

# ANALYSE DES RÉSEAUX DE PERSONNAGES DANS *LES CONFESIONS* DE JEAN-JACQUES ROUSSEAU

YANNICK ROCHAT

FRÉDÉRIC KAPLAN

Cet article étudie le concept de centralité dans les réseaux de personnages apparaissant dans *Les Confessions* de Jean-Jacques Rousseau. Notre objectif est ainsi de caractériser certains aspects des rôles des personnages du récit sur la base de leurs cooccurrences dans le texte.

DOI:10.3166/LCN.10.3.109-133 © 2014 Lavoisier

## 1. Introduction

Dans une œuvre littéraire, les « unités constitutives » de l'histoire, que nous désignons aussi comme des entités narratives, sont les événements, personnages, lieux et les temps de l'action (Bal, 1984). Le réseau formé par les apparitions et interactions des personnages en particulier est souvent secondaire dans l'analyse, ou n'est que partiellement exploré lorsqu'une étude porte sur la composition d'une intrigue et l'intrication des personnages dans celle-ci.

Dans cet article, nous proposons une analyse du réseau des personnages du texte sur la base de leurs cooccurrences. Les personnages sont ainsi étudiés non pas en fonction de leur rôle dans l'intrigue, mais en fonction de leurs apparitions dans l'ouvrage. À partir de l'index des personnages, nous générons un réseau représentant le système des personnages (Woloch, 2003).

Nous appliquons les méthodes de l'analyse de réseaux en littérature (Moretti, 2011 ; Elson, 2012 ; Sack, 2013) aux *Confessions* de Rousseau, texte de fiction autobiographique (Lejeune, 1975) composé d'un grand nombre de personnages. Cette richesse a été un facteur déterminant dans le choix de cette œuvre pour notre étude. Dans cet article, nous ne discutons pas du caractère autobiographique des *Confessions*, ni des relations entre le récit des *Confessions* et la vie de son auteur.

Nous basons nos travaux sur l'édition des œuvres complètes parue aux éditions Slatkine<sup>1</sup> en 2012. L'index de cette édition identifie 584 personnages sur un total de 774 pages. Rousseau n'apparaît pas lui-même dans l'index. Nous faisons usage du concept de centralité, une famille de mesures mettant en évidence les propriétés structurelles des sommets du réseau. Ces mesures permettent de définir des types de comportements ou de détecter les rôles joués par les personnages de l'œuvre en fonction de leurs positions dans le réseau. La prise en compte simultanée de plusieurs indices de centralité permet de proposer une caractérisation des personnages.

---

1. Accessible sur [rousseau.slatkine.com](http://rousseau.slatkine.com) Nous avons utilisé cette édition plutôt que d'autres versions du texte disponibles en ligne (Wikisource, Gallica) pour son index.

Nous commençons par définir l'analyse de réseaux en littérature. L'objectif est d'étudier l'organisation et la place des personnages dans le réseau. Puis, nous présentons la méthode de construction du réseau des cooccurrences. Nous justifions nos choix et décrivons les étapes nécessaires pour l'obtenir, et commentons ensuite les propriétés de ce réseau. Nous calculons les indices de centralité des personnages et montrons comment, en composant ces indices, cette méthode permet de caractériser les rôles narratifs des personnages d'un roman. Finalement, nous discutons le rôle préliminaire de cette étude pour mener des travaux à l'échelle de grandes collections de textes, sur la base des index.

## 2. Analyse de réseaux en littérature

### 2.1. Définition

L'analyse de réseaux en littérature étudie l'organisation des entités narratives au moyen de la théorie des graphes (Berge, 1958). Un réseau est constitué de sommets représentant les entités narratives et d'arêtes qui les lient entre elles, le cas échéant. La présence ou l'absence d'une arête entre deux sommets dépend de la méthode d'identification des relations. Quand aux méthodes d'analyse, elles sont inspirées de celles de l'analyse des réseaux sociaux (Wasserman et Faust, 1994 ; Borgatti *et al.*, 2009).

Dans ce travail, nous limitons l'étude des entités narratives aux seuls personnages.

### 2.2. Travaux existants

Il existe de nombreux modèles d'extractions et d'études de réseaux de personnages tirés d'œuvres littéraires. Les domaines dans lesquels ces recherches sont conduites comprennent les sciences humaines, tout comme la physique statistique et l'informatique. Nous remarquons une augmentation de la popularité de ce type de recherche.

De nombreux travaux (Pentland et Feldman, 2007 ; Gil *et al.*, 2011 ; Elson, 2012 ; Corman *et al.*, 2013 ; Bolioli *et al.*, 2013 ; Nalisnick et Baird, 2013) s'intéressent à l'étape préliminaire de construction du réseau, notamment grâce à des méthodes d'extraction automatique des relations entre personnages, voire des contextes de l'action ou des relations.

D'autres se focalisent sur l'analyse du ou des réseaux obtenus, et éventuellement les rôles des personnages (Stiller *et al.*, 2003 ; Moretti, 2011 ; Mac Carron et Kenna, 2012 ; Agarwal *et al.*, 2012 ; Sack, 2012, 2013 ; Grandjean, 2014 ; Sparavigna, 2013). Enfin, choisissant un point de vue plus vaste, diverses études (Alberich *et al.*, 2002 ; Gleiser, 2007 ; Rydberg-Cox, 2011 ; Hutchinson *et al.*, 2012) proposent des analyses de réseaux de personnages à partir d'ensembles de récits tirés chacun d'un univers cohérent, par exemple des sagas islandaises (Mac Carron et Kenna, 2013).

En physique statistique, l'utilisation de tels réseaux sert d'abord à l'expérimentation de méthodes algorithmiques (Newman et Girvan, 2004). Le réseau n'est donc plus un objet narratif, ainsi que le signale Donald E. Knuth à propos d'un célèbre jeu de données tiré des *Misérables* de Victor Hugo (Knuth, 2013) :

*Since the data files were prepared by hand, they are subject to human error. [...] They are intended simply as forever-frozen examples of typical data that is more or less accurate.*

*In particular, I recently learned that I forgot to include any connection between Fantine and her infant daughter Cosette, when I summarized the encounters between the characters of Les Misérables in the data file jean.dat. Fantine and Cosette appear together in chapter 1.4.1 (and then, of course – alas – they never see each other again). [...] However, I shall never update the file jean.dat, because it is “correct by definition.”*

### 2.3. Le système des personnages

La construction d'un réseau de personnages à partir d'une œuvre littéraire s'inscrit dans le cadre de la théorie des personnages définie par Alex Woloch (2003). Nous fournissons à cette théorie un outil – l'analyse des réseaux en littérature – permettant de conduire une telle analyse de manière quantitative. Dans cet article, nous utilisons les termes « système » et « structure » au sens de Woloch.

Tout d'abord, nous définissons l'*espace du personnage* :

*That particular and charged encounter between an individual human personality and a determined space and position within the narrative as a whole.* (Woloch, 2003, p. 14)

L'espace du personnage est la projection d'un individu de l'histoire<sup>2</sup> dans le récit<sup>3</sup>. Par exemple, on attend de l'espace du personnage protagoniste qu'il soit plus riche que l'espace d'un personnage secondaire. La narration implique que les scènes où apparaît un personnage secondaire soient plus rares, donc qu'une partie de l'information existante dans l'histoire soit omise ou réduite dans le récit. Woloch forme ensuite l'ensemble de tous les espaces de personnages. Il le nomme le *système des personnages* et le définit comme :

*The arrangement of multiple and differentiated character-spaces — differentiated configurations and manipulations of the human figure — into a unified narrative structure.* (Woloch, 2003, p. 14)

Dans le même temps, il définit une opération sur le système des personnages qui s'apparente à une forme mixte des concepts de réunion et d'intersection tels qu'on les comprend dans le domaine de la théorie des ensembles :

*For the character-system offers not simply many interacting individuals but many intersecting character-spaces, each of which encompasses an embedded interaction between discretely implied person and the dynamically elaborated narrative form.* (Woloch, 2003, p. 17)

Cette approche topologique<sup>4</sup> fournit un cadre théorique bien défini pour l'observation de la narration, transformation de l'individu du monde virtuel qu'est l'histoire en un personnage du récit. À ce stade, pour décoder le rôle du personnage dans la narration, il est nécessaire de considérer sa position dans le flot du récit relativement aux entités narratives, dont, en particulier, les autres personnages :

2. « Une histoire est une série d'événements logiquement reliés entre eux, et causés ou subis par des acteurs. » (Bal, 1984, p. 4)

3. « Un récit est le signifié d'un texte narratif. » (Bal, 1984, p. 4) « L'histoire ordonnée, structurée est le récit. » (Bal, 1984, p. 8)

4. Topologique car nous avons défini un ensemble accompagné d'un système de partitions de cet ensemble lui-même muni de règles d'union et d'intersection. S'il s'agissait d'un véritable espace topologique, l'union tout comme l'intersection d'espaces de personnages formeraient des espaces de personnages à leurs tours. Donc chaque élément de l'espace topologique doit toujours contenir au moins un personnage.

*[...] all character-spaces inevitably point us toward the character-system, since the emplacement of a character within the narrative form is largely comprised by his or her relative position vis-à-vis other characters. (Woloch, 2003, p. 18)*

Nous représentons le système des personnages à l'aide de l'analyse de réseaux en littérature. Celle-ci étudie la structure formée par les personnages, leurs relations et l'organisation de ces relations. Nous déduisons ensuite une partie des propriétés narratives d'un personnage à partir de sa position dans le réseau des interactions des personnages.

#### 2.4. Centralité

La centralité quantifie des propriétés structurelles des sommets du réseau interprétées comme des formes d'importance au sein de celui-ci (Hennig *et al.*, 2012). Par exemple, la centralité de degré mesure le nombre de connexions d'un sommet, et donc l'intrication de l'espace d'un personnage dans les espaces d'autres personnages. La centralité d'intermédiation, quand à elle, mesure la fréquence de placement d'un personnage sur les cheminements narratifs qui permettent de passer d'un personnage à un autre à travers le réseau des cooccurrences. Ainsi, entre autres possibilités, elle permet de mettre en évidence les personnages dont le rôle est central, ou ceux dont le rôle est de faire avancer le récit.

L'historique du développement du concept de centralité est de moindre importance ici en comparaison des questions que soulève ce concept, à savoir, entre autres, « que mesure la centralité ? », ou « comment utiliser et comparer ces mesures ? » (Sabidussi, 1966 ; Ruhnau, 2000 ; Koschützki *et al.*, 2005 ; Borgatti, 2005 ; Borgatti et Everett, 2006 ; Brandes *et al.*, 2012). De nombreuses formules mathématiques, établies à partir des réseaux, peuvent être considérées comme des indices de centralité. En l'absence d'une théorie de la centralité qui uniformiserait ces formules et permettrait de travailler dans un cadre mathématique clairement défini, c'est un article de Linton C. Freeman présentant les centralités de degré, de proximité et d'intermédiation (Freeman, 1978) et les travaux de Phillip Bonacich présentant la centralité par vecteurs propres (Bonacich, 1972, 1987) qui font référence et permettent de considérer ces quatre indices comme canoniques.

### 2.4.1. Degré

La centralité de degré d'un sommet s'exprime généralement par la somme des arêtes incidentes, ou de manière équivalente par la somme des sommets adjacents. C'est un indice ne prenant en compte que les éléments les plus proches du sommet, et ignorant le reste de la structure du réseau. Exprimé dans les termes de Degenne et Forsé (1994) :

La centralité de degré privilégie le point de vue local et mesure l'activité ou la capacité de communication ou d'échange de chaque individu au sein du réseau, en ne tenant pas compte de sa capacité à contrôler ces communications.

C'est un indice simple à mesurer et dont la lecture est intuitive. Mathématiquement, pour un sommet  $x_i$  dans un réseau d'ordre<sup>5</sup>  $n$  non dirigé dont la matrice d'adjacence est  $\{a_{ij}\}_{1 \leq i, j \leq n}$ , la centralité de degré est :

$$C_D(x_i) = \sum_j a_{ij}.$$

### 2.4.2. Proximité

La centralité de proximité consiste à évaluer la distance séparant un sommet de tous les autres. Cet indice est à l'origine du concept même de centralité (Bavelas, 1950). La centralité de proximité mesure l'effort nécessaire à un acteur du réseau pour atteindre tous les autres acteurs, par exemple lors de la diffusion d'une information. Dans l'analyse de réseaux littéraires appliquée au personnage, cet indice propose une forme de proximité narrative d'un personnage avec tous les autres. Pour ce travail, nous utilisons la centralité harmonique, dont les propriétés sont très proches et les résultats fortement corrélés avec la centralité de proximité (Rochat, 2009 ; Boldi et Vigna, 2013), mais dont la variance plus grande offre une meilleure lisibilité. Nous la définissons, pour un sommet  $x_i$ , par

$$C_H(x_i) = \sum_{i \neq j} \frac{1}{dist(x_i, x_j)}$$

avec  $dist(x_i, x_j)$  la distance du sommet  $x_i$  au sommet  $x_j$ .

---

5. En théorie des graphes, l'ordre est donné par le nombre de sommets.

### 2.4.3. Intermédiation

La centralité d'intermédiation mesure la propension d'un nœud à apparaître sur les plus courts chemins reliant les sommets du réseau. En sciences sociales tout comme dans des applications aux réseaux informatiques, les valeurs obtenues mettent en évidence les acteurs ou objets qui contrôlent l'information, ou par lesquels celle-ci transite.

Dans un réseau littéraire, la centralité d'intermédiation met en évidence les personnages apparaissant dans des événements du récit communs à plusieurs personnages. S'il n'y a qu'un seul personnage dans ce cas, sa mesure de centralité d'intermédiation sera significativement plus élevée que celles de ses voisins. Elle est donnée par :

$$C_B(x_i) = \sum \sum \frac{g_{jk}(i)}{g_{jk}}$$

avec  $g_{jk}$  le nombre de plus courts chemins allant du sommet  $x_j$  au sommet  $x_k$ , et  $g_{jk}(i)$  le nombre de plus courts chemins allant du sommet  $x_j$  au sommet  $x_k$  et contenant le sommet  $x_i$ . La double somme est calculée sur toutes les paires  $(j, k)$  telles que  $j = i = k$  et  $j < k$ .

### 2.4.4. Centralité par vecteurs propres

La centralité par vecteurs propres<sup>6</sup> attribue aux sommets une valeur tenant compte des centralités des sommets adjacents du réseau. Ainsi, il n'est pas possible de calculer la valeur de l'indice pour un sommet sans connaître celle des autres, et vice versa. Ce problème est résolu par l'utilisation de méthodes d'algèbre linéaire<sup>7</sup>.

Cet indice mesure l'importance d'un sommet en fonction de celle de ses voisins et propose, à l'instar des centralités de proximité et d'intermédiation, une description des sommets basée sur l'ensemble de la structure du réseau. Il est sensible à l'agglomération des sommets entre eux. Dans le contexte d'une étude structurelle du personnage, ceux qui ont

6. Ou centralité de pouvoir, ou centralité de Bonacich.

7. Pour permettre la résolution d'un système de  $n$  équations à  $n$  inconnues, pour  $n$  le nombre de sommets du réseau.



le plus de connexions, tout en étant reliés aux personnages les plus importants, obtiennent de plus hautes valeurs. Les mesures obtenues ne permettent pas de délimiter des groupes de personnages fortement connectés<sup>8</sup>, mais mettent en évidence les personnages qui sont aux centres de ces groupes.

### 3. Le réseau des cooccurrences

Dans cette section, nous décrivons le processus de génération du réseau des cooccurrences des personnages. Les choix effectués sont à prendre en compte lors de l'interprétation des résultats de l'analyse des centralités. Par exemple, les mesures de l'importance des personnages dans le système qui les réunit ne se lisent pas de la même manière si les relations sont basées sur une interprétation des contextes dans le texte narratif, ou sur une distance tenant compte des positions dans le texte, c'est-à-dire uniquement de la forme du récit. De telles mesures de distance peuvent reposer sur les mots, les phrases, les paragraphes, ou, dans notre cas, sur les pages.

Pour construire le réseau, nous recensons les cooccurrences des entités-personnages à partir des occurrences recensées dans l'index. Nous sommes tributaires de l'édition choisie, ainsi que des conventions adoptées par l'indexeur<sup>9</sup>, qui a choisi d'inclure les auteurs cités.

#### 3.1. Affinement

Nous ne suivons pas le scénario simpliste consistant à associer deux personnages s'ils apparaissent sur la même page<sup>10</sup>, car cette approche ne lierait pas deux personnages apparaissant par exemple dans la même phrase, l'un en bas d'une page et l'autre au sommet de la suivante. Pour générer le réseau des cooccurrences à partir de l'index des personnages, nous le transformons en un index de couples de pages : un personnage

8. Sur ce point, consulter la littérature sur les algorithmes de détection de communautés, par exemple (Fortunato, 2010).

9. Par exemple, l'« index des noms de personnes » des *Confessions* dans l'édition Classiques Garnier Poche (2011) contient des entrées pour des noms tels que Minerve ou Orphée.

10. Ce qui correspond à l'information brute telle qu'elle est fournie par un index.

apparaissant sur une page  $n$  dans le premier index appartient dans le second index aux couples de pages  $\{n-1, n\}$  et  $\{n, n+1\}$ . Cette méthode permet de substituer aux pages du premier index un système d'occurrences couplées et superposées<sup>11</sup>. Elle tient mieux compte de la proximité entre les personnages : deux sommets apparaissant sur la même page (cooccurrence forte) sont ainsi considérés par construction comme plus proches que deux sommets sur des pages consécutives (cooccurrence faible).

Cette approche par couples de pages résout le problème du format trop strict de la page, mais introduit du « bruit » : en admettant des occurrences plus larges, les cooccurrences qui en sont déduites sont également plus nombreuses. Pour pallier ce problème, nous fixons un seuil à atteindre pour qu'une relation soit prise en compte. En effet, lors de la compilation des cooccurrences, nous définissons un attribut de poids pour chaque relation, égal à la somme des cooccurrences faibles (dont la valeur ajoute une unité) et fortes (dont la valeur ajoute deux unités<sup>12</sup>). Une valeur minimale du poids égale à trois est requise pour la création de l'arête correspondante. La relation entre deux personnages atteint ce seuil lorsqu'ils partagent au moins une cooccurrence forte et une cooccurrence forte ou faible, ou au moins trois cooccurrences faibles.

En résumé, un personnage est retenu dans le modèle s'il existe au moins un autre personnage dans le roman partageant avec lui une relation jugée significative. C'est l'imbrication d'un personnage dans l'espace d'un ou d'autres personnages qui importe lors de la génération du réseau et de la définition des personnages à retenir. Ainsi, l'apparition dans des événements impliquant d'autres personnages, relatés sur plusieurs pages, garantit la présence du personnage dans le réseau final.

Le réseau des cooccurrences obtenu ne conserve pas tous les personnages présents dans l'index. Une analyse de l'ensemble des personnages de la narration est aussi possible : nous ne fixons alors pas de seuil. Ce réseau doit alors être interprété différemment, et les conclusions que nous en tirerions concerneraient d'autres questions, mais le problème du « bruit » demeurerait.

---

11. En effet, superposées car il n'y a plus de « vide » entre deux pages successives. Il n'est plus possible que deux personnages cités côte-à-côte ne puissent pas être considérés ensemble.

12. Puisqu'une cooccurrence forte et la réunion de deux cooccurrences faibles.

### 3.2. Premières propriétés du réseau

Le réseau que nous obtenons contient 226 personnages sur les 583 apparaissant dans le texte, ainsi que 613 arêtes. Il est simple et non dirigé. Sur les 357 personnages non retenus, 310 n'apparaissent qu'une seule fois dans le texte, 36 deux fois et 11 trois fois. En comparaison, aucun personnage n'apparaissant qu'une seule fois dans le texte n'est dans le modèle, mais 64 personnages apparaissant deux fois et 49 personnages apparaissant trois fois le sont. Dès lors, il faut prendre en compte le fait que les personnages situés en périphérie du réseau ne sont pas nécessairement des personnages anecdotiques dans le texte, puisqu'ils font partie des moins de 40 % retenus.

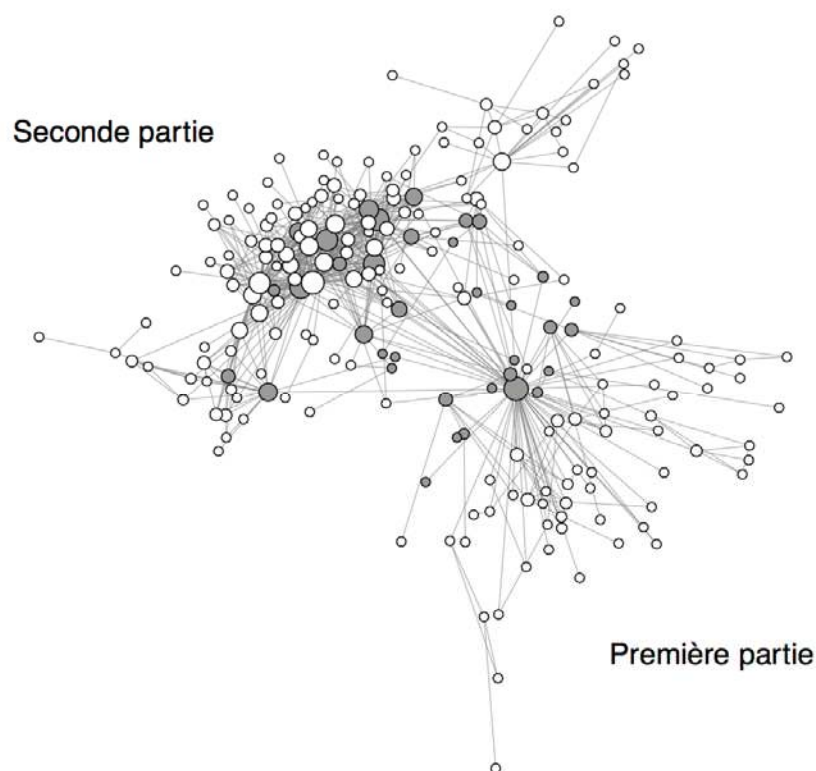
À ce stade, le réseau n'est pas connexe, car cinq couples de personnages gravitent hors de la composante géante<sup>13</sup>. Ces couples illustrent des événements narratifs importants, mais n'ayant pas ou trop peu d'interventions de personnages en commun avec les autres événements composant l'œuvre. Pour des raisons de calcul, nous restreignons le réseau des cooccurrences de personnages à la composante géante, et mettons donc de côté dix sommets et cinq arêtes qui étaient déconnectés du reste du réseau. Le résultat obtenu est montré à la figure 1.

### 4. Structure narrative et centralités

Nous avons exposé précédemment les divers indices de centralité nécessaires à cette étude. Ils nous permettent maintenant de décrire l'organisation de la structure narrative des personnages par l'auteur. Nous discutons les valeurs obtenues et en déduisons quelques résultats en nous basant sur une interprétation de ces valeurs et de diverses propriétés du réseau. Attention toutefois à faire la distinction : certains des résultats présents dans cette section sont entièrement déduits du réseau et renseignent sur le texte, d'autres sont au contraire des illustrations par le réseau de faits tirés du texte. Ainsi, le rôle du réseau est de soulever des questions comme d'y répondre, il peut être aussi bien variable expliquée que variable explicative.

---

13. Le plus grand sous-ensemble connecté de sommets, c'est-à-dire qu'il existe un chemin reliant chaque couple d'éléments de cette composante.



*Figure 1. Le réseau des cooccurrences des personnages, après restriction à la composante géante. Nous indiquons dans l'image où se situent les personnages des deux parties du livre. 16.5 % des personnages représentés dans ce réseau apparaissent dans les deux parties des Confessions, ils sont représentés en gris*

Des 216 sommets retenus dans le réseau des cooccurrences, le tableau 1 montre les douze sommets les plus centraux pour chacune des quatre mesures de centralité, réunis en un seul tableau afin de pouvoir les comparer. Nombre d'entre eux figurent plusieurs fois parmi les plus centraux, d'où le nombre réduit de sommets présentés. Les sommets sont ordonnés selon leurs degrés.

À l'aide du tableau 1 et de la figure 2, nous observons que la distribution des degrés décroît rapidement. En effet, quelques sommets

possèdent un grand nombre de connexions, tandis que la majorité en possède peu : la médiane est égale à trois, et le minimum est égal à un. Il existe des sommets beaucoup plus importants dans la trame que d'autres.

*Tableau 1. Mesures de centralité des douze sommets les plus centraux pour chaque indice, réunis en un seul tableau et classés en fonction de leurs degrés. Entre parenthèses, leurs rangs relativement à chaque indice*

Personnages	Degré	Interm.	Harm.	V.prop.
Mme de Warens	64 (1)	14926.69 (1)	34.9 (1)	0.19 (18)
Thérèse Levasseur	48 (2)	1231.19 (11)	26.3 (9)	0.79 (4)
Mme de Luxembourg	43 (3)	2969.67 (3)	29.7 (2)	0.48 (7)
Denis Diderot	40 (4)	2171.27 (5)	27.9 (4)	0.96 (3)
Grimm	38 (5)	1693.00 (6)	26.9 (6)	0.98 (2)
Louise d'Epinay	36 (6)	1282.83 (8)	25.6 (11)	1.00 (1)
Duc de Luxembourg	34 (7)	406.12 (33)	22.7 (44)	0.47 (8)
Mme Dupin	25 (8)	1295.08 (7)	27.2 (5)	0.28 (11)
George Keith	21 (9)	2198.60 (4)	26.3 (7)	0.08 (31)
Comtesse d'Houdetot	20 (10.5)	1079.13 (12)	26.3 (8)	0.64 (5)
Jean Le Rond d'Alembert	20 (10.5)	359.60 (35)	23.5 (30)	0.21 (14)
M. et Mme Levasseur	19 (12)	774.45 (16)	23.7 (24)	0.51 (6)
Voltaire	17 (15)	1259.88 (10)	27.9 (3)	0.12 (24)
Comtesse de Boufflers	17 (15)	125.62 (68)	22.8 (43)	0.20 (15)
Comte de Montaigu	16 (17.5)	3298.87 (2)	23.7 (23)	0.01 (123)
Charles Pinot Duclos	15 (20)	515.56 (24)	22.7 (45)	0.25 (12)
Baron d'Holbach	13 (22.5)	47.47 (87)	20.8 (74)	0.43 (9)
Marquise de Verdelin	13 (22.5)	474.45 (26)	25.0 (12)	0.05 (46)
Saint-Lambert	9 (28)	13.18 (103)	18.6 (128)	0.39 (10)
M. de La Poplinière	8 (36)	135.29 (66)	25.8 (10)	0.05 (44)
Mme de Larnage	7 (48.5)	1272.00 (9)	17.0 (161)	0.01 (106)

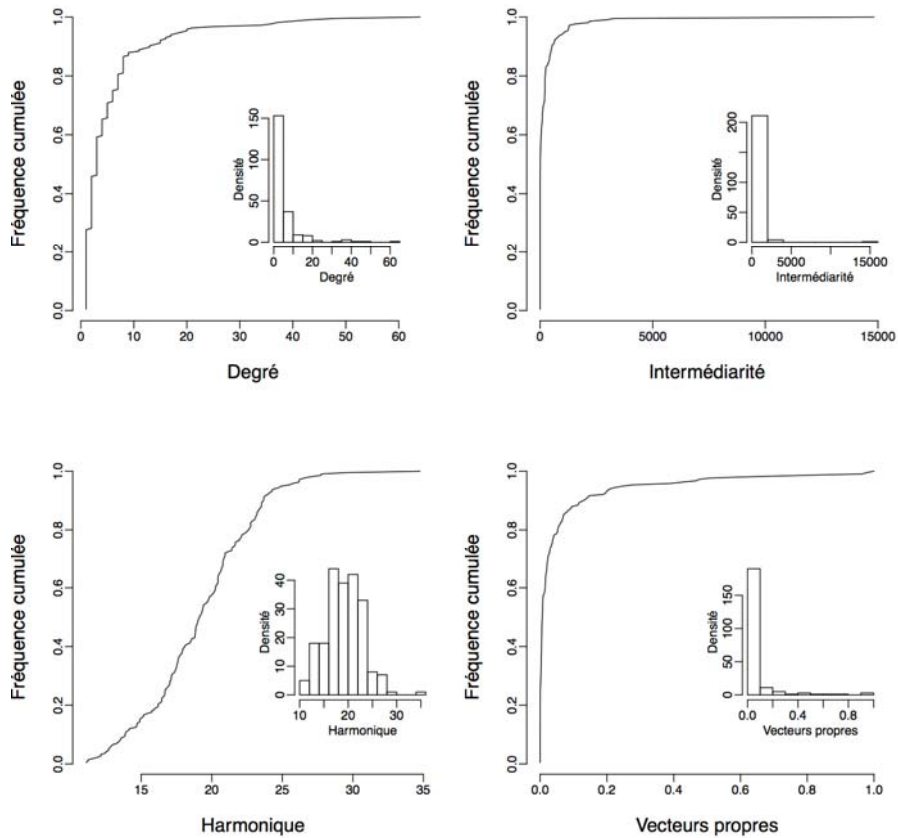


Figure 2. Fréquences cumulées et densités des mesures de centralité

La centralité d'intermédiarité et la centralité par vecteurs propres décroissent également rapidement. Dans ces deux cas, quelques sommets maximisent les propriétés mesurées respectivement par ces indices<sup>14</sup>, tandis qu'à l'opposé se trouvent des sommets pour lesquels les mesures sont nulles.

14. La valeur maximale de la centralité par vecteurs propres est égale à un. Elle est toujours atteinte par au moins un sommet.

La centralité harmonique suit une distribution dont la densité est symétrique. Les valeurs les plus hautes sont attribuées aux sommets proches du centre du réseau, les valeurs les plus basses aux sommets en périphérie. La majorité se situe dans une zone à mi-distance. Pour cet indice, médiane et moyenne sont égales.

À partir de ces mesures, nous discutons les rôles narratifs de quelques personnages ou groupes de personnages dans le récit.

#### 4.1. *Mme de Warens*

Les *Confessions* sont divisées en deux parties, chacune divisée en six chapitres. Il s'avère que Mme de Warens obtient des scores significativement plus élevés que les autres personnages pour les centralités de degré, d'harmonique et d'intermédiation. Elle est la figure principale de la première partie des *Confessions*, apparaissant majoritairement dans les chapitres deux à six<sup>15</sup>. Les sommets suivants dans les classements selon les indices correspondent tous à des personnages apparaissant pour la plupart dans la seconde partie, plus dense, comportant plus de personnages et induisant plus de relations dans le réseau des cooccurrences.

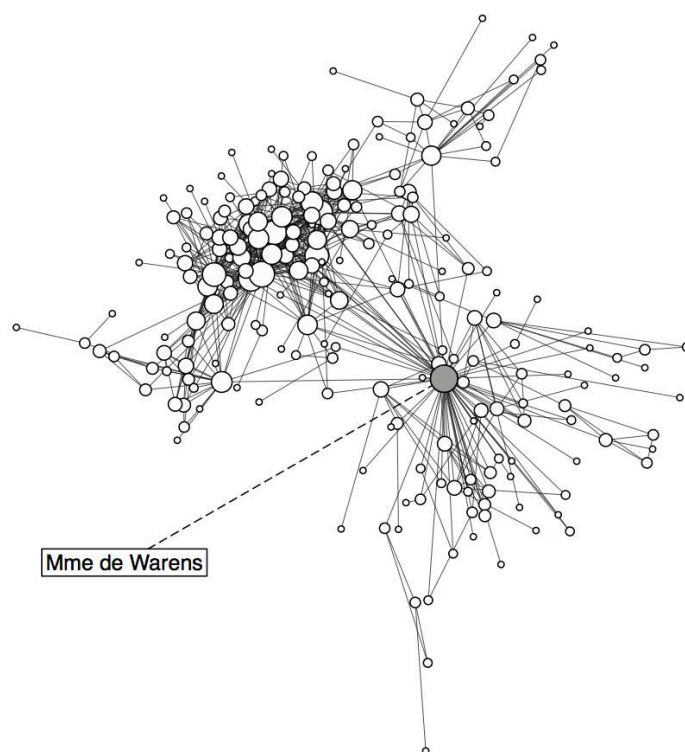
Il y a une absence de concurrence dans la première partie, construite autour du personnage de Mme de Warens. Son rôle de mentor auprès de Rousseau est décrit dans le texte. Il apparaît lié à de nombreux autres personnages dans le réseau, et par là à de nombreuses communautés. Son rôle de passeur narratif, qui se traduit par une centralité d'intermédiation très forte, est visible dans la figure 3.

Un indice fait exception dans ces classements : il s'agit de la centralité par vecteurs propres. En effet, cet indice renforce les scores des sommets dont les voisins ont des scores élevés. Le personnage de Mme de Warens est peu présent dans la seconde partie du roman, là où la vie sociale de Rousseau et des personnages qu'il côtoie est plus dense. Cette absence de relations vers des personnages issus de communautés fortes se trahit dans son score de centralité par vecteurs propres relativement bas. Il s'agit d'une différence importante dans les approches narratives respectives des deux

---

15. 87 % de ses apparitions se font dans la première partie. Dans la seconde partie, elle apparaît au moins une fois dans tous les chapitres, sauf au chapitre dix où elle n'apparaît pas.

parties : dans la première, Mme de Warens présente Rousseau à des personnages auxquels il ne s'attache jamais longtemps dans la narration. Les événements qui les contiennent sont souvent disjoints. Cela se retrouve dans le réseau, où ils occupent des positions relativement secondaires, traduites par des scores de centralité bas. Pour autant, cela ne signifie pas qu'ils sont moins présents en termes d'occurrences.



*Figure 3. Le réseau des cooccurrences des personnages, avec mise en évidence du sommet représentant Mme de Warens*

#### **4.2. Le métapersonnage du complot**

Nous avons vu que le coefficient de centralité par vecteurs propres ne considère pas Mme de Warens comme centrale (voir tableau 1). Au



contraire, et à l'exception de ses intimes<sup>16</sup>, les personnages obtenant des scores élevés sont principalement des membres de ce complot. Cela met en évidence un sous-groupe plus dense, et possédant des propriétés communes, du réseau des personnages. Sur la base de leurs scores, nous détectons notamment comme membres de celui-ci Louise d'Épinay, Friedrich Melchior Grimm, Denis Diderot, Jean le Rond d'Alembert, le baron d'Holbach, et la comtesse de Boufflers.

Au niveau de la narration, ces personnages forment un méta-personnage composé de figures interchangeables. Cette affirmation se justifie par le fait que l'existence de ce groupe et de ses ramifications demeure si nous en retirons un des personnages, leurs voisinages étant en partie communs. Dans le texte, Rousseau décrit ces individus à plusieurs reprises de manière groupée, tel un ennemi sans visage. Ce fait se trouve vérifié lorsque nous étudions la contraction des sommets représentant les membres du complot en un seul sommet. Cette opération consiste à remplacer les sommets choisis par un seul sommet héritant de toutes les connexions de ceux qui le composent. Ce que nous obtenons est alors la représentation du métapersonnage du complot dans le réseau. Elle nous permet d'observer l'emprise que possède cette réunion d'espaces de personnages au sein du système entier. Les mesures de centralité de ce sommet sont proches de celles du sommet représentant Mme de Warens (à l'exception toujours de la centralité par vecteurs propres, voir tableau 2).

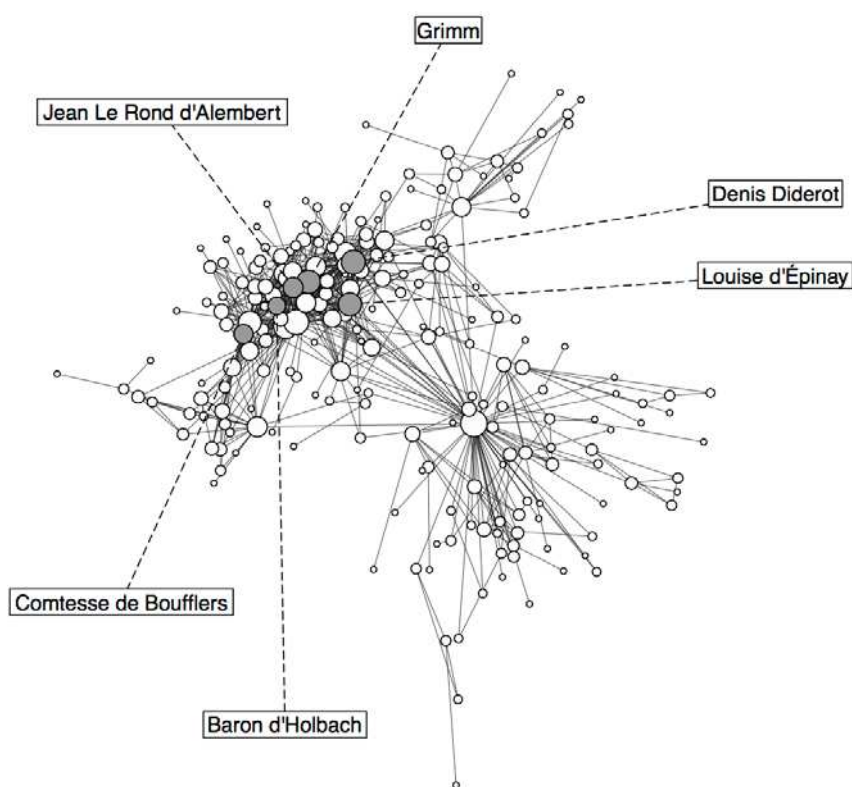
*Tableau 2. Mesures de centralité du sommet représentant Mme de Warens et du sommet réunissant les personnages du complot. Entre parenthèses, leurs rangs pour chaque indice respectivement*

Personnages	Degré	Interm.	Harm.	V.prop.
Complot	74 (1)	5910.47 (2)	29.95 (2)	1.000 (1)
Mme de Warens	62 (2)	13631.65 (1)	33.95 (1)	0.212 (12)

Leurs positions en première ou seconde place nous permettent d'affirmer que l'importance narrative donnée au groupe de personnages

16. Les Levasseur, M. et Mme de Luxembourg, la comtesse d'Houdetot, ...

derrière le complot est similaire dans la seconde partie à celle donnée à Mme de Warens dans la première partie. Néanmoins, il ne s'agit pas de prétendre que leurs rôles narratifs sont équivalents. Nous montrons dans la figure 4 la répartition des sommets des personnages décrits comme adhérant à ce complot. Nous observons la séparation du réseau des cooccurrences en deux parties, l'une disposée autour de Mme de Warens (voir figure 3), l'autre, plus diffuse, répartie le long des personnages du complot.<sup>17</sup>



*Figure 4. Le réseau des cooccurrences des personnages, avec mise en évidence de membres supposés du complot*

17. L'algorithme de répartition des sommets est donné par Fruchterman et Reingold (1991).

### 4.3. Principaux espaces de personnages

Dans cette section, nous analysons les mesures de centralité de quelques personnages plus en détail. Nous montrons ainsi comment l'analyse des réseaux en littérature permet d'avancer des hypothèses sur les aspects narratifs d'un texte. Les mesures de centralité mentionnées dans la suite de cette section se trouvent dans le tableau 1.

#### 4.3.1. Le baron d'Holbach, en retrait

Le baron d'Holbach est un des personnages que Rousseau soupçonne d'appartenir au complot. Il apparaît sur dix-huit pages, toutes dans la seconde partie des *Confessions*. Son degré de centralité est élevé en comparaison des autres sommets du réseau, ce qui signifie qu'il apparaît aux côtés de nombreux autres personnages et qu'il est impliqué dans cette société des personnages du roman. Cependant, pour ce qui est des centralités harmonique et d'intermédiarité, il se trouve en retrait. Ces indices traduisent pourtant des formes d'influence au sein du réseau. Ainsi, par ses choix narratifs, Rousseau présente ce personnage comme un suiveur : présent et bien connecté, il est peu influent. S'il fallait déduire des *Confessions* quel personnage Rousseau place à la tête du complot, il ne s'agirait probablement pas du baron d'Holbach.

#### 4.3.2. Louise d'Épinay - Friedrich Melchior Grimm - Denis Diderot

Ces trois personnages ont été parmi les plus proches amis de Rousseau, avant que ses relations avec chacun d'entre eux ne se détériorent. Ils possèdent, dans cet ordre, les trois premiers scores de centralité par vecteurs propres, et sont également considérés comme centraux selon tous les autres indices. Ainsi, nous en déduisons qu'ils occupent des positions narrativement influentes. Leur portée s'étend au-delà du groupe des « comploteurs » auquel ils appartiennent. Ils sont connectés à de nombreux personnages présents tout au long de la seconde partie des *Confessions*, ce qui dénote une influence potentiellement importante sur la vie de Rousseau, si l'on admet que celui-ci n'a pas consciemment mis en scène ce phénomène lors de la rédaction des *Confessions*.

#### 4.3.3. Le comte de Montaigu et George Keith, localement centraux

Le comte de Montaigu occupe un type de position quasiment unique dans le réseau, ce que l'on remarque facilement dans la figure 5. Sa

centralité d'intermédiation est élevée au point de le placer juste derrière Mme de Warens tandis que les autres indices ne le mettent pas en évidence. Cela s'explique par sa position d'ambassadeur de France à Venise, alors que Rousseau travaille pour lui comme secrétaire. À une exception près, il n'est mentionné que dans le chapitre sept, qui comprend justement les événements prenant place à Venise. Dans le réseau, il joue donc un rôle de relais par ses liens en France entre les personnages vénitiens et les personnages des autres événements relatés dans *Les Confessions*. Il est à l'écart des personnages constituant le noyau du réseau, donc doit être caractérisé comme personnage secondaire à l'échelle de l'ensemble du récit. Cependant, il émerge en leader d'un groupe clairement défini de sommets, et selon ce point de vue occupe une position centrale, reléguant à des rôles secondaires les autres personnages au sein de ce contexte.

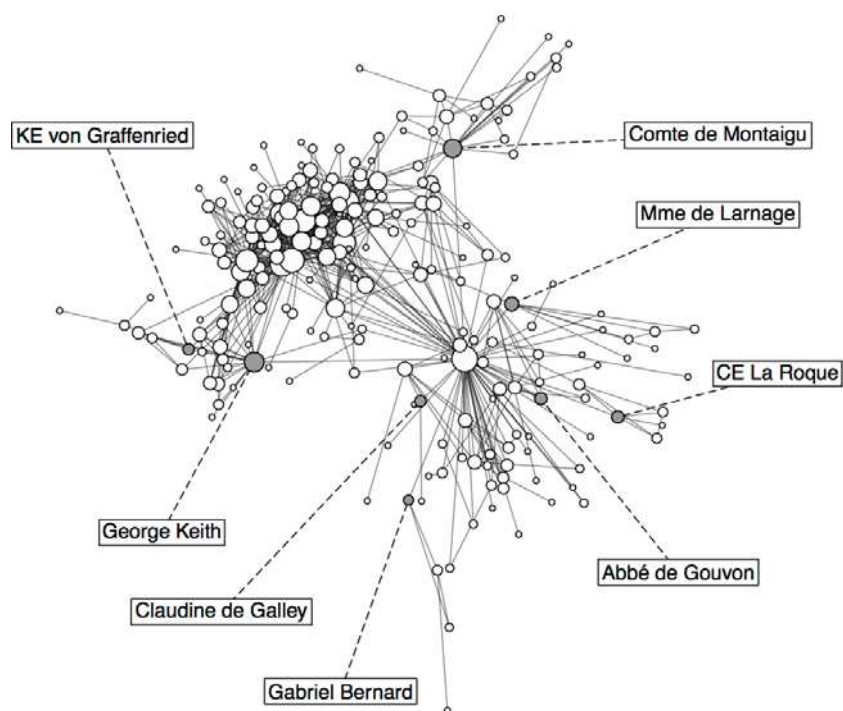


Figure 5. Le réseau des cooccurrences des personnages, avec mise en évidence du comte de Montaigu et de George Keith

De manière moins marquée, George Keith (alias Milord Marechal) occupe au chapitre douze<sup>18</sup> une position narrative similaire à celle du comte de Montaigu. Sa centralité d'intermédiation est très élevée, sa centralité par vecteurs propres très basse, et dans la figure 5 il apparaît dans un groupe à l'écart. Cependant, des connexions nombreuses entre les personnages de cette communauté des événements en Suisse<sup>19</sup> et les autres personnages de la seconde partie diminuent l'influence de George Keith sur ce groupe. Ces différences expliquent le contraste moins marqué entre les rangs de ses mesures de centralité qu'entre ceux du comte de Montaigu. Ce constat est peut-être même à relier avec la proximité géographique de la Suisse et des divers lieux visités en France, s'il s'avérait qu'elle augmente la densité du réseau des cooccurrences autour de George Keith, au contraire de Venise, plus distante.

## 5. Conclusion

Nous avons proposé une approche quantitative adaptée de la théorie des personnages d'Alex Woloch. De l'analyse des réseaux sociaux, nous avons importé et principalement utilisé le concept de centralité, à travers ses indices les mieux établis. Nous avons montré comment, au centre de chacune des deux parties du roman, Rousseau place un personnage autour duquel il structure le propos, et comment cela se retrouve dans chaque partie. Pour la première partie, on trouve au centre de la structure un personnage unique, mentor à l'aura positive. Pour la seconde partie, on identifie un collectif de personnages, négativement connoté, autour duquel les autres sommets sont distribués de manière compacte. Finalement, nous avons esquissé les principaux rôles narratifs que l'on peut déduire de cette analyse de réseaux.

Le travail que nous proposons dans cet article intègre lecture quantitative et qualitative. La même méthode pourrait sans aucun doute être appliquée à d'autres œuvres. Une question plus difficile est de savoir si une approche strictement quantitative, n'utilisant pas les connaissances spécifiques sur l'œuvre mais seulement l'analyse structurelle des réseaux de

---

18. Qui contient toutes ses occurrences (23), sauf une au chapitre deux qui n'est pas considérée comme significative par notre méthode d'extraction.

19. Plus précisément Neuchâtel.

personnages est concevable. Dans les approches de type *distant reading* qui envisagent l'analyse de milliers d'œuvres « sans ouvrir » le livre, il est en effet nécessaire de ne se baser que sur des éléments pouvant être calculés automatiquement. Les résultats que nous proposons dans cet article ne nous permettent pas de conclure sur la pertinence intrinsèque d'une approche uniquement quantitative, et donc applicable à des grands corpus d'œuvres. Néanmoins, il nous semble que la démarche proposée est suffisamment robuste en matière de centralité des personnages pour envisager avec sérénité une étude englobant plusieurs œuvres et qui nous permette de qualifier de manière plus précise leur pertinence à une autre échelle.

#### *Remerciements*

*Nous tenons à remercier le comité éditorial pour ses relectures attentives et son travail de coordination.*

#### **Bibliographie**

- Agarwal A., Corvalan A., Jensen J., Rambow O. (2012). Social network analysis of Alice in wonderland. *NAACL-HLT 2012*, p. 88-96.
- Alberich R., Miro-Julia J., Rossello F. (2002). *Marvel Universe looks almost like a real social network*. arXiv :cond-mat/0202174.
- Bal M. (1984). *Narratologie. Essais sur la signification narrative dans quatre romans modernes*. HES Publishers, Utrecht.
- Bavelas A. (1950). Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America*, vol. 22, p. 725-730.
- Berge C. (1958). *Théorie des graphes et ses applications*. Dunot, Paris.
- Boldi P., Vigna S. (2013). *Axioms for centrality*. arXiv :1308.2140 [physics].
- Bolioli A., Casu M., Lana M., Roda R. (2013). Exploring the Betrothed Lovers. In Mark A. Finlayson, Bernhard Fisseni, Benedikt Löwe et Jan Christoph Meister, éditeurs : *2013 Workshop on Computational Models of Narrative*, p. 30-35, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- Bonacich P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, vol. 2, n° 1, p. 113-120.

- Bonacich P. (1987). Power and centrality : A family of measures. *American Journal of Sociology*, vol. 92, p. 1170-1182.
- Borgatti S. (2005). Centrality and network flow. *Social Networks*, vol. 27, n° 1, p. 55 – 71.
- Borgatti S., Everett M. (2006). A graph-theoretic perspective on centrality. *Social Networks*, vol. 28, n° 4, p. 466-484.
- Borgatti S., Mehra A., Brass D., Labianca G. (2009). Network Analysis in the Social Sciences. *Science*, vol. 323, n° 5916, p. 892-895.
- Brandes U., Kosub S., Nick B. (2012). Was messen Zentralitätsindizes ? In *Die Integration von Theorie und Methode in der Netzwerkforschung*, p. 33-52. M. Hennig und C. Stegbauer (Hrsg.), Springer.
- Corman S., Ball B., Talboom K., Brewer G. (2013). Assessing Two-Mode Semantic Network Story Representations Using a False Memory Paradigm. In Mark A. Finlayson, Bernhard Fisseni, Benedikt Löwe et Jan Christoph Meister, éditeurs : *2013 Workshop on Computational Models of Narrative*, p. 53-61, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- Degenne A., Forsé M. (1994). *Les réseaux sociaux*. Armand Colin, Paris.
- Elson D. (2012). *Modeling narrative discourse*. Thèse de doctorat, Columbia University, New York City.
- Fortunato S. (2010). Community detection in graphs. *Physics Reports*, vol. 486, n° 3–5, p. 75-174.
- Freeman L. (1978). Centrality in social networks : conceptual clarification. *Social Networks*, vol. 1, n° 3, p. 215-239.
- Fruchterman T., Reingold E. (1991). Graph drawing by force-directed placement. *Software : Practice and experience*, vol. 21, n° 11, p. 1129-1164.
- Gil S., Kuenzel L., Suen C. (2011). *Extraction and Analysis of Character Interaction Networks From Plays and Movies*. CS 224W final project report, Stanford.
- Gleiser P. (2007). How to become a superhero. *Journal of Statistical Mechanics : Theory and Experiment*, vol. 2007, n° 09, P09020.
- Grandjean M. (2014). Comparing the relational structure of the Gospels. *Network Analysis as a tool for biblical sciences*. Harvard. Center for Hellenic Studies. Soumis.
- Hennig M., Brandes U., Pfeffer J., Mergel I. (2012). *Studying Social Networks : A Guide to Empirical Research*. Campus Verlag.

- Hutchinson S., Datla V., Louwerse M. (2012). Social networks are encoded in language. *Proceedings of the 34th annual conference of the cognitive science society*, p. 491-496.
- Knuth D. (2013). *The Stanford GraphBase : A Platform for Combinatorial Computing*. <http://www-cs-faculty.stanford.edu/~uno/sgb.html>, consulté le 13.02.2014.
- Koschützki D.; Lehmann K., Tenfelde-Podehl D., Zlotowski O. (2005). Advanced Centrality Concepts. In Ulrik Brandes et Thomas Erlebach, éditeurs : *Network Analysis*, vol. 3418, p. 83-111. Springer, Berlin, Heidelberg.
- Lejeune P. (1975). *Le pacte autobiographique : Signes de vie*, volume 1. Seuil.
- Mac Carron P., Kenna R. (2012). Universal properties of mythological networks. *Europhysics Letters*, vol. 99, n° 2:28002.
- Mac Carron P., Kenna R. (2013). Network analysis of the Íslendinga sögur – the Sagas of Icelanders. *The European Physical Journal B*, vol. 86 n° 10, p. 1-9.
- Michel J.-B., Shen Y., Aiden A., Veres A., Gray M., Pickett J., Hoiberg D., Clancy D., Norvig P., Orwant J., Pinker S., Nowak M., Aiden E. (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, vol. 331 n° 6014, p. 176-182.
- Moretti F. (2011). Network theory, plot analysis. *New Left Review*, vol. 68.
- Nalisnick E., Baird H. (2013). Extracting Sentiment Networks from Shakespeare's Plays. In 2013 *12th International Conference on Document Analysis and Recognition*, p. 758-762.
- Newman M., Girvan M. (2004). Finding and evaluating community structure in networks. *Physical review E*, vol. 69 n° 2:026113.
- Pentland B., Feldman M. (2007). Narrative Networks : Patterns of Technology and Organization. *Organization Science*, vol. 18, n° 5, p. 781-795.
- Rochat Y. (2009). *Closeness centrality extended to unconnected graphs : the harmonic centrality index*. Applied Social Network Analysis, Zürich.
- Ruhnau B. (2000). Eigenvector-centrality — a node-centrality ? *Social Networks*, vol. 22 n° 4, p. 357-365.
- Rydberg-Cox J. (2011). Social networks and the language of greek tragedy. In *Journal of the Chicago Colloquium on Digital Humanities and Computer Science*, vol. 1.
- Sabidussi G. (1966). The centrality index of a graph. *Psychometrika*, vol. 31, n° 4, p. 581-603.



- Sack G. (2012). *Character Networks for Narrative Generation*. In Eighth Artificial Intelligence and Interactive Digital Entertainment Conference.
- Sack G. (2013). Character Networks for Narrative Generation : Structural Balance Theory and the Emergence of Proto-Narratives. In 2013 *Workshop on Computational Models of Narrative*, vol. 32, p. 183-197.
- Sparavigna A. (2013). On Social Networks in Plays and Novels. *The international journal of sciences*, vol. 2 n° 10, p. 20-25.
- Stiller J., Nettle D., Dunbar R. (2003). The small world of Shakespeare's plays. *Human Nature*, vol. 14, n° 4, p. 397-408.
- Voisine J.-J. (2011). Introduction. In *Les Confessions*. Garnier, Paris.
- Wasserman S., Faust K. (1994). *Social Network Analysis : Methods and Applications*. Cambridge University Press.
- Woloch A. (2003). *The One Vs. the Many : Minor Characters and the Space of the Protagonist in the Novel*. Princeton University Press.

