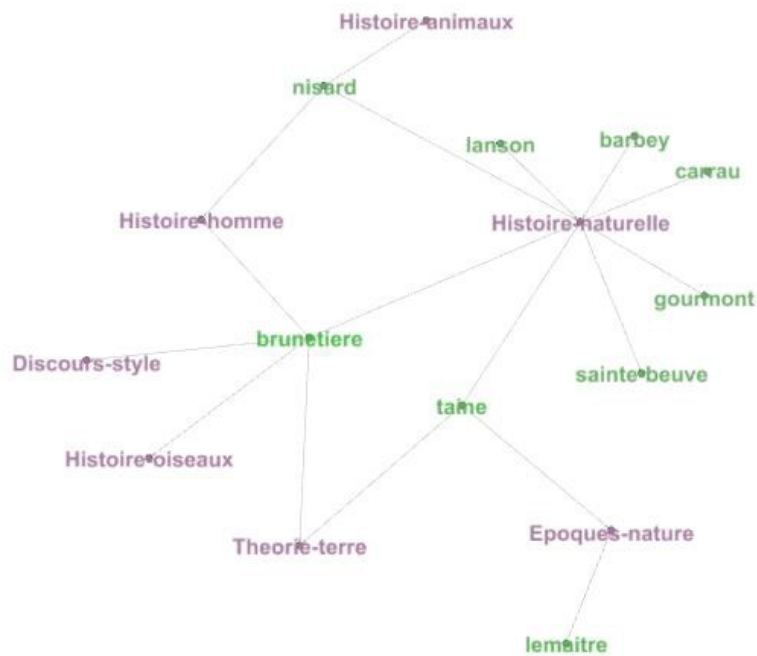
Bien que le *Novum Organum* reste la référence centrale, l'ensemble de l'œuvre baconienne est évoquée par les critiques comme fer de lance d'une Renaissance qui campe les fondements des sciences positives. Bacon incarne « le plus compréhensif, le plus sensé, le plus novateur des esprits du siècle » [TAI 66, 390] rappelle Taine, il est pour Brunetière le « “maître à penser” de la génération nouvelle » [BRU 98, 314], et tient à ce titre une place d'honneur dans l'histoire des sciences et des idées d'Europe. Mais surtout, la portée de son enseignement se redécouvre à une époque où l'on discute les méthodes scientifiques. Qu'elle soit directement lue ou enseignée, l'œuvre de Bacon continue de fournir aux penseurs et savants du XIX<sup>e</sup> siècle les procédés logiques du modèle empirique. Paul Janet insiste sur cette actualité des écrits baconiens au regard des récents travaux de Claude Bernard :

Pour ma part, en comparant le *Novum Organum* aux méthodes modernes, je suis beaucoup moins frappé de ce qu'il y a de suranné que de ce qu'on y trouve au contraire de neuf, de vivant, d'applicable encore. [JAN 72, 226]

Sa méthode, fondée la première sur l'observation et l'expérience, sert en quelque sorte d'égide sous laquelle viennent se placer les nouvelles prises de position scientifiques. Ainsi se trouve-t-il en épigraphe de *L'Origine des espèces*, de même qu'il est cité par Durkheim – directement en latin – dans *Les Règles de la méthode sociologique*. La critique littéraire, également en quête de méthode et d'assises positives, n'est pas en reste : « et nous aussi, simple historien littéraire, il est un côté par lequel nous ne saurions assez vénérer Bacon et le saluer, comme notre premier guide et inventeur » [S-B 62], déclare Sainte-Beuve, en se proclamant lui-même « disciple de Bacon » [S-B 26]. En tant que père de la pensée moderne, le savant britannique figure de ce fait comme auteur source, et ses textes, bien connus, participent activement à la transtextualité qui nourrit la critique littéraire, à l'intersection des sciences et de la philosophie.

Il en va autrement de Buffon, pour lequel les références sont presque exclusivement ciblées autour de *L'Histoire naturelle*, tel que le synthétise la figure ci-contre :

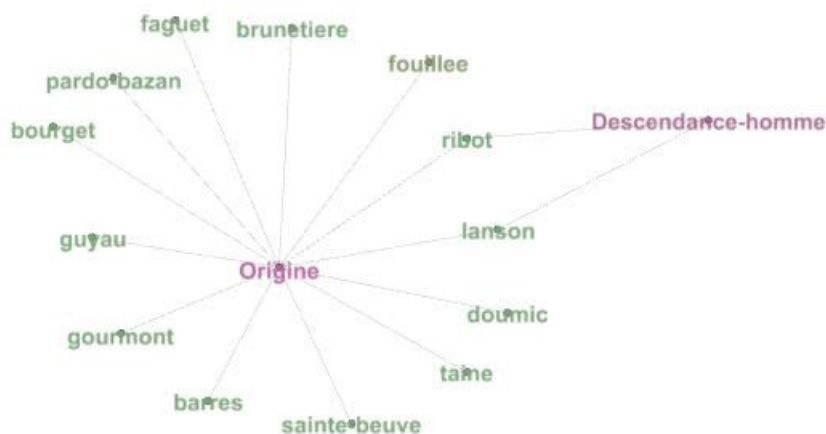




**Figure 13.** Répartition des titres de Buffon dans le corpus critique

L'inégale répartition des références bibliographiques dévoile en effet l'importance de son *Histoire naturelle*, auquel le nom de Buffon reste presque exclusivement attaché : c'est bien son effort de classification des espèces animales et minérales qui le rend avant tout célèbre, entreprise gigantesque qui lui vaut en 1753 d'être élu membre de l'Académie française. L'ouvrage demeure, un siècle plus tard, un monument des sciences naturelles, et incarne plus généralement les progrès réalisés par l'homme dans son acquisition encyclopédique de la connaissance. Souvent citée par les critiques au cours de leurs études historiques, *L'Histoire naturelle* scelle la volonté persistante de classer et d'ordonner les phénomènes naturels, et vient en ce sens servir de modèle à la nouvelle critique historique en pleine émergence. Elle livre, surtout, les grilles conceptuelles d'une science liant la notion d'espèce à l'étude du particulier : « Avant Buffon, retrace Nisard, on n'étudiait que l'individu, on négligeait l'espèce ; il apprit à mieux étudier l'un, et il créa la science de l'autre. » [NIS 83, 390] Mais hors de l'agglomérat référentiel qui entoure *L'Histoire naturelle*, la visualisation ci-dessus rend également compte de la façon dont Taine, Brunetière et Nisard multiplient les références bibliographiques à Buffon, en détaillant les titres des différents volumes qui composent son *Histoire* – enrichi des *Époques de la nature*, données en supplément, et de son fameux *Discours sur le style* prononcé lors de sa réception à l'Académie française. Cette diversité plus remarquable témoigne, outre une connaissance approfondie de l'œuvre elle-même, que les auteurs ont au moins partiellement lue, l'étude détaillée que chacun d'eux consacre au naturaliste dans leurs histoires : le *Manuel de l'histoire de la littérature française* pour Brunetière, *Les Origines de la France contemporaine* pour Taine, et *L'Histoire de la littérature française* pour Nisard.

Cet attachement du savant à une œuvre phare se retrouve, de façon plus magistrale encore, avec la figure de Darwin ; titre scientifique cité le plus fréquemment par la critique littéraire de la seconde moitié du XIX<sup>e</sup> siècle, *De l'origine des espèces* illustre le « canon scientifique » par excellence, au point que l'image de Darwin devient indissociable, dans le discours critique, de l'essai qui l'a rendu notoire.



**Figure 14.** Répartition des titres de Darwin dans le corpus critique

Cette répartition bibliographique trahit à la fois la célébrité particulière de ce best-seller, qui dépasse aussi bien les frontières disciplinaires que nationales, et la méconnaissance *de facto* d'une œuvre qui se passe souvent d'être lue. Peu, en effet, poussent l'étude de Darwin jusqu'à faire mention de ses essais suivants, pourtant accueillis avec enthousiasme par le grand public, parmi lesquels *La Descendance de l'homme et la sélection naturelle* (1871), traduit en 1872 par J.-J. Moulinié, et *L'Expression des émotions chez l'homme et les animaux* (1872), traduit en 1874 par les D<sup>rs</sup> Samuel Pozzi et René Benoît. Évoquer Darwin, c'est l'attacher – le cantonner, pourrait-on dire – à la théorie des espèces et de la sélection naturelle autour de laquelle se cristallise l'évolutionnisme des années 1870. Ici, le titre *De l'origine des espèces* active au sein du discours critique une double fonction référentielle et symbolique, en renvoyant à la théorie darwinienne des sciences de la vie tout en représentant plus largement le changement paradigmatique du siècle.

## Conclusion

En identifiant la nature et le poids des noms cités, mais aussi l'interaction que ceux-ci entretiennent dans les textes à travers leur coprésence, l'analyse de réseaux rend visible l'agencement interne d'une critique littéraire en formation, qui convoque des références exogènes et produit ses propres figures d'autorité afin d'appuyer la positivité de son discours. Du reste, l'intégration de savants au sein d'un discours à vocation littéraire pose clairement les revendications d'une critique qui se veut au cœur des savoirs. Si parler de science consiste d'abord à se faire l'écho d'une vision du monde, à promouvoir les découvertes scientifiques et à entériner les ruptures épistémiques du temps, une telle démarche ne s'envisage pas hors d'un questionnement ontologique de la critique littéraire. En s'arrogeant une place dans le paysage scientifique, cette dernière, jusqu'alors attachée à la littérature et, dans le meilleur des cas, à un pan de la philosophie, ne se contente pas d'élargir son champ : elle cherche consécutivement à conquérir l'autorité de ce qui fait savoir.

## Bibliographie

- [ABI 13] Abi Haidar, A., Six, A., Ganascia, J. G., & Thomas-Vaslin, V., « The Artificial Immune Systems Domain: Identifying Progress and Main Contributors Using Publication and Co-Authorship Analyses », in *Advances in Artificial Life*, ECAL, vol. 12, 2013, p. 1206-1217.
- [ABI 16] Abi-Haidar A., Wild, O., and Ganascia J.G, « A Simple yet Efficient Method for Unsupervised Named Entity Recognition and Disambiguation », *WIRES Data Mining and Knowledge Discovery* (DMKD), 2016.

- [BAS 09] Bastian, M., « Gephi : An Open Source Software for Exploring and Manipulating Networks », in *AAAI Publications, Third International AAAI Conference on Weblogs and Social Media*, 2009. La première version de Gephi date de 2008, et sa version la plus récente est téléchargeable en ligne. URL : <https://launchpad.net/gephi/0.6/0.6alpha1>
- [BRASS 92] Brass D. J. and Burkhardt M. E., « Centrality and Power in Organizations », in *Networks and Organizations*, Nohria N. et Eccles R.G. (ed.), Harvard Business School Press, Boston, 1992, p. 191-215.
- [BRU 98] Brunetière, F., *Manuel de l'histoire de la Littérature française*, Paris, C. Delagrave, 1898.
- [COL 17] Cole, F. J., et Eales, N. B., « The history of comparative anatomy. Part I: A statistical analysis of the literature », in *Science Progress*, 11, 1917, p. 578-596.
- [DES 37] Descartes, R., *Discours de la méthode*, Paris, 1637.
- [DOU 03] Doumic, R., *Hommes et idées du XIX<sup>e</sup> siècle*, Paris, Perrin, 1903.
- [FRE 79] Freeman, L. C., « Centrality in Social Networks : Conceptual Clarification », in *Social Networks*, vol. 1, 1979, p. 215-239.
- [GA 79] Garfield, E., *Is Information Retrieval in the Arts and Humanities Inherently Different from That in Science?* in *The Library Quarterly: Information, Community, Policy*, vol. 50, n°1, *Proceedings of the Fortieth Annual Conference of the Graduate Library School*, 18-19 may, 1979.
- [GA 83] Garfield, E., *Citation Indexing. Its Theory and Application in Science, Technology and Humanities*, Philadelphia, ISI Press, 1983.
- [GEN 82] Genette, G., *Palimpsestes. La Littérature au second degré*, Paris, Seuil, 1982.
- [GOU 25] Gourmont, R. de, *Promenades philosophiques. Troisième série*, Mercure de France, Paris, 1925.
- [JAN 72] Janet, P., *Les Problèmes du XIX<sup>e</sup> siècle : la politique, la littérature, la science, la philosophie, la religion*, Paris, Michel Lévy frères, 1872.
- [JORG 08] Jörg, B., *Towards the Nature of Citations*, German Research Center for Artificial Intelligence, Language Technology Lab, Saarbrücken, Germany, 2008.
- [LAN 95] Lanson, G., *Histoire de la littérature française*, Paris, Hachette, 1895.
- [LIU 05] Liu, X., Bollen, J., Nelson, M. L., & Van de Sompel, H., « Co-authorship networks in the digital library research community », in *Information processing & management*, 41(6), 2005, p. 1462-1480.
- [MEN 11] Mendes, P. N., et al, « DBpedia spotlight: shedding light on the web of documents », *Proceedings of the 7th International Conference on Semantic Systems*, ACM, 2011.
- [MOS 14] Mosallem, Y. Abi-Haidar, A. and Ganascia J.-G., « Unsupervised Named Entity Recognition and Disambiguation: An Application to Old French Journals », in *Lecture Notes in Computer Science series*, vol. 8557, *Proceedings of ICDM 2014*, St. Petersburg, Russia, 2014.
- [NAD 06] Nadeau, D., Turney, P., & Matwin, S., *Unsupervised named-entity recognition: Generating gazetteers and resolving ambiguity*, 2006.
- [NIS 83] Nisard, D., *Histoire de la littérature française*, t. IV, 10<sup>e</sup> éd., Paris, Firmin Didot, 1883.
- [RAI 62] Raising, L. M., « Statistical bibliography in the health sciences », in *Bulletin the medical library Association*, 50, p. 450-461. URL : <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC197860/>
- [REN 87] Renan, E., *Discours et conférences*, Paris, Calmann Lévy, 1887.
- [S-B 26] Sainte-Beuve, C. A., *Mes Poisons : cahiers intimes inédits*, Paris, Plon, 1926.
- [S-B 62] Sainte-Beuve, C. A., *Portraits littéraires*, t. II, Paris, Garnier frères, 1862.
- [S-B 67] Sainte-Beuve, C. A., *Nouveaux lundis*, t. VIII, Paris, Michel Lévy frères, 1867.
- [SEG 11] Séginger, G., « Présentation. Penser et rêver le vivant », in *Le Vivant, Revue Romantisme*, n°154, 4<sup>e</sup> trimestre 2011, p.3-20.
- [STA 93] Stapfer, P., *Des Réputations littéraires. Essais de morale et d'histoire. Première série*, Paris, Librairie Hachette et C<sup>ie</sup>, 1893.
- [TAI 66] Taine, H., *Histoire de la littérature anglaise*, t. I, 2<sup>e</sup> éd. revue et augmentée, Paris, Hachette, 1866.

- [TEU 06] S. Teufel, A. Siddharthan, D. Tidhar, *An annotation scheme for citation function*, in *Proceedings of Sigdial-06*, Sydney, Australia, 2006.
- [THI 30] Thibaudet, A., *Physiologie de la critique*, Paris, Éditions de la Nouvelle Revue Critique, coll. « Les Essais critiques », 1930.
- [WAS 94] Wasserman S. and Faust K., *Social network analysis : Methods and applications*, vol. 8, Cambridge university press, 1994.
- [WHI 04] White, H. D., « *Citation Analysis and Discourse Analysis Revisited* », in *Applied Linguistics*, 2004, p. 89-116.