



**HAL**  
open science

**Les “ Morceaux choisis ” comme enjeu dans la  
constitution de corpus textuels pour les sciences sociales  
post-positivistes**

Dimitri Della Faille

► **To cite this version:**

Dimitri Della Faille. Les “ Morceaux choisis ” comme enjeu dans la constitution de corpus textuels pour les sciences sociales post-positivistes. Usages de la lexicométrie en sociologie, Jun 2013, Versailles, France. hal-01620762

**HAL Id: hal-01620762**

**<https://hal.science/hal-01620762>**

Submitted on 20 Oct 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Les « Morceaux choisis » comme enjeu dans la constitution de corpus textuels pour les sciences sociales post-positivistes

*della Faille, Dimitri*<sup>1</sup>

Professeur

Département des sciences sociales  
Université du Québec en Outaouais

Des nombreux enjeux que pose la constitution de corpus, nous avons choisi de réfléchir ici aux « morceaux choisis » qui sont ces segments, extraits ou fragments de discours sélectionnés par idéologie ou par subterfuge qui confortent l'analyste dans ses attentes, hypothèses ou certitudes formées par ailleurs. Ces « morceaux choisis » sont, en fait, une opportunité d'amorcer une réflexion plus fondamentale sur l'accès à la connaissance par l'analyse du discours, sa nature et son utilité.

La constitution de corpus textuels relève d'une pratique proche de celle de l'artisan ou du corps de métier. Dans ce texte, nous allons effectuer le parallèle entre la constitution du corpus par l'analyste et l'abattage et le commerce des animaux par les garçons-bouchers. Le texte se développe en trois temps. Premièrement, il s'agira d'envisager la constitution d'un corpus comme une entreprise de violence et de prédation dans lequel l'analyste joue un rôle primordial. Nous montrerons comment, à l'image de l'abattage d'un bœuf, le discours est transformé en morceaux et comment l'analyste triomphe, peut-être un peu trop facilement, face au discours. Deuxièmement, il s'agira de proposer des pistes de réflexion sur les différentes approches positivistes, néo-positivistes et post-positivistes qui peuvent fonder le sens et la méthode des différentes manières d'envisager l'analyse du discours par lexicométrie. Troisièmement, il s'agira de montrer comment un ensemble de « morceaux choisis » peut être envisagé par certaines approches comme un corpus taré de vices rédhibitoires viciant le commerce de son analyse. En guise de conclusion, nous prendrons position contre une condamnation trop rapide des corpus de « morceaux choisis ».

## **Premier temps : Le découpage en morceaux du discours**

Afin d'être analysé, le discours textuel ou textualisé doit être réduit, ensuite, segmenté et puis encore, les morceaux rassemblés dans une opération dont la précision doit être chirurgicale, les procédures bien

---

<sup>1</sup> Pour prendre contact avec l'auteur : Dimitri della Faille Département des sciences sociales Université du Québec en Outaouais 283 boul. A. Taché CP 1250, succ. Hull Gatineau, Québec J8X 3X7 Canada [dimitri.dellafaille@uqo.ca](mailto:dimitri.dellafaille@uqo.ca)

maîtrisées et effectuées par des mains d'expérience. L'on peut, sans excès d'imagination, voir dans cette position à l'égard de l'analyse de discours par corpus textuel et lexicométrie un parallèle évident avec la transformation, par abattage, des animaux en viande de consommation.

Tel un garçon-boucher, l'analyste palpe le discours; il a probablement déjà choisi les morceaux du corps qu'il s'apprête à confronter. En 1837, l'*Encyclopédie d'agriculture pratique* sous la direction de Jacques Alexandre Bixio, décrit sous la signature du conservateur des abattoirs, Louis-Charles Bizet, en ces termes cette première opération qui voit le garçon-boucher se frotter au corps résistant du bœuf duquel l'on tirera les meilleurs morceaux :

Le maître garçon palpe les bœufs du bourgeois, et choisit celui qui lui paraît le mieux disposé à être fait. Le choix fixé, l'animal est coiffé du câble fatal, et conduit à l'échaudoir par le second garçon; il est suivi par le premier, qui, armé d'un gros bâton, frappe les pieds de derrière du bœuf, lorsqu'il ne marche pas avec une certaine bonne volonté. (Bixio, 1837, p. 537)

C'est que l'opération de réduction du discours au corpus d'analyse ne s'effectue pas sans une certaine force brute. Le corpus ne s'offre pas sans résister. Il nous faut faire preuve de force musculaire, mais aussi, beaucoup d'entraînement et de jugement. L'*Encyclopédie d'agriculture pratique* continue, en ces termes, sa description de l'abattage :

Le bœuf, ainsi dirigé, arrive à l'échaudoir avec plus ou moins de résistance, résistance d'ailleurs toujours vaincue par la force, l'adresse et le courage des garçons bouchers; il arrive et bientôt il est fixé, d'une manière à peu près inébranlable, à l'anneau d'abattage au moyen du câble doublement entrelacé dans ses cornes; le pauvre animal tombe étourdi avec un épouvantable fracas; cependant les coups de la masse se succèdent avec rapidité, jusqu'à ce que le bon soupir soit soufflé. (Ibid.)

C'est fait... Le discours a arrêté de résister! Le garçon-boucher a réussi à imposer la force percutante et contondante de ses outils et la maîtrise de son expérience face à un bœuf dont on dit que la masse, sans aucun doute des muscles et une structure osseuse impressionnante, vient d'emplir l'abattoir d'un grand bruit sourd lorsqu'elle s'écrasa sur le sol.

Avec sagesse, l'*Encyclopédie* nous indique cependant que l'opération n'est pas des moindres. Tel l'animal, le discours est tenace : « quelques fois les bœufs ne tombent pas sous les premiers coups de la masse; on en a vu résister au terrible choc répété plus de cent fois » (Ibid.). En effet, l'opération peut se révéler pleine de surprises. Il nous arrive d'interpréter hâtivement des signes. Alors qu'on le pensait maîtrisé, le discours nous

envoi des signaux contradictoires. Mais, bientôt il en sera fini de cette résistance.

Les yeux du bœuf expriment une vive douleur, ils sont tristes et languissants; le mouvement des membres antérieurs est totalement arrêté, mais celui des membres postérieurs ne l'est pas; les cuisses et les jambes sont assez vivement agitées. (Ibid.)

Maintenant, l'analyste triomphe, il avait espoir que la résistance soit vaine. Il a dompté l'animal, il a pris le dessus sur cette force sociale. Mais, cette opération n'est pas une mince affaire. Nombreux sont les analystes qui rencontrent de la résistance dans la constitution d'un corpus sous la forme, entre autres, de problèmes techniques de formatage du corpus textuel, de difficultés d'accès aux données textuelles, ou au contraire, d'une masse qui pourrait être jugée accablante. Avec l'expérience et en fonction de leur formation et de leur jugement les analystes développent des techniques propres.

Ces techniques et manières de faire varient, parfois même grandement. Divers outils s'offrent à l'analyste. Le garçon-boucher est lui aussi confronté à différentes écoles qui peuvent lui sembler étranges, barbares ou aller à l'encontre de son éthique. Ou bien, il fait face à de nouvelles manières de faire qui offrent des solutions concrètes à des problèmes pour lesquels il ne pouvait, jusqu'alors, envisager de dénouement facile. Face à des enjeux concrets, l'analyste doit garder une ouverture sur des façons de faire qui lui sont étrangères et se dire que les opérations peuvent être améliorées. Le garçon-boucher, en tous cas, lui, il le sait. Certains bœufs nécessitent des opérations particulières dont l'inspiration peut provenir d'ailleurs.

Pour abrégier la lente agonie des bœufs à tête molle, quelques bouchers emploient comme en Espagne l'énervation. C'est la section de la moelle épinière, opérée par l'introduction d'une sorte de stylet étroit et effilé, entre l'occipital et la première cervicale. A peine cet instrument est-il plongé, que le bœuf tombe avec une rapidité et une violence qui ferait croire que la foudre vient de l'écraser. (Ibid.)

Maintenant, le discours est sur le point d'être transformé en objet de recherche. L'analyste doit aller chercher dans le discours les segments ou morceaux nécessaires à son entreprise. Il réduit la réalité du discours à un ensemble de textes prêt pour l'analyse. Cette réduction demande autre chose qu'une force brute. *L'Encyclopédie* nous décrit les étapes auxquelles le garçon-boucher est confronté lorsqu'il retire du bœuf les premiers morceaux qui seront ensuite triés :

En maintenant la tête et en appuyant sur elle son genou droit, il ouvre le cou par une assez longue incision, légèrement cruciale, faite

auprès du larynx : il enlève d'abord le ris, puis plonge son couteau, qui va couper l'artère aorte : le sang alors s'écoule avec abondance; la quantité que fournit chaque bœuf peut être évaluée à deux seaux. Pendant cette opération, et avant même de la commencer, le second garçon passe une corde au pied gauche de devant du bœuf; il en tient l'extrémité en se posant sur le derrière de l'animal, et foule les flancs avec son pied droit pour faire sortir le sang avec plus de facilité. (Op. cit., p. 538)

Avant que le discours ne soit transformé en corpus, plusieurs opérations de tri sont nécessaires. Du corps du bœuf, tous les morceaux n'ont pas le même intérêt, la même destination, ni la même fonction. Leur traitement et leur extraction servent des objectifs distincts.

Les quatre pieds sont immédiatement coupés et séparés de leurs patins, qui ne sont autres que les tendons d'Achille, propriété des garçons qui les vendent aux fabricants de colle forte; les pieds appartiennent aux bouchers, ils servent à faire de l'huile et du noir animal. [...] La langue est d'abord enlevée; puis la toile, partie de suif qui enveloppe les intestins; un fort tinet; espèce d'anse en bois, est passé ensuite dans les jarrets de l'animal; ce tinet est accroché à la corde de treuil, et le bœuf est enlevé successivement à la hauteur nécessaire pour faciliter la vidange et le dépouillement. Les premiers travaux au moment de l'enlèvement sont de retirer les intérieurs; on commence par la vessie et les ratis; ensuite viennent les estomacs, le foie, la rate et l'amer; enfin un double coup de coteau détache le mou ou le poumon, et le cœur, qui tombent ensemble. (Ibid.)

Comme le bœuf qu'on égorge, qu'on énerve, la constitution d'un corpus de textes a comme effet de retirer la vie du discours. Mais, la métaphore bouchère a ses limites. Le discours n'est pas une force sociale à laquelle on peut mettre fin en procédant à l'analyse. Qu'il y ait intervention de l'analyste ou pas, le discours vit. Et l'analyste y contribue, du moins, une fois que le produit de son analyse soit réinjecté dans le monde du discours.

En présentant ces quelques extraits, nous avons voulu témoigner de la nature du travail, parfois minutieux, brusque ou déprécié, souvent physique qu'entreprend l'analyste lorsqu'il sélectionne, dans le discours, les morceaux qui constitueront son corpus d'analyse. S'il existe des procédures dont certaines sont standardisées et éprouvées, l'analyste doit continuellement faire appel à son jugement. C'est bien à ce titre que nous tâchons ici de réfléchir à la place des prises de décisions de l'analyste dans le dépeçage du discours en segments et à ses enjeux pour l'analyse par les sciences sociales interprétatives et empiriques.

## **Deuxième temps : L'accès à une connaissance des faits sociaux**

Une fois le discours violenté, que l'on l'aura retiré de son espace vital, l'analyste est confronté à un autre tracas, celui de la constitution de son corpus pour lequel il devra effectuer de nombreux choix. S'il est de bon aloi de dire que parmi ces choix, certains seront meilleurs que d'autres, il faut reconnaître que ce jugement sur la qualité du corpus dépend du statut qu'on lui accorde dans l'analyse. Ce statut de l'analyse est ancré dans la conception que se fait, de manière explicite ou implicite, l'analyste de sa position à l'égard de l'accès à la connaissance du social. Si l'on peut postuler des traits généraux aux sciences sociales, il existe pourtant plusieurs pratiques quant à l'utilisation et à l'objectif d'une analyse.

Ainsi, l'anthropologue français Jean-Pierre Olivier de Sardan postule que sciences sociales sont à la fois interprétatives et empiriques (Olivier de Sardan, 1996). Elles sont interprétatives, parce qu'elles travaillent à partir du sens qu'investissent les acteurs dans leur pratique. Pour avoir accès à ce domaine de signification, l'analyste passe nécessairement par la représentation ou par le langage qui sont autant d'éléments de pratique et d'action. Le langage est un élément de motivation de la pratique et de l'action ainsi qu'un moyen duquel elles tirent leur sens. Les sciences sociales sont empiriques parce qu'elles accèdent à la connaissance à partir de l'observation, des données et des faits sociaux. Pour l'analyste du discours cela veut dire que l'interprétation est liée à des contraintes empiriques dont le corpus fait partie. L'interprétation est fondée dans le corpus, envisagé comme des données sociales, et la constitution de celui-ci est une contrainte de l'accès à la connaissance. Une fois qu'on a postulé le caractère interprétatif et empirique des sciences sociales, il est nécessaire de situer sa position épistémologique afin d'éviter tout malentendu. Car en fait, le statut du corpus d'analyse, sa forme et ses limites sont autant de conséquences de la démarche de l'analyste.

Nous situons notre démarche en rupture avec une pratique dite positiviste des sciences sociales. Nous définissons cette pratique positiviste comme une démarche qui chercherait à établir des règles générales de la vie sociale et qui cherche à universaliser les observations (Fischer, 1998; Wildemuth, 1993). Le comportement social, les institutions et le langage appartiendrait à une réalité ordonnée en fonction de lois qui peuvent être observées de manière objective. Cette approche positiviste est dite objective car elle postule une indépendance entre les lois observées et les caractéristiques personnelles de l'analyste ou son appartenance à un groupe social. Les découpes du bœuf et son abattage sont standardisés. Les coupes ne sont pas le résultat de caractéristiques individuelles et du jugement moral du garçon-boucher. Depuis les années 1960, la pratique positiviste des sciences sociales est sous attaque. Les attaques les plus

discutées ont, sans aucun doute, été celles de Thomas S. Kuhn (1962), Paul Feyerabend (1975) et Imre Lakatos (1976) qui ont pavé la route à des ouvrages tels *La vie de laboratoire* de Bruno Latour et de Steve Woolgar (1979). Ces attaques ont provoqué des changements importants dans les sciences sociales, en particulier au sein des sciences sociales nord-américaines dès les années 1980 (Pleasant, 1997).

D'un bord, une salve de critiques du positivisme provient d'une approche qu'on peut qualifier de néo-positiviste qui reconnaît l'importance du projet positiviste, mais qui regrette que celui-ci n'ait pas été en mesure de systématiser ses concepts, techniques et attitudes scientifiques (Duchastel et Laberge, 1999; Fischer, 1998). Cette approche exalte des traits imaginés d'une certaine modernité non achevée dont on souhaiterait qu'elle triomphe de l'obscurantisme. Le néo-positivisme regrette que le positivisme et la modernité n'aient été, avant tout, que des attitudes et que le projet n'ait jamais été complètement réalisé. En lexicométrie, nous sommes souvent confrontés à de tels concepts, attitudes et techniques qui, par exemple, cherchent à percer les secrets des universels linguistiques ou encore qui cherchent à créer des chaînes automatiques de traitement des langues. Pour la constitution d'un corpus, cela signifie que les données textuelles devraient, par exemple, être exhaustives ou représentatives de la réalité sociale qui est analysée; qu'on constituerait des corpus de manière automatique ou encore que la différence entre l'échantillon et la population soit pratiquement nulle. Les avancées technologiques, matérielles et logicielles ont renforcé l'idée que de tels corpus était possibles, voir même souhaitables.

La sociologie et l'histoire des sciences nous offrent des pistes de réflexion nous permettant de mettre en perspective le profond sentiment d'accomplissement du travail de recherche par les valeurs d'une stricte méthode et d'une vigilance constante. Par exemple, le sociologue étatsunien Robert K. Merton établit un lien entre l'exaltation de la raison scientifique et l'importance croissante des valeurs puritaines en Angleterre au 18<sup>ème</sup> siècle (Merton, 1936). Merton nous dit que ces valeurs piétistes et puritaines ont été des forces motrices d'une nouvelle science qui exalte la raison. Bien que la thèse de Merton ait été critiquée (Becker, 1984), il faut reconnaître que l'importance que nous accordons à des principes de rationalité et d'empirisme est aussi un construit social. Tout autant que le sont les sentiments profonds qui peuvent nous habiter, en tant qu'analyste, pour le travail méthodique et la vigilance constante. L'observation, la déduction et l'accès à la connaissance par l'empirie et l'idée même d'une science sociale sont autant de produits de l'histoire et des dynamiques sociales dont il est difficile de s'abstraire. Il faut donc

reconnaître qu'il existe toujours « une idéologie scientifique avant une science dans le champ où la science viendra s'instituer » (Canguilhem, 1970, p. 12).

D'un autre bord, une salve de critiques du positivisme provient d'une approche que l'on peut qualifier de post-positiviste qui se place dans une rupture historique et épistémologique avec l'approche positiviste. Il faut cependant reconnaître que dans leurs concepts et techniques, les sciences sociales post-positivistes peuvent parfois reproduire certaines attitudes que l'on peut attribuer au positivisme (Clark, 1998). En particulier, lorsqu'il est question d'observer, de mesurer et de décompter. Cependant, le rapprochement s'arrête là. La mesure et la quantification de divers éléments du tissu social, incluant la lexicométrie, ne sont pas pour autant des attitudes positivistes ou néo-positivistes. Contrairement au positivisme et à ses formes contemporaines, les approches post-positivistes reconnaissent que le processus d'analyse est le résultat de positions prises continuellement qui peuvent être extérieures aux préoccupations liées à la nature de l'objet étudié. À toutes les étapes de l'analyse, incluant bien sûr la réduction du discours au texte, ou à une forme textualisée, l'analyste est appelé à prendre des décisions. Pour reprendre la métaphore bouchère, le garçon-boucher est un élément d'explication qui contribue à donner forme au résultat des coupes de viande. Les choix effectués sont liés au jugement de l'analyste et non à une imaginaire perfection cumulative des procédures méthodologiques qui serait à l'image de la réalité de laquelle l'analyste désire s'approcher à tâtons. Les sciences sociales post-positives postulent également le caractère subjectif de ce qu'on appelle « réalité ». Elles reconnaissent qu'on ne peut accéder qu'à des fragments de connaissance par l'interprétation et la cognition et que ceux-ci font partie d'une mosaïque de représentations de la « réalité » (Duchastel, 1993). Elles voient dans l'interprétation autre chose qu'une simple approche exploratoire ou un des multiples outils de l'analyste. L'analyse interprétative ne rejette pas la possibilité d'une recherche. Plutôt, elle découle du postulat que l'analyste ne peut pas accéder aux faits sociaux et ne les comprend qu'au travers d'une médiation dont il est un des éléments, un des filtres, un des facteurs d'explication (Letourneau et Allen, 1999). Les résultats de sa recherche sont liés au contexte de celle-ci, mais aussi, ils sont intimement liés à l'expérience d'observation par l'analyste. Pour la constitution d'un corpus, cela signifie que l'analyste visera avant tout à rassembler des textes qui lui permettent d'examiner les dimensions sociales qu'il cherche à examiner. Les données textuelles devraient, à cet égard, être pertinentes, cohérentes et aller dans le sens des objectifs de l'analyse.



Il nous faut donc reconnaître que l'analyse de corpus est une entreprise de prédation du fait social et que la constitution du corpus est un élément de procédure qui contient certaines dimensions de violence. Mais aussi, que la constitution de corpus est une étape qui passe par des prises de décision, par des jugements de l'analyste. Les « morceaux choisis » font partie des prises de décision qui mènent à la constitution d'un corpus; ils sont autant d'éléments en relation avec le jugement, l'éthique et l'épistémologie de l'analyste.

### **Troisième temps : Les « morceaux choisis » comme vices rédhibitoires**

La lexicométrie ou l'analyse statistique de texte vise, par exemple, à comparer ou vérifier des résultats ou encore à repérer des pistes de recherche (Leimdorfer et Salem, 1995). La lexicométrie appliquée aux sciences sociales vise, grâce à une analyse statistique contrastive ou descriptive du lexique et de sa répartition dans un corpus, à dégager certains aspects sociaux, politiques ou culturels. La littérature en lexicométrie offre plusieurs réflexions à l'égard de la question du corpus et de sa constitution. En lexicométrie, un corpus est entendu comme un ensemble textuel déterminé et un lien avec une méthode définie (Dubois, 1969). Entre autres choses, pour que cette méthode soit valide et efficace, la littérature en lexicométrie nous apprend qu'un corpus doit avoir certaines caractéristiques. Il doit être raisonné, cohérent, représentatif, homogène ou encore, en volume suffisant (Péry-Woodley, 2005; Pincemin, 1999). D'un point de vue heuristique, le corpus est le mode privilégié de l'étude des rapports sociaux envisagée du point de vue de la production de sens par le texte ou l'oral. Le corpus est ainsi une médiation entre l'analyste et la langue (Mellet, 2002). On peut aussi affirmer que le corpus est une incarnation spécifique ou une explicitation d'une problématique d'analyse (Charaudeau, 2009; Mayaffre, 2002). Mais, si pour la lexicométrie le corpus est tout cela, un corpus c'est également la matérialisation d'une épistémologie, d'une mise en pratique de concepts théoriques ou d'une position idéologique de l'analyste. C'est à ce titre que la question des « morceaux choisis » lors de la constitution d'un corpus pose certains enjeux.

Nous appelons « morceaux choisis » les éléments retenus des suites d'une sélection restreinte alors que l'analyste constitue un corpus en fonction d'une hypothèse de recherche établie. Ces éléments sélectionnés et ensuite analysés, font partie d'une réduction d'une vaste population ou de l'ensemble du discours à un échantillon dit favorable aux hypothèses de recherche. Comme nous l'avons illustré plus haut, une fois le bœuf abattu,

les morceaux sont répartis pour divers usages. Tout comme le corpus, les morceaux du cadavre du bœuf découpé ont des fonctions spécifiques, mais le cas des « morceaux choisis » dépasse la seule différence fonctionnelle des segments ou des éléments périphériques au texte ou contextuels. Ils sont une réduction du discours à un corpus qui pourrait être considéré comme tronqué. Ils peuvent avoir été choisis par maladresse, par problème logique, par ignorance des enjeux d'une telle sélection, par manque de données, par aveuglement idéologique ou encore par subterfuge. Dans le cas de l'aveuglement idéologique ou du subterfuge, l'on pourrait reprocher aux corpus constitués de « morceaux choisis », d'extraits ou de fragments, de donner des résultats qui vont dans le sens des attentes de l'analyste. Si l'on adopte une démarche d'exaltation de la rationalité scientifique et de neutralité des valeurs, ce corpus aurait comme fonction unique de défendre une thèse idéologique qu'il serait impossible d'invalidier. Ce corpus serait une fraude intellectuelle; il serait réputé avoir des vices rédhibitoires. L'analyste qui aurait commis le crime de choisir ses morceaux devrait être condamné. De retour à l'*Encyclopédie d'agriculture pratique*, on y apprend que ces vices, dans le cadre du commerce de la viande, sont passibles de poursuites :

Dans le commerce des animaux, comme dans celui de toute autre denrée, le vendeur est tenu, aux termes de l'article 1625 du Code civil de garantir à l'acquéreur la possession paisible de la chose vendue, et les défauts cachés de cette chose ou vices rédhibitoires. [...] On appelle vices rédhibitoires, aux termes de l'article 1641 du Code civil, tous les défauts cachés de la chose vendue qui la rendent impropre à l'usage auquel on la destine, ou qui diminuent tellement cet usage, que l'acheteur ne l'eût pas acquise ou n'en aurait donné qu'un moindre prix s'il les avait connus. (Bixio, 1837, p. 552)

Comme dans le commerce des animaux, la présentation des résultats d'une analyse d'un corpus de « morceaux choisis » contiendrait des vices rédhibitoires. C'est qu'une telle analyse irait à l'encontre d'une approche qui valorise la vérification ou la falsifiabilité des résultats de l'analyse. À ce titre, les « morceaux choisis » devraient être vigoureusement dénoncés.

### **En guise de conclusion**

Afin de conclure cette courte contribution, nous désirons prendre position à l'égard des corpus de « morceaux choisis » et critiquer leur dénonciation. Il ne faut pas se méprendre sur nos intentions. Ce n'est pas tant que nous désirions pousser l'analyste à constituer des corpus de « morceaux choisis » par idéologie ou utiliser des subterfuges, mais plutôt que nous intéressons aux arguments et aux objectifs de la dénonciation de telles approches. Nous ne sommes pas partisan de la fraude intellectuelle, ni même du fait

de cacher des vices du corpus. Il ne s'agit pas de faire l'apologie d'une « fausse » science qui ne rencontrerait jamais le faux et qui n'aurait à renoncer à rien (Canguilhem, 1970 p.7). Plutôt, nous sommes contre leur condamnation au nom d'un certain triomphe de la raison scientifique et de l'exaltation de valeurs puritaines. L'analyste doit reconnaître que toute entreprise d'analyse est empreinte de valeurs et d'idéologie. Certaines entreprises peuvent l'être plus explicitement que d'autres.

De manière ultime, le corpus et son analyse permettent de tenter d'expliquer les phénomènes sociaux. Une analyse, et la lexicométrie n'y échappe pas, est une réduction de la complexité de la « réalité » sociale. C'est une représentation du social qui tire sa légitimité de la sophistication de sa méthode et du capital symbolique qui est accordé à ses analystes. L'analyse permet l'établissement de typologies qui serviront d'autres explorations d'analyse. L'analyse permet aussi la montée en théorie qui sert l'objectif de prescrire une explication du comportement social, les institutions et le langage parmi les autres.

Les ouvrages classiques qui ont marqué l'histoire des sciences sociales nous servent de repères afin de comprendre le tissu social. Ils ont donné naissance à une science sociale interprétative et empirique qui contribue également à forger les relations sociales. Mais, ces ouvrages ne sont pas réputés pour la force de leur démonstration méthodologique et s'ils se sont imposés grâce aux outils qu'ils nous fournissent pour donner du sens au vivre en commun.

L'anthropologie (comme la sociologie ou l'histoire) a ainsi une composante « histoire des idées » où les livres réputés qui y ont pris place figurent, non en raison de leur pertinence empirique (souvent mise à mal par des travaux ultérieurs plus modestes mais mieux documentés), mais du fait de la nouveauté, de la séduction ou de l'originalité « en leur temps » des « thèses », des « paradigmes », des « postures heuristiques », et parfois des formules qui y figurent et qui leur ont valu la notoriété. [...] De par ce délestage empirique, [ces livres réputés] prennent place dans un circuit herméneutique savant (ou méta-interprétatif) dans lequel on ne leur demande plus de comptes quant à leur véridicité. (Olivier de Sardan, 1996, p. 54)

S'il est nécessaire de dénoncer le positivisme et l'idée d'une connaissance cumulative fondée dans une « réalité » vérifiée empiriquement et de questionner la base à partir de laquelle les sciences sociales opèrent leur interprétation du social, nous ne sommes pas totalement en mesure de nous abstraire du cadre de réflexion et des termes du débat que cette approche impose. Dans cet article nous avons quand même tâché de réfléchir à l'activité d'analyse comme un acte de violence et de prédation à l'égard du discours, mais surtout, nous nous sommes interrogé à propos de

la relation qui existe entre le statut accordé à l'analyse et aux jugements de l'analyste et la manière d'envisager le corpus et les formes souhaitables à donner à celui-ci afin d'accéder à la connaissance du social.

## **Bibliographie**

Becker, George (1984) « Pietism and Science: A Critique of Robert K. Merton's Hypothesis », *American Journal of Sociology*, vol. 89, no. 5, pp. 1065-1090.

Bixio, Jacques Alexandre (sous la dir.) (1837) *Maison rustique du XIX<sup>ème</sup> Siècle. Encyclopédie d'agriculture pratique. Tome second. Cultures industrielles et animaux domestiques*, Bureau du Journal d'agriculture pratique, Paris.

Canguilhem, Georges (1970) « Qu'est-ce qu'une idéologie scientifique? », *Organon*, no. 7, pp. 3-13.

Charaudeau, Patrick (2009) « Dis-moi quel est ton corpus, je te dirai quelle est ta problématique », *Corpus*, no. 8, pp. 37-66.

Clark, Alexander M. (1998) « The Qualitative-Quantitative Debate: Moving From Positivism and Confrontation to Post-Positivism and Reconciliation », *Journal of Advanced Nursing*, vol. 27, no. 6, pp. 242-249.

Dubois, Jean (1969) « Lexicologie et analyse d'énoncé », *Cahiers de lexicologie*, no. 2, pp. 115-126.

Duchastel, Jules (1993) « Discours et informatique: des objets sociologiques? », *Sociologie et sociétés*, vol. 25, no. 2, pp. 157-170.

Duchastel, Jules et Danielle Laberge (1999) « La recherche comme espace de médiation interdisciplinaire », *Sociologie et sociétés*, vol. 31, no. 1, pp. 63-76.

Feyerabend, Paul (1979 [1975]) *Contre la méthode, Esquisse d'une théorie anarchiste de la connaissance*, Seuil, Paris.

Fischer, Frank (1998) « Beyond Empiricism: Policy Inquiry in Postpositivist Perspective », *Policy Studies Journal*, vol. 26, no. 1, pp. 129-146.

Kuhn, Thomas S. (1983 [1962]) *La Structure des révolutions scientifiques*, Flammarion, Paris.

Lakatos, Imre (1984 [1976]) *Preuves et Réfutations : essai sur la logique de la découverte mathématique*, Éditions Hermann, Paris.

Latour, Bruno et Steve Woolgar (1988 [1979]) *La Vie de laboratoire. La Production des faits scientifiques*, La Découverte, Paris.

- Leimdorfer, François et André Salem (1995) « Usages de la lexicométrie en analyse du discours », *Cahiers des sciences humaines*, vol. 31, no. 7, pp. 131-143.
- Letourneau, Nicole et Marion Allen (1999) « Post-positivistic Critical Multiplism: A Beginning Dialogue », *Journal of Advanced Nursing*, vol. 30, no. 3, pp. 623-630.
- Mayaffre, Damon (2002) « Les corpus réflexifs : entre architextualité et hypertextualité », *Corpus*, no. 1, pp. 51-69.
- Mellet, Sylvie (2002) « Corpus et recherches linguistiques : introduction », *Corpus*, no. 1, pp. 5-13.
- Merton, Robert K. (1936) « Puritanism, Pietism and Science », *Sociological Review*, vol. 28, no. 1, pp. 1-30.
- Olivier de Sardan, Jean-Pierre (1996) « La violence faite aux données », *Enquête*, no. 3, pp. 31-59.
- Péry-Woodley, Marie-Paule (2005) « Discours, corpus, traitements automatiques », dans Anne Condamines (sous la dir.) *Sémantique et corpus*, Hermès, Paris, pp. 177-210.
- Pincemin, Bénédicte (1999) « Construire et utiliser un corpus : le point de vue d'une sémantique textuelle interprétative », dans Anne Condamines, Marie-Paule Péry-Woodley et Cécile Fabre (sous la dir.), *Atelier Corpus et TAL : pour une réflexion méthodologique, VI conférence annuelle sur le Traitement automatique des langues naturelles (TALN 99)*, Cargèse, France, 12-17 juillet 1999, pp. 26-36.
- Pleasants, Nigel (1997) « The Post-Positivist Dispute in Social Studies of Science and its Bearing on Social Theory », *Theory Culture Society*, vol. 14, no. 3, pp. 143-156.
- Wildemuth, Barbara M. (1993) « Post-Positivist Research: Two Examples of Methodological Pluralism », *The Library Quarterly*, vol. 63, no. 4, pp. 450-468.