



Analyse de données textuelles informatisée. Comment la pensée complexe et l'approche relationnelle peuvent nourrir quelques considérations méthodologiques

Computer-Assisted Textual Data Analysis. How Complex Thought and a Relational Approach can Feed some Methodological Considerations

Roger Gervais

Volume 11, numéro 1, novembre 2015

Sur le thème de l'analyse de données textuelles informatisée

URI : <https://id.erudit.org/iderudit/1035936ar>

DOI : <https://doi.org/10.7202/1035936ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Prise de parole

ISSN

1712-8307 (imprimé)

1918-7475 (numérique)

[Découvrir la revue](#)

Citer cet article

Gervais, R. (2015). Analyse de données textuelles informatisée. Comment la pensée complexe et l'approche relationnelle peuvent nourrir quelques considérations méthodologiques. *Nouvelles perspectives en sciences sociales*, 11(1), 181–215. <https://doi.org/10.7202/1035936ar>

Résumé de l'article

Cet article explore deux préoccupations méthodologiques liées à la méta-analyse de données textuelles : 1) le danger de la « décontextualisation » des idées comme conséquence d'une quantification des données textuelles; et 2) le principe selon lequel il importe de déterminer le cadre historique et l'origine du document avant de procéder à l'analyse. Pour ce faire, nous nous appuyons sur le traitement de 11 020 articles de périodiques parus au Canada et en France en 2005 effectué par le logiciel SPAD. Nous concluons que ce logiciel répond bien à la première préoccupation. SPAD produit des analyses de facteurs des données lexicales tout en offrant au chercheur la possibilité de retourner au texte et de vérifier le « sens » des mots. Toutefois, notre étude de cas montre aussi comment un échantillon de cette taille rend difficile la prise en considération *a priori* du cadre historique et de l'origine du document. Pour montrer comment il est possible de réaliser des méta-analyses en dépit de cette difficulté, nous nous référons à des principes proposés par les études relationnelles et par la systémique complexe.

Analyse de données textuelles informatisée. Comment la pensée complexe et l'approche relationnelle peuvent nourrir quelques considérations méthodologiques

ROGER GERVAIS

Université Sainte-Anne, Pointe-de-l'Église (Nouvelle-Écosse)

Les travaux de la systémique complexe et des études relationnelles nous encouragent à « surmonter la pensée limitative¹ » et à pousser plus loin notre compréhension de la société et des individus qui la composent. Les approches dites relationnelles², systémiques complexes³ ou conjonctives⁴ répondent à d'importantes questions philosophiques. Elles sont, entre autres, placées devant la nécessité, issue de la mondialisation, de dépasser les frontières liées à l'espace et au temps dans le but de comprendre des phénomènes dans une perspective internationale. Cette

¹ Citation tirée de *l'Appel à communications de la revue : Nouvelles perspectives en sciences sociales. Pour le colloque de son 10^e anniversaire sur le thème Pensée systémique complexe et relationnelle*, <http://www.acpcpa.ca/fr/posting.php?postid=201405231>, site consulté le 28 novembre 2014.

² Simon Laflamme, *Contribution à la critique de la persuasion politique*, Québec/Sudbury, Presse de l'Université du Québec/Université Laurentienne, 1987, et *Suites sociologiques*, Sudbury, Prise de Parole, coll. « Épistémè », 2006.

³ Edgar Morin, *La méthode - 1. La nature de la nature*, Paris, Seuil, 1977, et *La méthode - 2. La vie de la vie*, Paris, Seuil, 1980.

⁴ Jean-Louis Le Moigne, *La modélisation des systèmes complexes*, Paris, Dunod, 1999.

situation impose des échantillons de plus en plus volumineux. Elle oblige aussi les chercheurs à se doter d'outils très performants dans le but de surmonter les obstacles attachés aux échantillonnages anciens et de faire écho à des préoccupations mondialisantes. Il va sans dire que la révolution informatique représente un atout. Si, à ses débuts, l'ordinateur était rudimentaire, des informaticiens, partout dans le monde, l'ont énormément amélioré, offrant désormais aux chercheurs le fruit de leur labeur. Tel est le cas de l'analyse de données textuelles informatisées, qui s'est aujourd'hui dotée d'outils fort utiles pour le spécialiste qui se donne pour principe de respecter, autant que faire se peut, la complexité sociale.

Cet article propose d'utiliser les données d'une recherche antérieure pour donner lieu à une étude de cas qui explorera deux préoccupations méthodologiques liées à l'analyse de données textuelles qui sont centrées sur le « sens » des idées qui circulent dans les médias : la « décontextualisation » et le cadre historique.

Ces deux préoccupations méthodologiques demandent à être présentées. La première gravite autour de la notion du danger liée à la « décontextualisation » des idées, ou à la perte de sens, comme conséquence d'une quantification des données textuelles. Il ne s'agit pas là d'une nouvelle préoccupation puisque Dominique Maingueneau signalait ce problème déjà en 1976⁵. Cette préoccupation est d'ordre technique et devra trouver une solution technique. La seconde considération méthodologique qui nous intéresse ici est l'importance de déterminer le cadre historique et l'origine du document avant de procéder à l'analyse. Comme l'explique Paul Sabourin, « [d]u point de vue d'une analyse qui se pose d'abord comme sémantique, le sens n'est pas dans le texte, mais dans la relation entre le producteur d'un texte, le texte et un récepteur⁶ ». Alors qu'une telle approche peut produire des analyses intéressantes, elle se prête mal à l'étude de

⁵ Dominique Maingueneau, *Initiation aux méthodes de l'analyse du discours*, Paris, Hachette, 1976, p. 36.

⁶ Paul Sabourin, « L'analyse de contenu » (chapitre 15), dans Benoît Gauthier (dir.), *Recherche sociale : de la problématique à la collecte des données*, Québec, Presses de l'Université du Québec, 2006 [2003], p. 379.

grands échantillons, comme nous le démontrerons plus loin, et elle est très anthropocentrique de nature – elle préconise que nous ne puissions comprendre le social que par ce qui motive ou ce qui influence les humains. Si les logiciels de traitement de texte, comme SPAD, nous permettent de faire des méta-analyses sans pour autant tenir compte du cadre historique ni de l'origine du document en *a priori*, ce sont les modèles relationnels et systémiques complexes qui nous permettront de comprendre, épistémologiquement, comment cela est possible.

1. Notre méthode, notre échantillon, notre étude de cas

Dans le cadre de nos recherches, nous nous sommes intéressés à la question de l'homogénéisation et de la différenciation. Bien qu'il s'agisse d'un sujet en vogue, Émile Durkheim nous rappelle que les présocratiques en parlaient il y a plus de 2 000 ans⁷. Depuis, un grand nombre d'auteurs importants en sciences sociales en ont aussi discuté : Karl Marx⁸, Georg Simmel⁹, Vilfredo Pareto¹⁰, Max Weber¹¹, les structuro-fonctionnalistes¹²,

⁷ Voir Émile Durkheim, *De la division du travail social*, Paris, Presses universitaires de France, 1991 [1893], p. 18. Héraclite d'Éphèse, Méliossos de Samos, Empédocle d'Argentine, Anaxagore de Clazomènes et Démocrite de l'École d'Abdère ont abordé la question chacun à sa façon (voir, par exemple, Jean Voilquin, *Les penseurs grecs avant Socrate*, Paris, Garnier-Flammarion, 1964).

⁸ Karl Marx et Friedrich Engels, *Manifeste du Parti communiste I, Manifeste du Parti communiste I, les articles de F. Engels dans la « Réforme »*, traduction nouvelle par Charles Andler, Paris, G. Bellais, 1983, p. 21. Karl Marx, *Œuvres, I*, Paris, Gallimard, Bibliothèque de la Pléiade, 1965, p. 165.

⁹ Georg Simmel, *Le Conflit*, traduit de l'allemand par Sibylle Muller, Dijon-Quetigny (France), Circé, 2003 [1995], p. 22-23.

¹⁰ Vilfredo Pareto, *Traité de sociologie générale*, Paris, Droz, 1968, p. 2025-2031.

¹¹ Le concept de différenciation chez Weber est plutôt présenté par d'autres auteurs. Voir, par exemple, Richard Münch, « Differentiation, Rationalization, Interpenetration: Emergence of Modern Society », dans Jeffrey C. Alexander et Paul Colomy (dir.), *Differentiation Theory and Social Change*, New York, Columbia Press, 1990, p. 441-464.

¹² Voir Talcott Parsons, *Sociétés : essai sur leur évolution comparée*, Paris, Dunod, 1973 [1966], p. 28.

les néofonctionnalistes¹³, les penseurs de l'École de Francfort¹⁴, les systémiciens de la complexité¹⁵, Alain Touraine¹⁶, Naomi Klein¹⁷, Pierre Bourdieu¹⁸, Ulrich Beck¹⁹, Marshall McLuhan²⁰, Manuel Castells²¹, Philippe Breton²², entre autres. En gros, l'étude de l'ensemble de ces auteurs nous permet d'établir trois thèses générales : 1) le social tend vers une homogénéisation, c'est-à-dire une disparition de la différence, 2) le social tend vers une différenciation, ou un éclatement des repères sociaux, 3) le social est à la fois homogénéisation et différenciation. La recension des écrits dévoile aussi que seule la troisième thèse d'une double

¹³ Jeffrey C. Alexander et Paul Colomy (dir.), *Differentiation Theory and Social Change*, New York, Columbia Press, 1990.

¹⁴ Voir, par exemple, Theodor Adorno et Max Horkheimer, *La dialectique de la raison : fragments philosophiques*, Paris, Gallimard, 1974, p. 281; Herbert Marcuse, *L'homme unidimensionnel : Essai sur l'idéologie de la société industrielle avancée*, Paris, Minuit, 1968, p. 19; Jürgen Habermas, *Théorie de l'agir communicationnel, Tome 2, Pour une critique de la raison fonctionnaliste*, traduit de l'allemand par Jean-Louis Schlegel, Paris, Fayard, 1987, p. 439. Comme plusieurs, Éric Maigret (*Sociologie de la communication et des médias*, Paris, Armand Colin, 2004, p. 17 et p. 72), Michel Lallement (*Lallement, Michel, Histoire des idées sociologiques, tome 2, de Parsons aux contemporains*, Paris, Nathan, 1993, p. 141 et p. 152) et Jean-Marie Vincent (*La théorie critique de l'École de Francfort*, Paris, Galilée, 1976) associent ces auteurs à l'École de Francfort. Mais, il faut comprendre que Marcuse fait autre chose que publier selon les idées de l'École de Francfort et que Habermas critique beaucoup cette école de pensée.

¹⁵ Voir, par exemple, Edgar Morin, *La méthode – 1. La Nature de la Nature*, *op. cit.*, 1977, p. 39.

¹⁶ Alain Touraine, *Pourrions-nous vivre ensemble? Égaux et différents*, Paris, Fayard, 1997.

¹⁷ Naomi Klein, *No Logo. La tyrannie des marques*, traduit de l'anglais par Michel Saint-Germain, Arles, Actes Sud, 2001[2000], p. 29.

¹⁸ Pierre Bourdieu, *La Distinction : critique sociale du jugement*, Paris, Minuit, 1979.

¹⁹ Ulrich Beck, *Pouvoir et contre-pouvoir à l'ère de la mondialisation*, traduit de l'allemand par Aurélie Duthoo, Paris, Flammarion, 2003, p. 8 et p. 16.

²⁰ Marshall McLuhan, *Pour comprendre les médias*, cité par Éric Maigret (*Sociologie de la communication et des médias*, Paris, Armand Colin, 2004, p. 26).

²¹ Manuel Castells, *La société en réseaux : l'ère de l'information*, tome 1, traduit de l'anglais par Philippe Delamare, Paris, Fayard, 1998.

²² Philippe Breton, « La société de l'information : de l'utopie au désenchantement », *Revue européenne des sciences sociales*, tome XL, n° 123, 2002, p. 35-39.

tendance se construit à partir de preuves empiriques. Nous pensons, par exemple, aux travaux de Simon Laflamme et d'Ali Reguigui²³, de Philippe d'Iribarne²⁴, de Jean-Pierre Poulain²⁵, de David Robinson²⁶, de Raymond Breton²⁷ et de Jimmy Thibeault²⁸, pour n'en nommer que quelques-uns. Compte tenu des arguments présentés par ces nombreux auteurs, nous avons constitué un échantillon de 11 020 articles de périodiques parus au Canada et en France en 2005. Notre objectif était de mieux comprendre les phénomènes d'homogénéisation et de différenciation en étudiant les idées politiques qui circulent au sein de périodiques provenant de deux pays. Le tableau 1 présente la distribution de ces articles selon le périodique. Nous avons choisi ces périodiques français et canadiens pour vérifier une thèse homogénéisante présentée par Pierre Bourdieu, tirée d'une entrevue télévisée, intitulé *Sur la télévision. L'emprise du journalisme*²⁹. Dans cet ouvrage, Bourdieu explique que les médias, et surtout la télévision, sont des outils de domination, contrôlés par la classe dominante. Il rappelle que les chaînes, comme ABC ou TF1, sont toutes les propriétés d'entreprises et que, par conséquent, elles sont assujetties à la volonté du marché économique³⁰. Il souligne aussi que les

²³ Simon Laflamme et Ali Reguigui, *Homogénéité et distinction*, Sudbury, Prise de parole, 2003, p. 218.

²⁴ Philippe d'Iribarne, *La logique de l'honneur : gestion des entreprises et traditions nationales*, Paris, Seuil, 1989, et Philippe d'Iribarne, Alain Henry, Jean-Pierre Segal et Sylvie Chevrier, *Culture et mondialisation. Gérer par-delà les frontières*, Paris, Seuil, 2002 [1998], p. 7.

²⁵ Jean-Pierre Poulain, *Sociologies de l'alimentation*, Paris, Presses Universitaires de France, 2003 [2002].

²⁶ David Robinson, *Muslim Societies in African History*, Cambridge (UK), Cambridge University Press, 2004.

²⁷ Raymond Breton, « Introduction: Rethinking the Impact of Globalization Processes - Differentiation as well as Convergence », Raymond Breton et Jeffrey G. Reitz (dir.), *Globalization and Society. Process of Differentiation Examined*, Westport Connecticut, London, Praeger, 2003, p. 1-10.

²⁸ Jimmy Thibeault, *Des identités mouvantes. Se définir dans le contexte de la mondialisation*, Montréal, Nota bene, coll. « Terre américaine », 2015.

²⁹ Pierre Bourdieu, *Sur la télévision/L'emprise du journalisme*, Paris, Raison d'agir, 1996.

³⁰ Voir, par exemple, *ibid.*, p. 14.

grands hebdomadaires imposent leur contenu aux plus petits³¹. Ces affirmations en tête, nous avons sélectionné deux hebdomadaires français à grand tirage, un quotidien français à grand tirage et un quotidien français à plus petit tirage (ou à tirage régional plus restreint). Ensuite, nous avons fait une sélection semblable avec des périodiques canadiens dans une optique de double vérification.

Tableau 1

Distribution des articles selon le périodique

Périodique canadien	N	Périodique français	N
<i>Time</i>	295	<i>Le Figaro</i>	1517
<i>Macleans</i>	359	<i>Le Nouvel Observateur</i>	600
<i>The Globe and Mail</i>	1 005	<i>Le Monde</i>	909
<i>The Sudbury Star</i>	573	<i>Ouest France</i>	5 386
Total	2 232	Total	8 412
Moyenne	558	Moyenne	2 103
Écart-type	320,83	Écart-type	2 221,57

Cette sélection n'est pas irréprochable. Par exemple, on pourrait dire que nous avons forcé l'hétérogénéité entre les périodiques qui ne sont pas du même ordre; qu'un périodique local n'aura certainement pas le même contenu que le périodique national ou international. Nous étions conscients de ce danger lors de la constitution de notre échantillon. Si la thèse de l'homogénéisation, dans sa forme la plus extrême, était vraie, il n'y aurait pas de différence *a priori* entre les périodiques. De même, si la thèse de l'hétérogénéisation, dans sa forme la plus extrême, était vraie, il n'y aurait pas de ressemblance *a priori* entre les périodiques. Dans un cas, comme dans l'autre, il serait impossible d'insinuer quelque biais. Certes, la construction même d'un échantillon pour vérifier ces thèses révèle déjà l'absence d'absolus : il ne peut exister deux périodiques entièrement semblables ou entièrement différents *a priori*. Si nous avons poursuivi notre étude tout en étant conscient de ce genre de biais, c'était tout justement pour mieux voir de quelles autres façons cette ressemblance et cette différence se manifestaient.

³¹ *Ibid.*, p. 86-87.

Il est aussi vrai que les phénomènes d'homogénéisation et d'hétérogénéisation peuvent être étudiés en scrutant les pratiques culturelles, linguistiques ou religieuses, ou même en examinant les habitudes alimentaires. La société nous offre plusieurs objets légitimes en lien avec ce sujet. Toutefois, nous avons cru qu'une étude empirique des idées politiques qui circulent dans les périodiques ajoutait quelques valeurs à cette réflexion compte tenu de l'importance de ces idées et de ces périodiques au sein des sociétés contemporaines. Élaborons un peu plus longuement cette idée.

Dans *le modèle de la circulation des biens, des idées et des personnes*, Laflamme explique qu'il est possible de catégoriser des idées à partir de la problématique politique³² qui anime une région, c'est-à-dire à partir des idées qui la constituent et qui en animent la sphère politique. C'est dans cette logique qu'il peut y avoir une relation trialectique entre les biens, les idées et les personnes. Il s'agit tout simplement de penser à la relation qu'entretient une opinion politique avec l'économie d'une région et avec sa population. Or, ce ne sont pas toutes les idées qui sont étudiées, mais bien celles qui sont suffisamment récurrentes³³, et forcément évidentes, pour que les politiciens ne puissent pas ne pas en tenir compte dans leurs interventions, dans leur rapport au pouvoir. Le but de l'exercice de repérage n'est pas d'identifier

³² « La problématique politique est le lieu où la société conçoit son intervention sur elle-même, où les organismes politiques proposent un avenir collectif. Contexte intériorisé, elle est de la même façon le point de jonction de la tradition commune et de la contingence. C'est là que la société, par des initiatives, s'organise en fonction des accidents ou bien qui surgissent en elle ou bien dans lesquels elle est, soit partiellement soit entièrement, mise en cause, victime ou responsable. C'est là qu'est éprouvé l'événement, événement dont l'impact sur les phénomènes de persuasion politique peut être déterminant, pouvant régler le sort d'un persuadeur, pouvant transformer ou même engendrer une problématique politique » (*Contribution à la critique de la persuasion politique*, Québec/Sudbury, Presse de l'Université du Québec/ Université Laurentienne, 1987, p. 23-24).

³³ La récurrence nécessaire pour qu'une idée fasse partie de la problématique politique élimine dès le départ la possibilité d'observer une hétérogénéité éclatante. Ce problème est inhérent à tout objet social puisque dès que cet objet devient social, même reconnaissable, c'est qu'il ressemble à autre chose.

la totalité de ce qui peut être dit à l'intérieur d'une région donnée – il est certain, alors, que les médias seraient réducteurs puisqu'il y a des événements qui leur échappent –; il est de cerner la problématique politique, c'est-à-dire le discours qui ne peut pas ne pas être partagé par un nombre, même marginal, d'individus dans une région donnée et sans que cela intervienne sur les valeurs qui animent la population qui l'habite. En correspondance avec ces pensées, nous avons cru que l'étude des idées associées à la problématique politique pouvait non seulement produire une recherche intéressante pour mieux comprendre la circulation des biens, des idées et des personnes, mais qu'elle pouvait aussi éclairer les phénomènes d'homogénéisation et d'hétérogénéisation.

Afin de relever les idées qui sont en lien avec la problématique politique, nous avons choisi d'étudier les journaux. Face à ce choix, la question suivante pouvait se poser : pouvons-nous nous contenter des journaux pour comprendre les idées qui composent la problématique politique d'une région? En considérant certaines preuves, par contre, il ne devrait pas y avoir de doute sur le fait que l'étude de périodiques (et donc de journaux) permet de comprendre les idées qui composent la problématique politique. Par exemple, les gens considèrent plus les journaux que la télévision comme source fiable d'information³⁴. Les journaux doivent nécessairement influencer autant, sinon plus, la sphère politique que la télévision, de par le fait qu'ils sont corrélés avec l'intérêt public d'une région³⁵, et de par le fait qu'ils diffusent des informations qui sont plus faciles à retenir que celles qui sont diffusées

³⁴ Voir, par exemple, Elizabeth Taylor (*A Cross-Media Study of Audience Choice : The Influence of Traits, Needs, and Attitudes on Individual Selection of « media repertoires »*, Unpublished Dissertation, Philosophy, The University of Texas at Austin, 1999, 253 p.) et Dane S. Claussen (« Cognitive Dissonance, Media Illiteracy, and Public Opinion on News Media », *American Behavioral Scientist*, vol. 48, n° 2, Octobre 2004, p. 212-218) qui montrent que les journaux sont considérés comme la source d'information la plus importante des personnes qu'ils ont recensées.

³⁵ Kevin G. Barnhurst, and Ellen Wartella, « Newspapers and Citizenship: Young Adults' Subjective Experience of Newspapers », *Critical Studies in Mass Communication*, vol. 8, n° 2, 1991, p. 195–209.

par la télévision ou par la radio³⁶. Joseph P. Bernt et ses collaborateurs ont montré que les rédacteurs arrivent à correctement anticiper la volonté et les intérêts de leurs lecteurs, ce qui va à l'encontre de l'opinion que les rédacteurs sont des personnes autonomes qui agissent à leur propre gré indépendamment de la volonté de leurs lecteurs³⁷. Cette observation va aussi à l'encontre de l'idée qui veut que les entreprises manipulent ce qui est diffusé par les journaux indépendamment de la volonté des lecteurs³⁸. Par ailleurs, les journaux comportent plus d'informations que les médias comme la télévision et la radio ne parviennent à en diffuser; or, si une information se rapporte à la problématique politique, elle apparaîtra fort probablement dans les journaux. C'est en vertu de tels liens entre les idées, les personnes et les périodiques que nous avons cru que les journaux sont suffisants pour étudier la problématique politique d'une région.

Ajoutons à ces raisons théoriques celles qui sont plutôt d'ordre logistique. Nous avons rappelé, dans un travail antérieur³⁹, que Laflamme suggère le journal pour dégager les idées qui circulent dans une région puisque, en raison de son format écrit, il est plus facile à traiter que ne le sont les contenus des médias comme la télévision ou la radio. Nous avons aussi fait valoir qu'il est plus facile d'identifier les thèmes qui sont pertinents pour une région en scrutant les journaux qu'Internet, puisque ce dernier diffuse

³⁶ Melvin L. Defleur, Lucinda Davenport et Mary D. Cronin, « Audience Recall of News Stories Presented by Newspaper, Computer, Television and Radio », *Journalism Quarterly*, vol. 69, n° 4, 1992, p. 1010-1022.

³⁷ Joseph P. Bernt *et al.*, « How Well Can Editors Predict Reader Interest in News? », *Newspaper Research Journal*, vol. 21, n° 2, 2000, p. 2-10.

³⁸ Éric Maigret explique clairement que la thèse qui veut que les médias déterminent le restant de la société est une thèse simpliste qui arrive mal à saisir la relation complexe qu'entretiennent les médias avec leur société (*Sociologie de la communication et des médias*, Paris, Armand Colin, 2004, p. 164-165).

³⁹ Roger Gervais, *Les effets de la mondialisation sur la relation entre les idées et les personnes : une comparaison entre les régions toulousaine et sudburoise*, mémoire de maîtrise en sociologie, Université Laurentienne, Sudbury, 2005 et, Roger Gervais, « La mondialisation : vers une compréhension duelle de l'homogénéisation et de la différenciation », *Nouvelles perspectives en sciences sociales*, vol. 2, n° 1, 2006, p. 69-100.

beaucoup d'informations disparates, dont bon nombre ne sont pas nécessairement fiables. Ainsi, pour des raisons purement techniques, les journaux représentent le média le plus facile à traiter si l'on souhaite étudier les idées⁴⁰.

À n'en pas douter, il y a plusieurs façons d'étudier les phénomènes d'homogénéisation et d'hétérogénéisation. Cependant, pour les raisons présentées ci-dessus, il peut être propice de travailler à partir de la problématique politique, telle qu'on peut la résumer à partir de messages diffusés par les périodiques. Un échantillonnage de ces messages devrait permettre d'aborder efficacement le double fait de l'homogénéisation et de l'hétérogénéisation.

Afin de satisfaire aux besoins de leurs différents publics, de rallier et de captiver un auditoire qu'ils souhaitent sans cesse croissant, et pour faire face à la concurrence effrénée qu'ils se livrent entre eux, les médias se font un devoir de parler de tout. Ils traitent de tous les sujets qui font la nouvelle et donnent sur chacun le plus de renseignements possible, tout en tenant compte du mode de diffusion et du style qui leur est propre⁴¹.

Cette citation annonce assez bien une difficulté liée à l'analyse des contenus médiatiques : la diffusion d'un vaste nombre d'informations. Même si le format écrit et la méthode de distribution des journaux rendent plus accessible l'information, comment se doter d'un échantillon qui soit tout à la fois abordable et adéquat? Un chercheur pourrait choisir tout simplement de sélectionner les informations qui figurent sur la page couverture⁴². Il pourrait cibler un seul thème et étudier son évolution⁴³. Il pourrait aussi scruter les publicités, l'emplacement des textes ou retenir exclusivement le titre des articles. Ces techniques réduisent l'information à appréhender en même temps qu'elles offrent des corpus intéressants. Toutefois, une étude qui met en son centre la problématique politique ne peut pas recourir à une telle

⁴⁰ *Ibid.*

⁴¹ Lise Chartier, *Mesurer l'insaisissable : méthode d'analyse du discours de presse*, Sainte-Foy, Presses de l'Université du Québec, 2003, p. 29.

⁴² *Ibid.*, p. 22

⁴³ Cette méthode est souvent employée par les chercheurs pour faire une étude, par exemple, sur la démocratie, le sexisme ou le racisme que l'on retrouve dans les journaux.

méthode puisque chacune d'elles élimine, à sa façon, une partie de ce qui compose la problématique politique. Par exemple, la sélection des articles situés sur certaines pages éliminerait des articles qui ne sont pas considérés captivants⁴⁴; le choix de thèmes *a priori* éliminerait le caractère évolutif et dynamique des idées liées à la problématique politique⁴⁵; concentrer l'analyse sur la publicité forcerait à éliminer des propos essentiels d'intérêt politique; et la sélection exclusive de titres éliminerait le contexte qui les justifie. Une recherche sur la problématique politique ne doit donc pas discriminer selon les procédés signalés ci-dessus.

Dans l'optique de rendre notre échantillon abordable et pertinent, il est possible de procéder à une première sélection d'articles en excluant ceux qui se rapportent à des thématiques culturelles, sportives, scientifiques⁴⁶, médicales⁴⁷, technologiques⁴⁸, ou à des faits divers. Sur ces thématiques, ne sont retenus que les articles qui s'inscrivent dans le discours politique. Pour qu'un article, qui serait en principe exclu à cause de son propos *a priori* non politique, soit retenu, il doit faire partie du discours politique – nous notons par exemple les moments où une maladie (le SRAS ou la maladie de la vache folle) ou une nouvelle technologie (celles de la reproduction) oblige les politiciens à en tenir compte lors de leurs décisions administratives. Dans cet esprit, nous avons donc écarté des articles qui ne touchent pas à la problématique politique.

⁴⁴ Lors d'une catastrophe, comme l'explosion d'une usine chimique à Toulouse ou l'effondrement du viaduc à Sudbury, les articles qui rapportent des événements moins tumultueux, comme l'approbation du Conseil municipal pour le financement d'une piscine publique, pourraient fort bien se situer sur les pages arrière du périodique.

⁴⁵ Ce qui est d'intérêt un jour, peut ne pas l'être trois mois plus tard. De plus, ce qui est d'intérêt un jour, peut ne pas l'être pendant les trois prochains mois et le redevenir ensuite. Or, le choix de thèmes *a priori* limite la prise en considération de ces changements.

⁴⁶ Par exemple les articles qui présentent les nouvelles découvertes en sciences.

⁴⁷ Par exemple les articles qui présentent les nouvelles découvertes en médecine ou les risques pour la santé.

⁴⁸ Par exemple les articles qui informent les consommateurs des nouveaux produits qui se trouvent sur le marché.

Ensuite, la sélection de dates d'impressions (par exemple, traiter seulement les publications autour du 10 de chaque mois) peut aussi aider à réduire l'information à traiter. Nous pensons que les propos qui sont structurellement de l'ordre de la problématique politique s'observeront mieux en étudiant des dates éloignées les unes des autres, parce que la récurrence de ces propos dans le temps les distinguera des idées éphémères. Le danger avec cette méthode, par contre, est semblable à celui qui est associé à la sélection *a priori* de thèmes – nous pouvons introduire un biais en sélectionnant certaines dates qui, pour une raison ou une autre, ne sont pas représentatives d'une autre journée et il se peut qu'un journal ne rapporte pas les mêmes événements que les autres. Nous ne serons répréhensibles sur ce point que si nous n'observons aucune ressemblance ou aucune différence entre les périodiques. Dès qu'il est possible de faire ces types d'observation, nous aurons la preuve que notre échantillon est suffisamment grand pour nous donner accès aux informations nécessaires à la vérification de nos hypothèses. Nous avons donc constitué l'échantillon en nous concentrant sur des dates.

La mobilisation en France contre la Constitution européenne ou contre le Contrat de première embauche (aussi connu sous l'abréviation CPE) ne sont que deux exemples qui témoignent de l'engagement politique des Français. À première vue, de tels événements sont beaucoup moins fréquents chez les Canadiens, qui, lorsqu'on les compare aux Français, paraissent beaucoup moins emportés sur le plan politique. Cette observation sommaire en tête, nous avons choisi de comparer les problématiques politiques qui circulent en France et au Canada. L'opération permettait de vérifier le degré d'homogénéité et d'hétérogénéité, tant à l'intérieur des deux pays, qu'entre les deux pays, et, éventuellement, de critiquer l'impression première. Il était possible, par exemple, que le positionnement politique des Français traduise une très grande hétérogénéité – l'opinion étant pluralisée par la diversité des groupes d'intérêt – et que le calme canadien sous-tende une grande homogénéité étant donné que peu de personnes s'affirment. Le contraire était tout aussi

possible. La grande mobilisation politique en France supposait peut-être une plus grande homogénéité puisque les Français peuvent s'associer à des groupes de pression et la tranquillité des Canadiens peut imposer une plus grande hétérogénéité de par le fait que les manifestants ne peuvent que s'affirmer en tant qu'individus. Quel que soit le scénario que nous dessinions pour expliquer le contraste entre la France et le Canada, nous avons cru que la comparaison était la bienvenue dans une analyse de l'homogénéité et de l'hétérogénéité.

Il nous est impossible de présenter l'ensemble de nos résultats dans cet article et nos conclusions liées aux thèses de l'homogénéisation et de la différenciation ne font pas l'objet de la réflexion en cours. Notre objectif, comme nous l'avons indiqué dans l'introduction, est d'étudier l'impact qu'a l'analyse de données textuelles informatisées sur deux préoccupations méthodologiques de départ liées au sens. Nos résultats seront donc affinés en fonction de ce questionnement et serviront à mieux comprendre ce qui se passe dans une analyse qui repose sur de grands échantillons.

2. Le danger de perdre le sens des mots

La « décontextualisation » des mots soulève une inquiétude chez les chercheurs, et ce, depuis les débuts de l'analyse de discours. Les extraits que nous avons ciblés font état, à cet égard, de la pertinence de la préoccupation de Maingueneau :

L'analyse factorielle, précisons-le bien, n'est qu'une méthode de traitement des données, et nullement de constitution des données : autrement dit, elle est passible des mêmes critiques que toute entreprise opérant sur les mots à la surface du discours; en outre, c'est d'une statistique « hors contexte » qu'il s'agit, avec les inconvénients que cela implique⁴⁹.

⁴⁹ Dominique Maingueneau, *Initiation aux méthodes de l'analyse du discours*, op. cit., p. 36. Nous reconnaissons que Maingueneau a depuis publié d'autres ouvrages sur l'analyse de discours. Si nous nous référons à ce classique, c'est en raison de la qualité de la démonstration des dangers liés à l'analyse lexicométrique. Quelques références additionnelles provenant de cet auteur : Dominique Maingueneau, « Présentation », *Langages*, n° 117, 1995, p. 5-11; Dominique Maingueneau, *L'analyse de discours*, Paris, Hachette, 1997 [1991]; Dominique Maingueneau, « L'Analyse du discours », dans Patrick Charaudeau

[...] L'analyse factorielle porte sur des textes, et non sur des positions [...]. Si nous avons seulement trois textes à analyser, l'un disant : *je suis partisan de la révision de la constitution*, le second : *je suis adversaire de la révision de la constitution* et le troisième ne soufflant mot du problème, l'analyse factorielle dirait, avec raison, que les deux premiers textes ont en commun *révision* et *constitution*, et qu'ils s'opposent ensemble au troisième.

[...] [L]'analyse factorielle ne donne [donc] pas des résultats « immédiatement parlants », mais n'est qu'une sorte de microscope obligeant à une lecture plus rigoureuse des textes d'une part, et de l'autre contribuant à valider les interprétations que cette lecture suggère⁵⁰.

[...] [L]'analyse lexicométrique est obligée d'annuler en quelque sorte de nombreux « effet de sens ». La mise en relation d'une étude en contexte, et d'une étude hors contexte peut toutefois être très féconde, les approches se corrigeant réciproquement. [Or] la lexicométrie ne se veut pas une approche exclusive et son grand intérêt est de constituer un ensemble de modèles cohérent, dont on peut discuter la validité avec précision, ensemble qui doit être confronté aux conclusions d'autres approches et ne peut jamais suffire à fonder à lui seul la pertinence des conclusions⁵¹.

Si cette observation est toujours juste, les logiciels contemporains les plus performants permettent à la fois des analyses lexicales et des analyses contextuelles, comme le demande Maingueneau. Par conséquent, les logiciels comme SPAD réduisent les dangers liés à une analyse statistique sans « sens ». Le chercheur peut donc, tout en demeurant maître de son analyse⁵², repérer rapidement des tendances à l'intérieur de corpus énormes; le chaos initial peut ainsi prendre quelque forme. Les logiciels permettent ensuite au chercheur de saisir le sens de ces objets et de ces tendances, puisque le logiciel, lui, ne perd plus le contenant dans lequel l'information a été initialement organisée.

et Dominique Maingueneau (dir.), *Dictionnaire d'analyse du discours*, Paris, Seuil, 2002, p. 41-45; ou Dominique Maingueneau, « Que cherchent les analystes du discours? Argumentation et analyse du discours », *Argumentation et analyse de discours, la revue électronique du groupe ADARR*, n° 9, 2012.

⁵⁰ *Ibid.*, p. 40.

⁵¹ *Ibid.*, p. 44.

⁵² Cette idée s'inspire de l'*Appel à communications de la revue : Nouvelles perspectives en sciences sociales. Pour le colloque de son 10^e anniversaire sur le thème Pensée systémique complexe et relationnelle*, op. cit.

3. Un problème technique corrigé par la technique : comment maintenir le sens lors de l'analyse de données textuelles informatisées

Dans le cadre de nos recherches, le logiciel SPAD a été utilisé non seulement pour produire des analyses factorielles, mais aussi pour dégager le sens de concordance de mots. À titre d'exemple, nous avons été en mesure d'établir que le mot *défense*, tel qu'utilisé dans le *Macleans*, avait une double connotation, la première étant militaire, la seconde, juridique.

Examinons de plus près SPAD et ses outils d'analyse. Après chaque analyse factorielle, des graphiques et des tableaux comme ceux que nous avons présentés ci-dessous (graphiques 1 et 2; tableaux 2 et 3) peuvent être construits. Ces informations ainsi dépeintes, nous pouvons conclure qu'il y a une grande concentration du corpus lexical associé aux périodiques français et que cette ressemblance, ou homogénéité, ne se traduit pas par l'absence de distinctions. Par exemple, le premier facteur isole entièrement le quotidien local *Ouest France* des autres périodiques. De plus, le deuxième facteur sépare *Le Monde* et *Le Figaro* du *Nouvel Observateur* (voir graphique 2⁵³). La combinaison des facteurs 1

⁵³ **Valeurs propres** : Le premier facteur extrait environ 60 % de l'inertie totale, le deuxième, environ 19 % et le troisième, environ 18 %. 3 % de l'inertie totale ne sont extraits par aucun des trois facteurs.

Coordonnées des fréquences actives : Le premier axe sépare *Ouest France* (-0,39) des autres périodiques (*Le Nouvel Observateur* : 0,34; *Le Monde* : 0,27; *Le Figaro* : 0,28), le deuxième sépare *Le Nouvel Observateur* du *Monde* et du *Figaro*, tandis qu'*Ouest France* est sur l'axe. On voit, pour le troisième axe, une distance importante entre *Le Figaro* (-0,24) et *Le Monde* (0,34) avec *Le Nouvel Observateur* (0,04) et *Ouest France* (0,00) entre les deux.

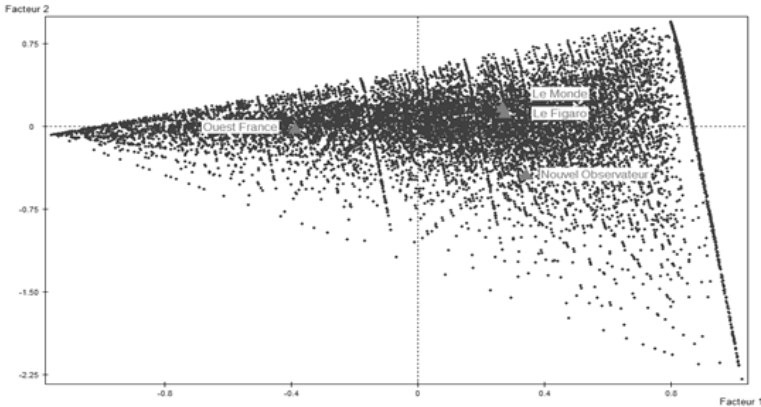
Contributions des fréquences actives : Si toutes les fréquences avaient la même contribution, celle-ci devrait osciller autour de 25 % (100 %/4). Les fréquences avec des contributions supérieures à 25 % ont une influence supérieure à la valeur théorique. L'analyse du corpus lexical pour les périodiques français montre qu'*Ouest France* explique presque 50 % de la totalité de l'inertie de l'axe 1, *Le Nouvel Observateur* explique 72 % de la totalité de l'inertie de l'axe 2 et *Le Monde* et *Le Figaro* partagent la contribution de l'axe 3.

Cosinus carrés des fréquences actives : *Ouest France* coïncide avec le premier axe (cosinus pour axe 1 : 0,95), *Le Nouvel Observateur* se dessine largement

et 3 ne modifie pas l'isolement d'*Ouest France*, mais révèle que *Le Monde* et *Le Figaro* ne sont pas aussi semblables qu'on pouvait le croire (graphique 2⁵⁴).

Graphique 1

Comparaison du corpus lexical des périodiques français
Facteurs 1 et 2



comme deuxième (cosinus pour axe 2 : 0,60), tandis que *Le Monde* (cosinus : 0,47) et *Le Figaro* (0,37) correspondent faiblement à l'axe 3.

Poids relatif : *Ouest France* a le poids le plus important (poids relatif 42,39 %). *Le Figaro* se place en deuxième (poids relatif : 25,04 %). *Le Monde* vient ensuite (poids relatif : 18,19 %), suivi en dernier par *Le Nouvel Observateur* (poids relatif : 13,57 %).

⁵⁴ Voir la note infrapaginale précédente pour les explications statistiques.

Graphique 2

Comparaison du corpus lexical des périodiques français
Facteurs 1 et 3

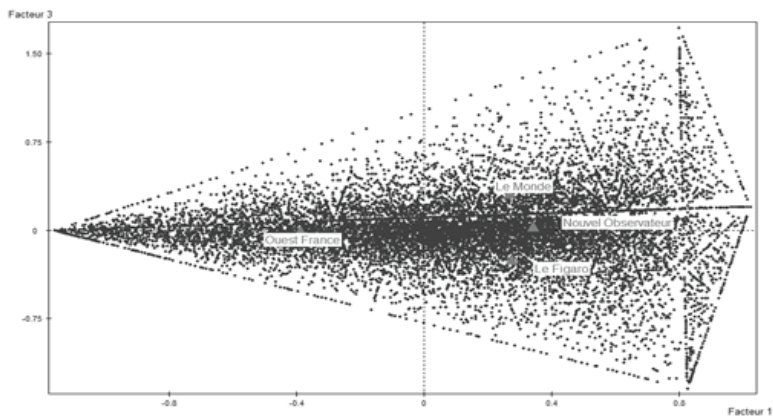


Tableau 2

Vocabulaire spécifique pour Ouest France				
+ = proximité significative entre vocabulaire et périodique - = distance significative entre vocabulaire et périodique				
	<i>Le Figaro</i>	<i>Le Nouvel Observateur</i>	<i>Le Monde</i>	<i>Ouest France</i>
Aménagement	-	-	-	+
Association	-	-	-	+
Centre	-	-	-	+
Commune(s)	-	-	-	+
Conseil	-	-	-	+
École	-	-	-	+
Maire	-	-	-	+
Mairie	-	-	-	+
Municipal	-	-	-	+
Rue	-	-	-	+
Salle	-	-	-	+
Travaux	-	-	-	+
Élus	-	-	-	+
Lotissement	-	-	-	+
Pompiers	-	-	-	+
Élève	-	-	-	+

Les tableaux 2 et 3 rendent clair l'éloignement d'*Ouest France* : les autres périodiques, sur le plan du vocabulaire spécifique, sont presque toujours à une distance notable d'*Ouest France* (voir tableau 2), et ce journal-ci est constamment à bonne distance des autres périodiques (voir tableau 3). C'est un lexique local (*aménagement, communes, conseil, maire, mairie, municipal, lotissement et pompiers*) et communautaire (*association, centre, école, salle, travaux, élève*) qui rend *Ouest France* particulier. C'est un vocabulaire institutionnel politique (*France, Paris, gouvernement et ministre*) qui rend *Le Figaro* unique; un recueil de mots associés à la chose et à des personnages politiques (*Chirac, politique* et *Sarkozy*) qui distingue *Le Nouvel Observateur*, et un regroupement de mots américano-économique (*Américain, États, Monde et dollars*) qui différencie *Le Monde*. Le partage des mots *millions* et *États-Unis* explique en partie la proximité du *Figaro* et du *Monde* dans le graphique 1.

Tableau 3

Vocabulaire non spécifique pour Ouest France				
+ = proximité significative entre vocabulaire et périodique - = distance significative entre vocabulaire et périodique				
	<i>Le Figaro</i>	<i>Le Nouvel Observateur</i>	<i>Le Monde</i>	<i>Ouest France</i>
Million(s)	+		+	-
États-Unis	+		+	-
Américain(e)			+	-
État			+	-
Monde			+	-
Dollars			+	-
Euros			+	-
Chirac		+		-
Politique		+		-
Sarkozy		+		-
France	+			-
Paris	+			-
Gouvernement	+			-
Ministre	+			-
Milliards	+			-

La préoccupation de la perte du sens est toujours très légitime si l'on part de ces informations seulement. Comment peut-on être certain que « Chirac » correspond vraisemblablement aux

personnages politiques de la période étudiée sans avoir plus d'information? Dans le but de corriger cette lacune, SPAD permet la construction de tableaux comme celui-ci :

Tableau 4

Exemples de concordances pour le mot *Chirac*

au président de la République, Jacques	Chirac	, au premier ministre, Jean-Pierre Raffarin,	Article 33 LF
Le gouvernement Rouge-Vert de Berlin et, à Paris, le président Jacques	Chirac	ont par opposition milité	Article 151 LF
que Jacques	Chirac	soit, ou non, candidat à sa propre succession	Article 356 LF
Entre le oui à la Constitution et le non à Jacques	Chirac	, la campagne de François Hollande n'était déjà pas facile.	Article 555 LF

Ces tableaux sont très utiles pour la personne qui veut assurer le sens des mots après avoir saisi le lexique spécifique de chaque périodique. Quand les mots *New* et *Orleans* apparaissent dans le *Times*, par exemple, on court le risque de joindre les deux mots, de penser que ce périodique s'intéresse à des événements spécifiques à la ville de *New Orleans*. L'analyse de concordances des mots a facilement redirigé cette pensée : si *New* et *Orleans* sont souvent évoqués, le premier mot est aussi en lien avec la ville de New York.

Tableau 5

Exemples de concordances pour le mot *New*

The city council of	New	York last month released the results of a local study	Article 10 TI
he took a job with TIME as a copyboy while earning a degree at	New	York University	Article 59 TI
It may be weeks before the lights come back on and months before	New	Orleans is mopped out	Article 219 TI
When I told him about the rising floodwaters throughout	New	Orleans, the rain coming in through the torn Superdome roof	Article 228 TI

À lui seul, le mot *Pantano*, retenu dans le vocabulaire spécifique du *Times*, ne voulait rien dire. Grâce aux tableaux de concordances, nous avons rapidement découvert qu'Ilaro *Pantano* était un soldat américain. Toujours dans le *Times*, le terme *troops* fait

parfois référence aux soldats canadiens ou à ceux de l'OTAN en Afghanistan, parfois aux soldats américains en Afghanistan et en Iraq. Cette distinction peut se comprendre seulement grâce aux tableaux de concordance. Il en va de même du lien entre les termes *water* et *tsunami* qui n'était pas clair sans l'analyse de concordances. Lorsque le *Time* emploie les mots *Administration* et *House*, c'est surtout en référence à l'administration de la Maison-Blanche, comme dans *Bush Administration*, *Nixon Administration* ou *White House*. Le mot *Bush* se réfère principalement au président *George W. Bush*, mais parfois aussi à son père, ancien président *George Bush*, une nuance qui est peu visible en l'absence des tableaux de concordances.

Il y a aussi le cas du mot « Amérique » qui est employé par *Ouest France* deux fois sur vingt-deux pour désigner les États-Unis⁵⁵, alors que *Le Figaro* s'adonne à ce glissement beaucoup plus souvent.

Nous avons ici un bel exemple des outils de SPAD qui permettent de produire des analyses factorielles et, par ailleurs, de prendre en considération la signification des mots en les renvoyant aux concordances. Par conséquent, on ne perd pas la complexité des idées par le seul fait qu'on manipule un corpus imposant et qu'on recourt à l'analyse statistique.

4. La détermination du cadre historique et de l'origine du document

En l'absence de logiciels permettant l'analyse d'un grand corpus textuel, le chercheur doit nécessairement s'armer de plusieurs informations avant même de plonger dans les textes. « De fait, le commentaire de document exige du chercheur une attention et une vigilance de tous les instants, une bonne connaissance de l'origine et du contexte de production du document ainsi qu'une grande capacité d'interrogation et d'imagination », nous dit Jocelyn Létourneau dans son manuel méthodologique intitulé

⁵⁵ Les autres fois qu'*Ouest France* utilise le terme Amérique, c'est pour désigner l'Amérique latine, l'Amérique du Nord (Québec), l'Amérique du Sud ou l'Amérique centrale.

le *Coffre à outils du chercheur débutant*⁵⁶. Plus loin dans ce même texte, ce même auteur explique que la deuxième étape du commentaire est « la détermination du cadre historique et de l'origine du document⁵⁷ ».

Dans cette section, il souligne comment le chercheur établira cette « bonne connaissance de l'origine et du contexte de production du document ». Avant même d'analyser le document, il faut connaître l'auteur du document, sa notoriété et son « envergure sociale au moment de la rédaction du document »⁵⁸. Il faut juger de l'importance et la crédibilité du témoignage de ce même auteur, de son objectif avoué et de son objectif non avoué, ainsi que du possible « intérêt supplémentaire ou particulier » lié au fait que le document analysé a été produit par cet individu. L'analyste doit aussi déterminer le cadre historique originel du document afin de saisir les particularités de la forme, du contenu, du ton, de la présentation, du discours, de l'organisation générale du document. Paul Sabourin et Vincent Ross avant lui expliquent qu'une préparation méticuleuse assure que le chercheur connaît l'énonciateur, les destinataires visés et les particularités de l'écriture du document⁵⁹.

⁵⁶ Jocelyn Létourneau, *Le coffre à outil du chercheur débutant. Guide d'initiation au travail intellectuel*, Montréal, Boréal, 2006, p. 77.

⁵⁷ *Ibid.*, p. 80.

⁵⁸ *Ibid.*, p. 80.

⁵⁹ Paul Sabourin, « Chapitre 15 : L'analyse de contenu », *op. cit.*, p. 371. On trouve un autre exemple d'un texte qui souligne l'importance d'étudier le cadre historique et de l'origine du document, mais qui en parle en relation avec les périodiques chez Micheline Cambron (dir.), *Le journal Le Canadien : littérature, espace public et utopie, 1836-1845*, Saint-Laurent (QC), Fides, coll. « Nouvelles études québécoises », 1999 : « Avant d'entreprendre l'analyse du discours dans un journal [...], il importe en conséquence de préciser ce que nous savons du dispositif éditorial qui l'informe, de la facture en tant qu'elle détermine la lecture et l'intelligence » (p. 77-78). Ou encore, chez Lise Chartier : « Afin de satisfaire aux besoins de leurs différents publics, de rallier et de captiver un auditoire qu'ils souhaitent sans cesse croissant, et pour faire face à la concurrence effrénée qu'ils se livrent entre eux, les médias se font un devoir de parler de tout. Ils traitent de tous les sujets qui font la nouvelle et donnent sur chacun le plus de renseignements possible, tout en tenant compte du mode de diffusion et du style qui leur est propre » (Lise Chartier, *Mesurer l'insaisissable : méthode d'analyse du discours de presse*, Sainte-Foy, Presses de l'Université du Québec, 2003, p. 29).

On peut admettre cette logique dans les cas où le chercheur se trouve devant un nombre limité de textes. Le fait de prendre en considération l'ensemble de ces informations peut alors enrichir la lecture des documents. On peut concevoir qu'un auteur s'exprime différemment s'il est journaliste, éditeur, chroniqueur, politicien, expert-conseil ou parent militant. Les propos de l'expert-conseil auront une crédibilité différente que ceux provenant des autres groupes identifiés. L'article qui se trouve dans un périodique renommée ou scientifique ne sera pas de même nature qu'une correspondance entre deux amis, deux collègues de travail ou deux avocats. L'ensemble de ces détails, s'ils sont connus à l'avance, outille l'analyste et lui permet de bonifier son interprétation des propos. L'analyse historiographique le montre depuis très longtemps. Des personnages historiques, comme Étienne Brûlé, peuvent passer d'inconnu, à traître, à héros en raison du cadre historique à l'intérieur duquel l'interprétation est faite⁶⁰.

5. La croisée des chemins

Que se produit-il quand on applique cette consigne méthodologique de la prise en considération du cadre historique et de l'origine de l'énonciation à l'analyse de journaux? Que faire quand le nombre d'articles analysés dépasse les 500, les 700, voire les 1 000? On peut difficilement, dans des situations de ce genre, recueillir l'ensemble des informations sur tous les auteurs, qu'il s'agisse de leur notoriété, de leur « envergure sociale au moment de la rédaction », de leur importance, de leur crédibilité en tant que journaliste individuel ou de leurs « objectif avoué et objectif non avoué », pour ne mentionner que quelques facteurs qui ont été mentionnés ci-dessus par des méthodologues de l'analyse de document. Notre échantillon contenait des centaines d'auteurs et la recherche de ces informations historiques et contextuelles aurait obligé à un travail extraordinaire, voire infini. Pour cette

⁶⁰ Stéphanie St-Pierre, « Étienne Brûlé : la création d'un personnage », *Revue du Nouvel-Ontario*, n° 29, 2004, p. 5-44. Voir aussi Colin M. Coates et Cecilia Morgan, *Heroines and History: Representations of Madeleine de Verchères and Laura Secord*, Toronto, University of Toronto Press, 2002.

raison, nous avons considéré nos périodiques, au point de départ, comme des véhicules de la circulation des idées, et non comme des lieux de diffusion des opinions d'individus spécifiques. Comme de nombreux chercheurs en sciences politiques et en analyse de discours⁶¹ l'ont déjà fait, le périodique a été traité comme un répondant unique. En dépit du fait qu'un média ne peut exister sans la contribution de ses auteurs individuels (et parfois collectifs), notre objet d'étude n'imposait pas qu'on obtienne sur ces auteurs des informations *a priori*. Nous n'avons pas non plus, dans le cadre de notre recherche, tenté de connaître la mission ou la description de nos périodiques. Seules les informations liées à leur tirage ont été retenues, à titre d'indicateur de l'importance de la circulation des informations dans une région donnée. Regrettons-nous notre décision? Aucunement. Aurait-il été nécessaire d'obtenir ces informations à l'avance? En comparant quelques-unes de nos conclusions aux descriptions de nos périodiques, nous verrons que ces informations n'auraient rien ajouté à nos analyses.

*Maclean's is Canada's only national weekly current affairs magazine. Maclean's enlightens, engages and entertains 2.4 million readers with strong investigative reporting and exclusive stories from leading journalists in the fields of international affairs, social issues, national politics, business and culture*⁶².

L'autodescription du magazine permet au chercheur de comprendre que le périodique traite des affaires commerciales et internationales, de problèmes sociaux, de politiques nationales et de culture. On constate aussi que la qualité du travail journalistique est d'une grande importance pour le périodique et qu'il s'adresse à un lectorat de 2,4 millions d'individus. En sommes, l'information de départ est très générale et livre peu d'informa-

⁶¹ Pour une liste de références à ces sujets, voir David Nivens, « Objective Evidence on Media Bias: Newspaper Coverage of Congressional Party Switchers », *Journalism & Mass Communication Quarterly*, June, vol. 80, 2, 2003, p. 311-326.

⁶² Maclean's, *About Us*, <http://www.macleans.ca/about-us/>, site consulté le 15 mars 2015.

tion au chercheur qui souhaiterait satisfaire aux exigences des méthodologues de l'analyse de document.

Dans la base de données *Eureka*⁶³, on trouve cette description à propos du périodique :

Macleans est un important périodique anglophone canadien. Il a été fondé en 1905 par le Torontois John Bayne Maclean. Depuis 1994, il appartient à Rogers Communications. Il est connu pour son classement annuel des universités et collèges canadiens⁶⁴.

La date de fondation, le nom du fondateur et le fait d'appartenir à Rogers Communications, bien que ce soient des informations intéressantes, aident-elles vraiment à l'analyse documentaire?

En traitant 359 articles du *Macleans* avec le logiciel SPAD, nous avons constaté que, pendant la période étudiée, ce périodique a un regard très axé sur le Canada et sur l'économie, qu'il mentionne souvent les villes *Québec*, *Ottawa* et *Montréal*, qu'il nomme souvent le premier ministre du Canada en 2005, *Paul Martin*, et le chef de l'opposition, *Stephen Harper*. Les mots *Canada* et *Canadians* paraissent aussi souvent dans ce périodique, ce qui révèle une inclination pour les contenus canadiens. Les mots *World*, *Market* et *Funds* témoignent d'un intérêt pour les questions d'ordre économique. On peut donc dire que le vocabulaire spécifique du *Macleans*, lorsqu'il est étudié à côté de celui de trois autres périodiques canadiens, est plutôt canado-économique.

L'information que nous retirons en analysant plus de 300 articles sur une période donnée, en comparaison avec trois autres périodiques canadiens, ne trahit pas ce que dit le *Macleans*

⁶³ Université Sainte-Anne, *Bases de données A-Z*, <https://www.usainteanne.ca/bases-de-donnees#e>, site consulté le 22 mars 2015 : « Eureka donne accès au texte intégral de plus de 400 titres canadiens et plus de 1 120 titres internationaux couvrant l'actualité internationale, nationale et régionale. On y trouve non seulement des journaux, mais aussi des publications spécialisées, des fiches biographiques et profils d'entreprises, des fils de presse, des émissions télévisées et radiophoniques transcrites, des blogues triés sur le volet et des sites *web* référencés. Texte intégral : Oui. Langues : Français-Anglais-Autres. Web contrôlé, RCDR ».

⁶⁴ Eureka, *Aide en ligne*, <http://aide.eureka.cc/default.aspx?Prod=Ekbiblio&lang=fr-ca&user=user>, site consulté le 23 mars 2015.

de lui-même⁶⁵. Était-il donc nécessaire de consulter la ligne éditoriale avant d'entamer l'analyse d'un grand corpus? Il est certain que le fait de connaître cette mission avant la lecture d'un petit corpus aiderait le chercheur, comme nous l'avons déjà souligné. La prise en considération d'un grand échantillon, toutefois, indique que l'on peut dégager la tendance annoncée par le périodique sans consulter sa mission : si le *Maclean's* fait ce qu'il dit faire, nous l'observerons.

L'audescription du quotidien *Le Monde* dit aussi peu de choses :

Depuis sa création par Hubert Beuve-Mery en décembre 1944, *Le Monde* s'est imposé comme le quotidien de référence aussi bien par sa rigueur dans le traitement de l'information que par sa tradition d'indépendance vis-à-vis des idéologies et des pouvoirs. *Le Monde* est une entreprise moderne de plus de mille personnes dont le siège social est à Paris. Des frais additionnels peuvent s'appliquer pour l'accès aux archives entre 1944 et 1987⁶⁶.

Nos analyses montrent que *Le Monde* est surtout américano-économique quand nous le comparons aux trois autres périodiques français sélectionnés. Il utilise, plus souvent que les autres périodiques, les mots suivants : *million(s), États-Unis, Américain, État, dollars et Euro*. Alors que nous pouvons retirer d'autres informations à partir du logiciel SPAD, on voit déjà que le fait d'avoir *a priori* certaines informations à l'égard du quotidien informe peu l'analyse de son contenu. Nous aurions pu, comme nous l'avons fait d'ailleurs, ignorer la description que nous propose *Le Monde* de lui-même.

Que se produit-il quand nous comparons notre méta-analyse à la description que donne le propriétaire d'un périodique? Le cas du *Nouvel Observateur* est intéressant à cet égard : la description de sa mission n'était pas facile d'accès, mais nous avons trouvé une description du journal offerte par le propriétaire. Le 8 décembre 2013, Claude Perdriel, fondateur et président du

⁶⁵ Si la culture ne figure pas dans notre analyse, c'est que nous n'avons pas retenu les articles portant sur ce thème lors de notre analyse initiale.

⁶⁶ Eureka, *Aide en ligne*, <http://aide.eureka.cc/default.aspx?Prod=Ekbiblio&lang=fr-ca&user=user>, site consulté le 23 mars 2015.

groupe *Nouvel Observateur* (Nouvel Obs, Challenges, Sciences et Avenir, Rue 89), a donné une entrevue à Enguérand Renault et Alexandre Debouté⁶⁷. Lors de cette entrevue, Perdriel a affirmé que : « [*Le Nouvel Observateur* est] un journal mendésiste, social-démocrate de gauche. C'est un journal du cœur qui défend depuis cinquante ans la veuve et l'orphelin, les pauvres et les gens victimes des excès de puissance de l'État⁶⁸ ». L'analyse de traitement de données de 600 articles du *Nouvel Observateur*, lorsque comparée à celle de plus de 7 812 articles provenant de trois autres périodiques français, révèle que *Le Nouvel Observateur* s'intéresse plutôt aux personnages politiques d'ordre national et français. De plus, on comprend qu'il traite le national, l'étatsunien et le local de la même façon que le font *Le Monde*, *Le Figaro* et *Ouest France*. Ce que le propriétaire en dit n'est nullement retenu par SPAD. Cela ne veut pas dire que le propriétaire se trompe. L'analyse d'un grand corpus dénote tout simplement une tendance qui est autre que ce que dit le périodique de lui-même et, dans le cadre de la thèse que nous avançons, montre bien que le sens que donne un acteur de son travail n'est pas entièrement représentatif de ce qui se passe à tous les niveaux de la production sociale.

À partir du site *web* de *Ouest France*, on apprend qu'il est le « premier quotidien français » et « est diffusé en Basse-Normandie, Bretagne et Pays de la Loire ». « Depuis l'assemblée générale extraordinaire du 9 avril 1990, S.A. Ouest-France dépend d'une association loi 1901 à but non lucratif, l'« Association pour le soutien des principes de la démocratie humaniste », présidée par Jacques Duquesne⁶⁹ ». Selon sa page Facebook, sa devise est « Justice et Liberté⁷⁰ ». Il est vrai que nous avons observé

⁶⁷ Enguérand Renault et Alexandre Debouté, « Claude Perdriel prêt à céder le contrôle du Nouvel Obs », *Le Figaro*, 8 décembre 2013, <http://www.lefigaro.fr/medias/2013/12/08/20004-20131208ARTFIG00080-claude-perdriel-pret-a-ceder-le-contrôle-du-nouvel-obs.php>, site consulté le 12 mars 2015.

⁶⁸ *Ibid.*

⁶⁹ Ouest France, *Qui sommes-nous*, <http://www.ouest-france.fr/qui-sommes-nous>, site consulté le 8 février 2015.

⁷⁰ Tiré de la page Facebook d'*Ouest France*, <https://fr-ca.facebook.com/ouest-france>, site consulté le 8 février 2015.

qu'*Ouest France* s'intéresse particulièrement au régional par comparaison aux autres périodiques analysés, une constatation que nous avons faite grâce au traitement de 5 386 articles et qui pouvait être facilement déduite de son public cible. Mais selon ce que dit *Ouest France* de lui-même, et toujours en lien avec notre problématique de départ, on ne fait que nommer les lieux de diffusion, l'accent régional n'étant pas souligné. La méta-analyse révèle que c'est un lexique local (*aménagement, communes, conseil, maire, mairie, municipal, lotissement et pompiers*) et communautaire (*association, centre, école, salle, travaux, élève*) qui rend *Ouest France* particulier. La méta-analyse ne s'attarde pas à savoir si *Ouest France* fait preuve de « Justice et Liberté », non plus. Nous ne savons pas si ce lexique est en lien avec la justice ou la liberté en étudiant un très grand corpus. Il est vrai que nous aurions pu cibler certains mots, comme « aménagement », afin de plonger plus profondément dans la façon dont on les utilisait, et vérifier ainsi le lien qu'il y a entre la devise du périodique et le vocabulaire. Cela étant dit, notre travail consistait à définir la double tendance de l'homogénéisation et de la différenciation sur la base d'un grand corpus de données textuelles. Par conséquent, la non-vérification de l'adéquation d'une devise avec son vocabulaire objectif n'a pas nui à l'étude.

On constate donc qu'il n'est pas facile de respecter les exigences méthodologiques du cadre historique et de l'origine des documents, que le nombre d'articles rend difficile la cueillette d'informations sur les auteurs et que la mission ou la description des périodiques peut poser certains problèmes⁷¹.

⁷¹ D'autres logiciels permettent aussi de faire des analyses de données textuelles intéressantes et pertinentes, comme Alceste, Tropes ou Lexico 3. La plupart d'entre eux offrent la possibilité de travailler sur des co-occurrences et de produire des plans factoriels. En principe, en comparant les résultats que livrent ces divers logiciels, on devrait obtenir des résultats concordants et une interprétation en accord avec les prémisses de la systémique complexe et des études relationnelles devraient être possible. La section suivante précise notre argumentaire épistémologique à ce sujet.

6. Analyse de données textuelles informatisées, la systémique complexe et les études relationnelles

L'analyse de données textuelles a beaucoup changé depuis l'avènement de l'informatique. Alors que, auparavant, la perte de sens était grande lors de la quantification des données textuelles, surtout pour de grands échantillons, des logiciels comme SPAD permettent, d'une certaine manière, de préserver le cadre à l'intérieur duquel l'information initiale a été générée. Les analyses de facteurs, les analyses de concordances et même l'accès rapide aux textes intégraux fournissent aux chercheurs les outils qui sont nécessaires pour appréhender des idées dans leur complexité à différentes échelles. L'opération demande des interventions en deux ou trois temps; mais on peut dénombrer les mots et étudier leurs relations, tout en conservant les idées qui correspondent à leur utilisation. À maints égards, en effet, on ne perd pas de vue la relation entre les mots et les idées.

Quant aux exigences de prise en considération du cadre historique et de l'origine des documents, nos observations font état d'une situation tendue : plus l'échantillon est grand, plus il est difficile de rassembler les informations nécessaires pour tenir compte du contexte sociohistorique de la production des documents. Recueillir des informations sur les auteurs et sur les destinataires visés puis sur les particularités de l'écriture des documents s'avère d'autant plus difficile que le corpus est grand et qu'il compte un grand nombre d'éléments individualisables, ce qui est le cas notamment d'un échantillon d'articles de journaux à grands tirage. L'avantage de la méta-analyse est qu'elle permet de dégager les tendances d'un périodique à partir de son contenu, et non en fonction des auteurs individuels ou de ce qu'en disent les propriétaires. On pourrait critiquer l'analyse informatique des données textuelles en affirmant qu'elle ignore la présence de l'individu subjectif, de l'acteur, en tant qu'il est soit producteur du texte, soit destinataire. On pourrait dire d'elle, comme on le ferait du calcul d'une moyenne statistique, que son résultat perd les caractéristiques des individus.

Nous pensons que c'est là où se trouve le lien entre nos observations et les théories de la systémique complexe et des études relationnelles. Ces théories ne permettent pas qu'un tout existe sans ses parties, de la même façon que la moyenne ne peut pas exister sans les unités qui la composent. Ces théories ne permettent pas, non plus, que les parties expliquent pleinement le tout. On n'est pas dans l'alternative; on est dans la conjonction.

On pourrait comparer les tendances que nous observons sur divers concepts à des catégories abstraites, sociales, construites en lien avec l'activité individuelle, mais qui ne renvoient plus strictement à l'activité individuelle. Sur ce point, on peut se référer à Bachelard quand il discute des concepts et de leurs relations avec les caractères individuels :

[L]e concept [de résistance en physique] est désormais l'élément d'une loi complexe, loi au fond très abstraite, loi uniquement mathématique, qui forme une sorte de *nœud de concepts*. Alors on conçoit que l'urine, le vinaigre, le lait puissent avoir des effets spécifiques, mais ces effets ne sont enregistrés que par l'intermédiaire d'une notion véritablement abstraite, c'est-à-dire sans signification immédiate dans la connaissance concrète, sans référence directe à la sensation première. La résistance électrique est une résistance épurée par une définition précise; elle est incorporée dans une théorie mathématique qui en limite toute extension abusive. L'empirisme est alors en quelque manière déchargé; il n'a plus à rendre compte à la fois de tous les caractères sensibles des substances mises en expérience⁷².

Les tendances que nous observons par l'entremise d'une analyse de données textuelles de grande envergure n'ont plus à rendre compte de tous les journalistes; elles sont une abstraction de ce qu'ils ont produit.

Nous pouvons aussi reprendre Edgar Morin qui, en conceptualisant ses notions sur la systémique complexe, utilise l'atome pour nous aider à comprendre la relation entre l'unité et le système.

L'atome, nous dit-il, perçu auparavant comme unité de base de toute matière, doit aussi être perçu comme un système en soi. Les unités de base de cet atome connaissent une existence stable grâce au système. Un

⁷² Gaston Bachelard, *La formation de l'esprit scientifique*, Paris, Vrin, 2004 [1938], p. 127.

système n'est donc pas plus important, ou moins important, que les parties qui la composent : les parties du système font intégralement partie du système⁷³.

Or, si les tendances que nous observons proviennent d'un périodique, et que l'individu se trouve parmi les éléments de base du périodique, il se trouve aussi dans les éléments de base de ses tendances qu'il contribue à dessiner. Il est possible de décomposer les tendances en leurs éléments constitutifs; pour reprendre l'analogie du fruit de Morin, d'étudier les motivations des acteurs et des contextes qui les influencent; mais, ce faisant, nous perdons de vue les tendances que révèlent de grands échantillons de données textuelles informatisées; nous nous interdisons d'accéder à un des aspects de la complexité sociale, selon une autre perspective.

C'est précisément ce qui se produit avec la moyenne en statistique. Alors qu'elle est la somme des valeurs qu'on divise par le nombre de cas, et donc, qu'elle est un résumé, une abstraction, de ce qui se passe sur le plan des individus, elle existe en relation avec les valeurs individuelles. Pour avoir la moyenne, nous n'effaçons pas les valeurs individuelles; nous identifions la tendance de l'ensemble qu'ils forment; nous travaillons avec leur abstraction. Et, comme il y a des contraintes associées à la modélisation qu'est la formule de la moyenne, les appareils interprétatifs que nous utilisons pour l'étude de grands échantillons peuvent limiter ce que nous analysons. Si les propriétaires et les éditeurs des périodiques ne décrivent pas leur contenu comme le permettent nos analyses avec SPAD, c'est qu'ils sont aussi contraints par leurs instruments d'observation; ils n'ont pas fait de méta-analyses de leur contenu, et donc ne voient pas nécessairement les méta-tendances. Cela ne veut aucunement dire que le mandat d'un périodique ou le contexte dans lequel les auteurs écrivent n'influencent pas leurs écrits. Cela ne veut pas dire, non plus, que les motivations individuelles n'influencent pas le contenu ou qu'une recherche qui cible les motivations comme objet d'étude est peu intéressante. Ce que nous affirmons est que plus on retient d'articles pour faire écho à des problématiques complexes –

⁷³ Edgard Morin, *La méthode - 1. La Nature de la Nature*, op. cit., p. 97-98.

et mondiales en ce qui nous concerne –, alors plus on est à même de dessiner un portrait général du contenu – ou de la problématique politique véhiculée par le média – et plus on cible « le lieu où la société conçoit son intervention sur elle-même, où les organismes politiques proposent un avenir collectif⁷⁴ ». Plus on est en mesure de faire tout cela, moins on a à s'attarder sur le cadre historique et sur l'origine des documents pour rendre compte des contenus. Comme l'explique Simon Laflamme :

« En se complexifiant, la compréhension ne redit pas tout ce qui a été dit ou ne dit pas tout ce qui peut être dit; elle reprend du déjà connu, mais elle produit aussi de nouvelles connaissances, elle se donne accès à une nouvelle façon de connaître et donne de nouvelles informations sur l'objet⁷⁵ ».

L'analyse de données textuelles informatisées montre qu'on peut régler des problèmes techniques de la « décontextualisation » en recourant à des techniques informatiques. Elle fait aussi état de la difficulté que posent les préoccupations méthodologiques anthropocentriques, celles qui misent fortement sur la détermination du cadre historique et de l'origine des documents. Si nous pouvons, par-delà cet impératif microméthodologique, poursuivre nos analyses, c'est que nous n'avons pas perdu le sens de notre objet social, que nous avons tout simplement tourné notre regard vers un autre objet d'étude, que nous nous concentrons sur ce qui émerge de l'activité humaine au lieu de scruter l'activité humaine en tant que telle.

⁷⁴ Simon Laflamme, *Contribution à la critique de la persuasion politique*, op. cit., p. 23-24.

⁷⁵ Simon Laflamme, *Suites sociologiques*, op. cit., p. 59.

Bibliographie

- Adorno, Theodor et Max Horkheimer, *La dialectique de la raison : fragments philosophiques*, Paris, Gallimard, 1974.
- Alexander, Jeffrey C. et Paul Colomy (dir.), *Differentiation Theory and Social Change*, New York, Columbia Press, 1990.
- Bachelard, Gaston, *La formation de l'esprit scientifique*, Paris, 2004 [1938].
- Beck, Ulrich, *Pouvoir et contre-pouvoir à l'ère de la mondialisation*, traduit de l'allemand par Aurélie Duthoo, Paris, Flammarion, 2003.
- Bourdieu, Pierre, *La Distinction : critique sociale du jugement*, Paris, Minuit, 1979.
- Bourdieu, Pierre, *Sur la télévision/L'emprise du journalisme*, Paris, Raison d'agir, 1996.
- Breton, Philippe, « La société de l'information : de l'utopie au désenchantement », *Revue européenne des sciences sociales*, tome XL, n° 123, 2002, p. 35-39.
- Breton, Raymond, « Introduction: Rethinking the Impact of Globalization Processes - Differentiation As Well As Convergence », Raymond Breton et Jeffrey G. Reitz (dir.), *Globalization and Society. Process of Differentiation Examined*, Westport Connecticut, London, Praeger, 2003.
- Cambron, Micheline (dir.), *Le journal Le Canadien: littérature, espace public et utopie, 1836-1845*, Saint-Laurent (QC), Fides, coll. « Nouvelles études québécoises », 1999.
- Castells, Manuel, *La société en réseaux : l'ère de l'information, tome 1*, traduit de l'anglais par Philippe Delamare, Paris, Fayard, 1998.
- Chartier, Lise, *Mesurer l'insaisissable : méthode d'analyse du discours de presse*, Sainte-Foy, Presses de l'Université du Québec, 2003.
- Coates, Colin M. et Cecilia Morgan, *Heroines and History: Representations of Madeleine de Verchères and Laura Secord*, Toronto, University of Toronto Press, 2002.
- Durkheim, Émile, *De la division du travail social*, Paris, Presses universitaires de France, 1991 [1893].
- Eurika, *Aide en ligne*, <http://aide.eureka.cc/default.aspx?Prod=Ekbiblio&lang=fr-ca&user=user>, site consulté le 23 mars 2015.
- Gervais, Roger, *Les effets de la mondialisation sur la relation entre les idées et les personnes : une comparaison entre les régions toulousaine et sudburoise, mémoire de maîtrise en sociologie*, Université Laurentienne, Sudbury, 2005.

- Gervais, Roger, « La mondialisation : vers une compréhension duelle de l'homogénéisation et de la différenciation », *Nouvelles perspectives en sciences sociales*, vol. 2, n° 1, 2006, p. 69-100.
- Habermas, Jürgen, *Théorie de l'agir communicationnel, Tome 2, Pour une critique de la raison fonctionnaliste*, traduit de l'allemand par Jean-Louis Schlegel, Paris, Fayard, 1987.
- Iribarne, Philippe d', *La logique de l'honneur : gestion des entreprises et traditions nationales*, Paris, Seuil, 1989.
- Iribarne, Philippe d', *et al., Culture et mondialisation. Gérer par-delà les frontières*, Paris, Seuil, 2002 [1998].
- Klein, Naomi, *No Logo. La tyrannie des marques*, traduit de l'anglais par Michel Saint-Germain, Arles, Actes Sud, 2001 [2000].
- Laflamme, Simon, *Contribution à la critique de la persuasion politique*, Québec/Sudbury, Presse de l'Université du Québec/Université Laurentienne, 1987.
- Laflamme, Simon et Ali Reguigui, *Homogénéité et distinction*, Sudbury, Prise de parole, 2003.
- Laflamme, Simon, *Suites sociologiques*, Sudbury, Prise de Parole, coll. « Épistémè », 2006.
- Lallement, Michel, *Histoire des idées sociologiques, tome 2, de Parsons aux contemporains*, Paris, Nathan, 1993.
- Le Moigne, Jean-Louis, *La modélisation des systèmes complexes*, Paris, Dunod, 1999.
- Létourneau, Jocelyn, *Le coffre à outil du chercheur débutant. Guide d'initiation au travail intellectuel*, Montréal, Boréal, 2006.
- Maclean's, About Us, <http://www.macleans.ca/about-us/>, site consulté le 15 mars 2015.
- Maigret, Éric, *Sociologie de la communication et des médias*, Paris, Armand Colin, 2004.
- Maingueneau, Dominique, *Initiation aux méthodes de l'analyse du discours*, Paris, Hachette, 1976.
- Maingueneau, Dominique, « Présentation », *Langages*, n° 117, 1995, p. 5-11.
- Maingueneau, Dominique, *L'analyse de discours*, Paris, Hachette, 1997 [1991].
- Maingueneau, Dominique, « L'analyse du discours », dans Patrick Charaudeau et Dominique Maingueneau (dir.), *Dictionnaire d'analyse du discours*, Paris, Seuil, 2002, p. 41-45.

- Maingueneau, Dominique, « Que cherchent les analystes du discours? Argumentation et analyse du discours », *Argumentation et analyse de discours, la revue électronique du groupe ADARR*, n° 9, 2012.
- Marcuse, Herbert, *L'homme unidimensionnel : Essai sur l'idéologie de la société industrielle avancée*, Paris, Minuit, 1968.
- Marx, Karl, *Œuvres*, I, Paris, Gallimard, Bibliothèque de la Pléiade, 1965.
- Marx, Karl et Friedrich Engels, *Manifeste du Parti communiste I, les articles de F. Engels dans la « Réforme »*, traduction nouvelle par Charles Andler, Paris, G. Bellais, 1983.
- Morin, Edgar, *La méthode - 1. La nature de la nature*, Paris, Seuil, 1977.
- Morin, Edgar, *La méthode - 2. La vie de la vie*, Paris, Seuil, 1980.
- Münch, Richard, « Differentiation, Rationalization, Interpenetration: Emergence of Modern Society », dans Jeffrey C. Alexander et Paul Colomy (dir.), *Differentiation Theory and Social Change*, New York, Columbia Press, 1990, p. 441-464.
- Nivens, David, « Objective Evidence on Media Bias: Newspaper Coverage of Congressional Party Switchers », *Journalism & Mass Communication Quarterly*, vol. 80, n° 2, 2003 p. 311-326.
- Ouest France, Qui sommes-nous*, <http://www.ouest-france.fr/qui-sommes-nous>, site consulté le 8 mars 2015.
- Parsons, Talcott, *Sociétés : essai sur leur évolution comparée*, Paris, Dunod, 1973 [1966].
- Poulain, Jean-Pierre, *Sociologies de l'alimentation*, Paris, Presses Universitaires de France, 2003 [2002].
- Renault, Enguérand et Alexandre Debouté, « Claude Perdriel prêt à céder le contrôle du Nouvel Obs », *Le Figaro*, 8 décembre 2013, <http://www.lefigaro.fr/medias/2013/12/08/20004-20131208ARTFIG00080-claude-perdriel-pret-a-ceder-le-contrôle-du-nouvel-obs.php>, site consulté le 12 mars 2015.
- Robinson, David, *Muslim Societies in African History*, Cambridge (UK), Cambridge University Press, 2004.
- Sabourin, Paul, « L'analyse de contenu » (chapitre 15), dans Benoît Gauthier (dir.), *Recherche sociale : de la problématique à la collecte des données*, Québec, Presses de l'Université du Québec, 2006 [2003], p. 357-385.
- Simmel, Georg, *Le Conflit*, traduit de l'allemand par Sibylle Muller, Dijon-Quetigny (France), Circé, 2003 [1995].
- St-Pierre, Stéphanie, « Étienne Brûlé : la création d'un personnage », *Revue du Nouvel-Ontario*, n° 29, 2004, p. 5-44.

- Thibeault, Jimmy, *Des identités mouvantes. Se définir dans le contexte de la mondialisation*, Montréal, Nota bene, coll. « Terre américaine », 2015.
- Touraine, Alain, *Pourrons-nous vivre ensemble? Égaux et différents*, Paris, Fayard, 1997.
- Université Sainte-Anne, Bases de données A-Z, <https://www.usainteanne.ca/bases-de-donnees#e>, site consulté le 22 mars 2015.
- Pareto, Vilfredo, *Traité de sociologie générale*, Paris, Droz, 1968.
- Vincent, Jean-Marie, *La théorie critique de l'École de Francfort*, Paris, Galilée, 1976.
- Voilquin, Jean, *Les penseurs grecs avant Socrate*, Paris, Garnier-Flammarion, 1964.