
Analyse de contenu et analyse lexicale, Le cas d'une étude en management public

Céline Desmarais

Institut de Recherche en Gestion et Economie (IREGE), Université de Savoie (Annecy)

Celine.Desmarais@univ-savoie.fr

Jean Moscarola

Institut de Recherche en Gestion et Economie (IREGE), Université de Savoie (Annecy)

Jean.Moscarola@univ-savoie.fr

ABSTRACT. Lexical statistics and textual data analysis open new opportunities for content analysis. However, the use of these tools also raises questions. How should they be used and how can they be given scientific legitimacy? Which methods and which research strategies should be privileged?

In order to answer these questions, this article discusses a research project in public management based on interviews with municipal workers and elected representatives. This example is used to demonstrate how lexical analysis and content analysis can be combined, and to what extent each methodology brings specific insights to the corpus.

KEY WORDS: Content analysis, Lexical analysis, Public administration

RÉSUMÉ. Les outils de traitement informatisés des données textuelles ouvrent de nouvelles perspectives pour les méthodes d'analyse des données qualitatives. Cependant l'utilisation de ces outils n'est guère stabilisée et leur validité pose question. Comment s'y prendre pour en tirer une connaissance, et comment garantir à celle-ci une légitimité scientifique ? Quelles méthodes privilégier parmi l'arsenal dont le chercheur dispose désormais ? Pour apporter des réponses à cet ensemble de questions, cet article analyse un processus de recherche ayant eu pour objectif d'appréhender l'évolution des rôles de l'encadrement des villes, en contexte de modernisation. Il expose plus particulièrement deux méthodes : l'analyse de contenu et l'analyse lexicale, dont les potentialités et la complémentarité sont développées.

MOTS-CLÉS : Analyse de contenu, analyse lexicale, administration publique

1. Introduction

L'analyse des données textuelles connaît un succès croissant dans les sciences humaines. Ce succès est renforcé par la vogue pour les approches qualitatives ou constructivistes. Analyse de contenu, analyse de discours, statistique lexicale, lexicographie, toutes ces méthodes sont de plus en plus faciles à mettre en œuvre grâce aux progrès des logiciels (In Vivo, Hyperbase, Alceste, Spad T, Sphinx Lexica....). Ils font gagner beaucoup de temps, permettent le

traitement de corpus abondants et, par le biais de la numération, apportent aux études qualitatives le crédit accordé aux approches quantitatives. Les outils informatisés ouvrent ainsi de nouvelles perspectives dans le traitement et l'analyse des données qualitatives. Cependant quelle que soit leur efficacité, ces méthodes, pour être bien utilisées, doivent assurer leur rigueur scientifique. Elles nécessitent une expérience difficile à acquérir. L'utilisation des outils informatisés doit encore faire l'objet d'une réflexion de la communauté scientifique pour affiner les méthodes et valider des procédures robustes, garantes de gains de temps pour les chercheurs.

Le but de cet article est d'aider le chercheur dans l'exercice des choix méthodologiques nécessaires à l'approche des corpus textuels : comment s'y prendre pour en tirer une connaissance, et comment garantir à celle-ci une légitimité scientifique ? Quelles méthodes privilégier parmi l'arsenal dont le chercheur dispose désormais ? Afin d'apporter des réponses à ces questions, cet article analyse un processus de recherche portant sur l'évolution des rôles des cadres dans les collectivités territoriales désireuses de moderniser leur gestion, pour illustrer ces méthodes et préciser leur pertinence. Pour cela nous évoquerons les différentes méthodes d'analyse des données textuelles, avant de nous pencher sur deux méthodes centrales : L'analyse de contenu et l'analyse lexicale. Nous montrerons la complémentarité de méthodes souvent présentées comme concurrentes.

2. Analyse des données textuelles, des techniques aux choix méthodologiques.

Les différentes méthodes de traitement des données qualitatives se rattachent à des traditions intellectuelles et des prémisses méthodologiques divergentes.

2.1 La tradition du commentaire

Une première approche qu'on peut rattacher à la tradition de la critique littéraire ou artistique consiste à développer un commentaire par lequel le critique fait part des impressions, interprétations, réflexions que lui suggère sa lecture. Sur le mode du résumé et de la citation, il renvoie au texte pour illustrer son propos et gagner ainsi l'adhésion du lecteur. Tout son art consiste à développer une argumentation en se référant aux idées admises ou débattues sur le sujet et à en proposer de nouvelles. Dans cette démarche, les extraits du texte sont utilisés comme illustrations ou preuves du commentaire. La qualité de l'analyse repose entièrement sur le crédit accordé au chercheur quant à sa compréhension et son intelligence du corpus, à la confiance faite dans ses choix de citations et dans son habileté à écrire et argumenter. Le courant de l'analyse sémiologique [BARTHES ; FLOCH 88...] s'apparente à cette très longue tradition du commentaire.

Les méthodes que nous proposons ici se démarquent de cette tradition, dans la mesure où toutes font appel à un outillage informatique et statistique. Elles imposent au chercheur un ensemble de contraintes qui peuvent le conduire à rendre mieux compte du processus d'analyse. Mais nous n'irons pas jusqu'à opposer ces approches, tant l'expérience nous a appris, qu'en matière de construction du sens, il est illusoire de prétendre éliminer toute subjectivité.

2.2 L'analyse de contenu

Les méthodes classiques de l'analyse de contenu [BELERSON 54] préconisent une approche rigoureuse et méthodique fondée sur l'explicitation des règles de lecture, d'interprétation et de codage. Elles reposent sur l'élaboration d'un cadre de référence auquel on confronte le contenu du texte. Ce cadre peut être établi à priori et se fonder sur des modèles généraux indépendants du corpus particulier ou au contraire être progressivement construit en cours de lecture et d'analyse.

Quelle que soit la démarche suivie (que le cadre de référence soit modifié ou enrichi en cours d'analyse), on procède toujours par une lecture visant à repérer dans le texte les éléments du modèle ou de la grille d'analyse utilisée. Dans un deuxième temps intervient l'analyse statistique. Elle porte sur l'analyse des éléments de la grille de référence : fréquence d'apparition, variation selon les locuteurs, selon les contextes, interdépendance entre les éléments du modèle.

2.3 L'analyse des données textuelles

Le paradigme de l'analyse de données textuelles offre une autre voie basée sur l'analyse de la fréquence des éléments (mots ou formes graphiques) présents dans le texte et considérés comme indicateurs des actes de langage [AUSTIN 70 ; SEARLE 72] ou comme trace des modèles cognitifs [Johnson-Laird 94]. Il s'agit alors de dresser la statistique des décisions de langue [MULLER 93 ; HARRIS] pour remonter aux facteurs susceptibles de les expliquer ou tracer la cartographie des associations lexicales (BENZECRI 81) en révélant ainsi les réseaux sémantiques ou modèles mentaux dont le texte porte la trace. Ces méthodes opèrent sur les seules propriétés statistiques des textes et peuvent mettre au jour les structures linguistiques qui les caractérisent. Il appartient alors à l'analyste de prendre connaissance de ces substituts lexicaux : les mots les plus fréquents, les usages sur représentés selon les locuteurs, les propension à associer certains termes et à en opposer d'autres....

Dans le cas de très grands corpus, la statistique permet de tirer parti de la redondance de la langue pour réduire considérablement l'effort de la lecture ou orienter le retour au texte, à partir d'entrées lexicales que seule la statistique est en mesure de révéler à la curiosité du chercheur.

L'analyse de ces deux principales méthodes faisant appel à l'outillage informatique sera exposée, dans les développements suivants, à travers l'exemple d'une recherche en management public. Elle débouchera sur la mise en valeur de deux types de complémentarités très caractéristiques des méthodes d'analyse de données textuelles :

- La première renvoie à la posture du chercheur qui peut s'appuyer sur un modèle préexistant ou au contraire explorer le corpus sans aucune préconception.
- La seconde concerne la manière dont il envisage le matériau textuel, support d'énoncés et de contenus ou trace des processus d'énonciation.

3. L'étude de cas : Le discours des acteurs du management public

La réflexion menée sur l'analyse des données qualitatives s'est appuyée sur une recherche doctorale dont nous brosons, pour la compréhension du lecteur, la problématique, le terrain et la nature et l'étendue du corpus analysé.

3.1 La problématique : l'évolution du rôle de l'encadrement

La reconfiguration de l'action publique face à un environnement perçu comme turbulent est au cœur des questionnements du management public. Après une décennie d'effervescence et d'expérimentations visant à transformer le management dans les organisations publiques il convient d'évaluer les transformations en profondeur qui en résultent. Pour apporter une contribution à ce débat récurrent, l'étude porte sur une évaluation des changements engagés, à travers l'analyse des transformations de la fonction d'encadrement.

Les références théoriques de cette recherche se situent au croisement d'un ensemble de domaines, complexes et mal formalisés : les principes diffusés par les doctrines de la modernisation administrative et les règles de fonctionnement émanant des systèmes de gestion des ressources humaines. Ces références théoriques conduisent à analyser les évolutions de la fonction d'encadrement à travers le concept de rôle. Ce concept a tout d'abord été opérationnalisé suivant les approches classiques en gestion des entreprises :

L'approche par les activités *qui définit le rôle comme une émanation des interactions* informelles entre les acteurs. Les rôles se déduisent des activités observables des managers, selon une méthodologie popularisée par Henry Mintzberg (1984).

L'approche par les fonctions qui voit le rôle comme un système de contraintes émanant de la structure formelle de l'organisation. Les rôles sont alors un ensemble de fonctions de l'organisation : planifier, organiser, coordonner, commander et contrôler [FAYOL 16].

C'est suivant cet éclairage qu'ont été menés les entretiens à l'origine du corpus étudié.

Tableau 1 : lignes directrices des entretiens.

Eléments de contexte : ville, état des méthodes, identité de l'interviewé, statut, compétence...	
Fonctions	Mode d'exercice (activités)
Rôle d'expertise Rôle de gestion : organisation allocation des moyens Rôle managérial animation contrôle finalisation	Décision (degré d'autonomie) Information (mode d'accès et de diffusion) Rôles interpersonnels verticaux horizontaux externes

3.2 Le terrain : 160 interviews dans 7 villes moyennes

La gestion des collectivités locales a constitué le terrain concret de cette recherche. Sept villes moyennes ont été sélectionnées. Toutes ces villes ont affiché, au cours des 10 dernières années, une ambition de moderniser le management des hommes en utilisant l'apport des outils de GRH. Dans chaque ville, une vingtaine de personnes a été interrogée : 93 cadres intermédiaires, 42 agents, 13 dirigeants, 6 élus, et 9 représentants syndicaux, soit au total 160 interviews d'une durée moyenne d'1 heure 30. Ces acteurs ont ainsi été invités à discourir sur

leur rôle, vu comme « *un ensemble de représentations partagées et de relations d'interdépendance, attachées à une position dans l'organisation* ».

C'est ce matériau qui a été analysé, avec la difficulté de tenir compte du caractère convenu des discours tenus sur le management dans les organisations. En effet, il est fréquent que, notamment dans les niveaux hiérarchiques supérieurs, le discours tenu sur les rôles de l'encadrement reflète davantage « *ce qu'il est de bon ton de dire* » plutôt qu'une réalité de terrain. La perspective multi-acteurs permet de confronter et d'éprouver les discours recueillis.

3.3 De la retranscription des interviews à la structuration du corpus

La retranscription des 160 interviews représente plus de 500 pages. Ce matériau considérable, collecté en suivant un guide d'entretien, reproduit tous les aléas et les détours de la conversation : digressions par rapport aux questions posées, retours sur des points déjà évoqués... Le regroupement de ces éléments épars, selon les dimensions et les concepts qui guident la recherche, a conduit finalement à une base structurée selon 19 variables de contexte situant la ville et le répondant et 18 variables texte dans lesquelles a été ventilé le contenu des interviews. Ces variables reprennent l'approche en terme d'activités et de fonctions qui a guidé les entretiens et organisent ainsi le discours dans un format plus propice à l'analyse. Le corpus total obtenu représente 232 633 mots (91 272 sans les mots outils), qui constituent la base du travail d'analyse de contenu et d'analyse lexicale que nous allons maintenant décrire.

4. Analyse de contenu. : Confronter le corpus à un modèle de référence

4.1 Le codage selon le modèle des « fonctions / activités »

La masse des données recueillies nécessite de s'attacher dans un premier temps à réaliser deux phases essentielles [MILES 93] :

- Une phase de condensation des données visant à synthétiser, élaguer, trier et organiser les données pour pouvoir ensuite tirer des conclusions.
- Une phase de présentation des données visant la création d'un nouveau format qui présente de manière systématique des informations véhiculées par le discours.

Ce travail a été effectué fonction par fonction, en cherchant à identifier 2 aspects :

- Celui du point de vue traditionnel/bureaucratique
- Celui du discours de la modernisation.

A titre d'illustration nous allons ci après développer ce travail en ce qui concerne le rôle de **contrôle**.

Celui-ci peut être défini a priori selon deux extrêmes :

Le point de vue traditionnel bureaucratique voit dans ce rôle la garantie de la transmission, sans trahison de la volonté des représentants du peuple par des fonctionnaires, considérés comme des « automates aveugles » [BODIGUEL 96]. La conséquence est que le rôle de contrôle des encadrants se traduit par des contrôles pointilleux de l'exécution stricte des

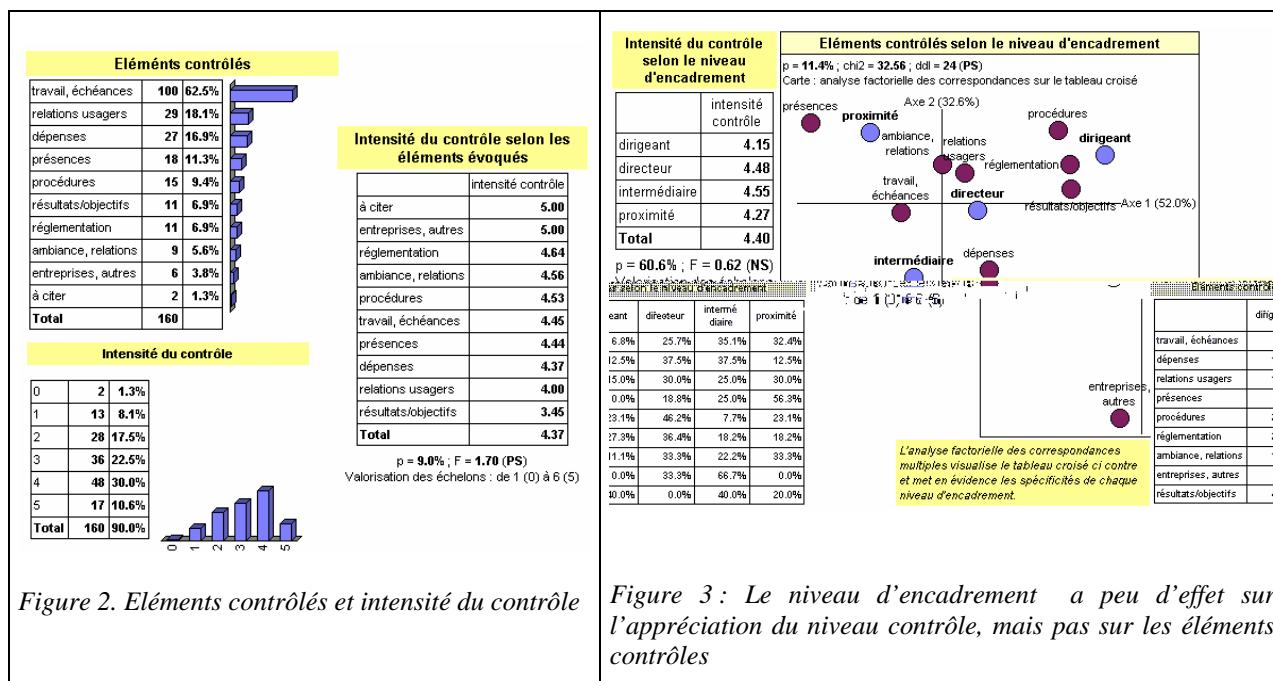
consignes à tous les niveaux de la ligne hiérarchique, vue comme une courroie de transmission de l'impulsion politique et repérable par les éléments sur lesquels ce contrôle s'exerce.

Le discours de la modernisation promeut une conception du contrôle centrée sur l'autonomie des acteurs, à tous les niveaux de l'organisation, et sur la mise en place d'un cycle de management qui suppose un contrôle, centré non plus sur les actes quotidiens, mais sur les résultats à moyen terme, voire un auto contrôle de fonctionnaires responsabilisés.

La confrontation de ces modèles conduit à repérer dans les discours analysés la nature des éléments contrôlés et l'intensité du contrôle exercé, afin de déterminer comment situer le contrôle exercé entre les deux extrêmes identifiées. Cette étape conduit donc à la construction de deux variables. La figure ci-dessous illustre cette première étape du travail d'analyse de contenu par lequel le chercheur prend connaissance du discours relatif à chaque fonction et code son contenu dans les 2 variables « *éléments sur lesquelles porte la fonction* » et « *intensité de l'implication dans cette fonction* ». Il a également la possibilité de noter les idées ou formulations remarquables qui pourront être ultérieurement reprises. On passe ainsi du discours des acteurs aux conclusions de l'analyste. Ceci pose bien sûr le problème de la fatale subjectivité des interprétations. Elle est d'autant mieux contrôlée que le chercheur peut se référer à l'ensemble des énoncés du répondant.

Figure 1 : Analyse de contenu de la fonction de contrôle

L'analyse individuelle de chaque interview selon des catégories conceptuelles plus abstraites permet ensuite d'utiliser la statistique pour synthétiser, observer les tendances et les convergences et enrichir ainsi la compréhension par une vue d'ensemble, tant de la nature des contenus codés que de leur croisement avec des variables de contexte.



On constate ainsi sur l'exemple de la fonction de contrôle (Fig. 2.) qu'elle porte très majoritairement sur les tâches et les échéances et beaucoup moins sur les objectifs, la réglementation, ou l'ambiance... D'autre part il est remarquable que, lorsque les objectifs sont évoqués, l'intensité du contrôle apparaît beaucoup moins forte, ce qui correspond assez bien aux idées de la modernisation...

Cette première confrontation relative à la consistance interne du discours peut être complétée par le contexte de sa production en examinant les variations selon le niveau d'encadrement (Fig. 3.). Ceci conduit à constater qu'il n'y a pas de différence significative sur l'appréciation de l'intensité du contrôle selon les niveaux d'encadrement.

4.2. L'émergence d'un nouveau modèle d'analyse des rôles des cadres territoriaux en contexte de modernisation.

Le modèle de référence initial auquel le corpus a été confronté a également montré les limites de ses vertus explicatives et compréhensives et de sa capacité à rendre compte du contexte public et des mouvements de modernisation. C'est pourquoi, celui-ci a été remis en cause par les résultats de l'analyse de contenu qui a été menée initialement. En effet la première grille d'analyse s'est révélée trop éloignés des réalités émergents des rôles décrits par les acteurs.

Ces conclusions, réalisées à partir d'un premier niveau d'analyse, ont alors permis de proposer une grille d'analyse des rôles davantage conforme au contexte public, l'approche par les processus qui se pose en alternative à l'approche par les fonctions et les activités.

Tableau 2 : Une nouvelle approche des rôles par les processus

Objet du rôle	Référentiel traditionnel (- -)	Référentiel adapté (+ +)
Consignes	Rôle de transmission Mise en œuvre des consignes, contrôle de leur réalisation et retour au niveau supérieur. Expertise de mise en œuvre et d'information.	Rôle de traduction Pilotage itératif, prise en compte des problèmes d'utilisateurs et de collaborateurs, expertise décisionnelle. Utilisation d'outils synthétiques.
Externalités	Surveillance Contrôle par les écrits et les procédures budgétaires. Nature des externalités prises en compte : respect de la réglementation, des procédures.	Maîtrise des externalités Par la culture et les interactions informelles. Nature des externalités prises en compte : qualité du service public, image de la ville.
Relations humaines	Supervision : Evitement hiérarchique Motivation à travers les moyens octroyés, appel au sens du service public. Pilotage individuel des collaborateurs. Coupure hiérarchique et entre les services.	Pilotage relationnel Pilotage par l'écoute, la convivialité, l'information, l'esprit d'équipe. Accompagnement vers l'autonomie, et la responsabilisation, souci d'implication. Attention aux relations avec l'extérieur
Ressources	Exécution des moyens Mise en œuvre des ressources sous un contrôle pointilleux.	Adaptation des moyens Conception et négociation des moyens, optimisation des ressources.

Ce nouveau modèle a permis un réexamen des réponses à **l'ensemble des questions posées**, en faisant abstraction de la structure initiale du questionnement semi-directif. Sa mise en œuvre empirique a donc été éprouvée par l'intermédiaire d'un codage de second niveau. Sur chaque rôle, nous avons attribué la notation (+ +) en cas de présence importante du rôle rénové, (+) en cas de présence plus mitigée, (- -) en cas de présence manifeste du rôle traditionnel et (-) en cas de présence mitigée¹.

La figure ci-dessous montre l'un des écrans qui a permis la relecture du corpus à travers cette codification non prévue a priori et qui a permis de confronter d'une manière différente au discours recueillis, les conclusions ayant émergé du premier niveau d'analyse.

¹ Pour ce codage, compte tenu du niveau d'interprétation élevé, nous avons réalisé un double codage des données avec calcul de la fiabilité intercodeurs et réduction des divergences par discussion sur cas d'espèces.

Situer par rapport aux référentiels empiriques

Rôles

Quelle est la nature et quels sont les principaux acteurs du contrôle effectué sur l'activité du service que vous encadrez ?

Il y a forcément un rôle de contrôle car c'est moi qui suis responsable devant les élus et donc je dois me donner les moyens de ne pas avoir à répondre de choses qui ne me plairaient pas. Je m'assure que chacun est au courant de ce qu'il faut faire. Très souvent quand je me promène dans les équipements j'évoque certains problèmes avec les responsables. Le contrôle est de s'assurer que tout le monde est au courant des règles et les applique. Je fais des informations, des réunions pour connaître où on en est... Il y a aussi le contrôle budgétaire exercé par l'attachée. On nous demande de plus en plus de rigueur par rapport au mouvement des fonds publics.

Comment sont déterminés et diffusés les objectifs s'appliquant au service que vous encadrez ?

A Annecy la situation est facile car la ville s'intéresse à la politique culturelle. On n'a pas besoin d'aller les convaincre sans arrêt. Les points d'achoppement sont plutôt sur les modes de fonctionnement. Les élus font plutôt confiance aux fonctionnaires. on a une assez grande marge de proposition. La situation au niveau des moyens est assez bonne mais il est difficile de faire des projets qui demanderaient des moyens supplémentaires. Il y a quelques années le maire nous a incité à développer un programme d'éducation artistique et on a redéployé des moyens. En ce qui concerne les objectifs de la bibliothèque c'est difficile car le quotidien est lourd. A Annecy la lecture est sur sa lancée. Il y a eu 20 ans d'expansion en partant de très bas et en arrivant très haut. Il y a eu beaucoup de moyens mais la bibliothèque vit sur sa lancée. J'ai souvent signalé qu'il faudrait réformer l'organisation de la bibliothèque qui date des années 60, car la ville a évolué et il n'y a plus adéquation entre la physionomie de la ville et le réseau de lecture publique. Pour les élus pour le moment il n'y a pas grand chose à faire pour la bibliothèque. Pour le moment je ne suis pas entendue.

Implications des situation étudiées

Rôle traduction

Maîtrise externalités

Pilotage relationnel

Adaptation moyens

Réglementation
travail, échéances
relations usagers
présences
procédures
résultats/objectifs
entreprises, autres
dépenses
ambiance, relations
à citer

Référentiel traditionnel --, -
Référentiel adapté ++

Figure 4 : Analyse de contenu transversale en vue de documenter le référentiel empirique

Au terme de cette nouvelle analyse de contenu, un traitement statistique a pu être mené permettant d'affiner et d'éprouver les rôles découlant du modèle (Fig. 5).

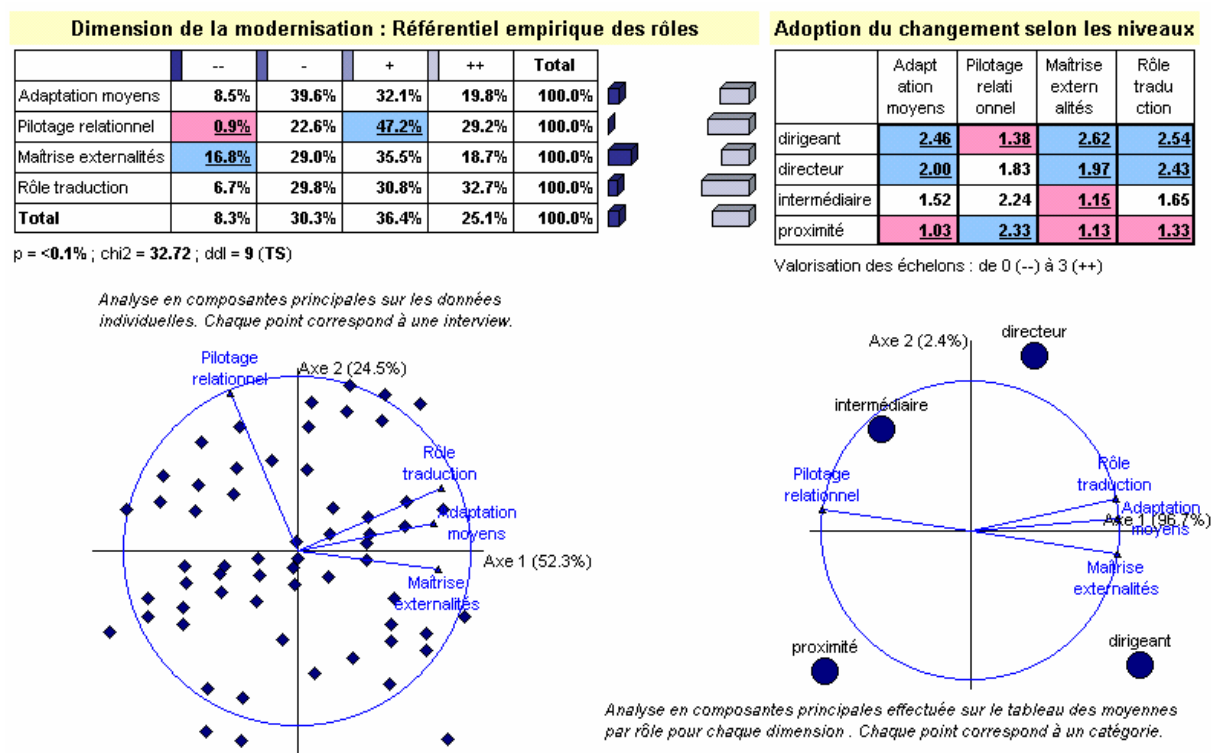


Figure 5 : L'évolution des rôles selon les dimensions de l'analyse transversale

Cette analyse montre d'une part que l'adaptation aux nouveaux challenges est plus nette pour ce qui concerne le pilotage relationnel et le rôle de traduction que pour l'adaptation des moyens et la maîtrise des externalités. D'autre part, on peut constater que le référentiel adapté concerne essentiellement les franges élevées de l'encadrement, sauf pour le rôle de pilotage relationnel, mieux adopté par l'encadrement intermédiaire ou de proximité.

Cela nous a permis de conclure sur l'existence d'un effet différencié des démarches de modernisation par la GRH :

- ✓ En ce qui concerne les franges élevées de l'encadrement, cela se traduit par une réelle évolution des rôles qui s'éloignent distinctement des caractéristiques traditionnelles pour consacrer le statut d'acteur à part entière des cadres concernés.
- ✓ En revanche, l'encadrement intermédiaire et de proximité demeure cantonné dans des rôles de transmission des consignes et de contrôle de leur exécution. Ces aspects fort traditionnels sont cependant amendés par le rôle de pilotage relationnel qui instaure des relations hiérarchiques et transversales amendant partiellement les dysfonctionnements engendrés par la persistance de l'organisation bureaucratique.

4.3 La dynamique de l'analyse de contenu : d'un modèle à l'autre, la construction progressive du sens

Les étapes qui précèdent montrent bien que le raisonnement a progressé de la manière suivante :

- 1/ Le premier niveau d'interprétation a eu lieu conformément à l'a priori de la collecte et au modèle « fonctions/activités »
- 2/ L'analyse des données a remis en cause cette première vision
- 3/ Cette remise en cause a permis l'émergence d'un nouveau modèle d'interprétation confronté à nouveau au corpus mais d'une manière plus globale, affranchie du découpage initial.

La recherche a ainsi progressé par affinement et remise en cause du cadre conceptuel. Cette démarche montre tout le gain que le chercheur peut retirer de la flexibilité d'une approche qualitative, en remettant en cause les prémices de départ et en ayant la possibilité de conforter les nouvelles conclusions en les confrontant à nouveau au terrain. Ce travail de construction progressive nécessite les moyens informatiques qui permettent tout à la fois de manipuler de grande quantité de données, de les restructurer et de modéliser les interfaces qui permettent d'organiser la lecture et le travail d'analyse de contenu et de codification. Enfin, ces outils doivent également intégrer tous les moyens de traitement statistique qui permettent de valider ou d'invalider les modèles et de faire ainsi progresser la compréhension des phénomènes.

5. L'analyse lexicale : Prendre de la distance et contrôler l'interprétation

En se focalisant sur les actes de langage, l'analyse lexicale offre une alternative à l'analyse des textes par leur contenu. Dans ce qui précède, l'analyse de contenu a été utilisée pour

analyser l'énoncé, c'est à dire ce que les acteurs ont voulu donner comme représentation de leur action dans l'organisation. Même si le dispositif empirique a été organisé pour permettre l'expression sincère des acteurs et le croisement des représentations, celui-ci n'est pas exempt des limites habituelles d'une recherche largement fondée sur des discours. Par ailleurs, comme nous l'avons soulevé ci-dessus, les discours sur les rôles d'encadrement sont particulièrement sujets à ce type de distorsion. De la part des encadrants il s'agit de montrer que l'on est un bon chef, performant et humain. Les discours se font donc largement le relais des présupposés sur ce que devrait être un « bon cadre », véhiculés par les approches managériales. Le risque est alors que le cadre serve au chercheur le discours qu'il attend, en lien avec sa propre connaissance des dernières tendances managériales. Ce biais est donc susceptible de prendre une ampleur particulière, dans le cadre du thème de notre recherche.

Pour limiter cette difficulté, nous avons utilisé l'apport de l'analyse lexicale. Celle-ci permet d'éclairer les différents aspects du langage susceptibles de fournir des indices sur la pensée, les sentiments ou les représentations du locuteur. Cela suppose un plus grand avancement sur la voie de l'interprétation, à travers l'analyse des adaptations personnelles que les individus font subir, consciemment ou non à l'outil linguistique. Le chercheur analyse les modalités du dire [GAVARD-PERRET 98].

Ceci passe par la **distinction entre énoncé et énonciation** qui explique les usages différenciés que nous avons attribués à l'analyse de contenu (post-codages sémantiques) et aux analyses lexicales. Alors que le travail de post-codage prend peu de recul par rapport au caractère littéral de l'énoncé et aux présupposés du chercheur, les analyses lexicales peuvent être utilisées pour prendre de la distance par rapport au texte et rechercher, dans les traits de l'énonciation, les représentations portées par les catégories d'acteurs.

5.1 L'approche exploratoire

L'avantage de l'analyse lexicale, telle qu'elle est très couramment pratiquée, est tout d'abord de faire l'économie d'une réflexion formalisée à priori. En cela elle s'apparente à la tradition du commentaire. Elle ne s'en distingue que par le fait de remplacer la lecture du texte par ses substituts lexicaux dérivés des procédés de la statistique lexicale. En tirant parti de la redondance de la langue pour examiner les termes les plus fréquemment employés, on se fait beaucoup plus rapidement une idée. L'utilisation d'analyseurs syntaxiques et la lemmatisation perfectionnent l'analyse. La lemmatisation est une opération qui permet de regrouper dans de mêmes unités des formes graphiques qui correspondent aux différentes flexions d'un même lemme [LEBART 94]. Cette opération est un préalable indispensable à la constitution de lexiques, car elle permet de regrouper les termes employés à travers leur racine commune et d'agréger ainsi une partie des propos tenus pour une plus grande pertinence de l'analyse². Examiner alors les substantifs, les verbes, les adjectifs les plus fréquents permet une approche très efficace du corpus et donne une idée de son contenu (l'énoncé), mais met à jour les traits caractéristiques de l'énonciation comme le montrent les tableaux et figures ci dessous.

² Le logiciel de traitement de données ramène ainsi : les formes verbales à l'infinitif, les substantifs au singulier, les adjectifs au masculin singulier et les formes élidées à la forme sans élision.

5.2 Caractériser les énoncés : se faire une idée des contenus

service	575	gestion	99	mairie	58	activité	45	cas	37	horaire	28
travail	317	dossier	98	jour	57	moment	45	collaborateur	37	qualité	28
budget	270	besoin	97	responsabilité	56	politique	45	encadrement	37	quotidien	28
problème	266	année	92	plus	55	gars	42	accueil	36	signature	27
objectif	254	réunion	90	adjoint	54	quelque_chose	42	point	36	souci	27
élu	243	ca	89	jury	54	compétence	41	règle	36	chantier	26
agent	233	fois	89	priorité	54	crédit	41	emploi	35	façon	26
rôle	217	planning	87	avis	53	domaine	41	hiérarchie	35	groupe	26
directeur	205	cadre	85	chacun	53	partie	41	avancement	34	place	26
équipe	204	maire	85	conseil	52	quartier	41	finance	34	sport	26
personnel	202	question	85	tâche	52	autonomie	40	ligne	34	accord	25
chose	185	direction	84	terrain	52	celui	40	situation	34	aspect	25
personne	182	heure	83	autre	51	commande	40	suivi	34	bâtiment	25
contrôle	180	chef	82	nous	51	eux	40	même	33	entreprise	25
temps	142	secteur	80	sécurité	51	fonction	40	produit	33	fonctionnaire	25
cela	135	congé	79	confiance	50	manière	40	engagement	32	relation	25
gens	133	an	72	maîtrise	50	mois	40	choix	31	réponse	25
organisation	124	information	69	moyen	50	achat	39	tout_le_monde	31	mal	24
courrier	118	matériel	66	proposition	50	bon_de_commande	39	intervention	30	décal	23
demande	113	collègue	65	mission	49	bureau	39	nombre	30	monsieur	23
recrutement	113	fonctionnement	61	compte	48	coup	39	semaine	30	orientation	23
poste	111	formation	61	décision	48	difficulté	39	système	30	parent	23
projet	107	niveau	61	dépense	48	franc	39	évaluation	29	structure	23
responsable	105	procédure	59	école	48	remplacement	39	matin	29	délégation	22
ville	101	entretien	58	enfant	48	investissement	38	action	28	gardien	22

Figure 6 : 150 premiers substantifs pour se faire une idée de l'énoncé

La liste ci dessus (Figure 6), produite à partir de la totalité du texte de chaque interview, réduite après lemmatisation aux seuls substantifs, met en évidence les champs lexicaux suivants :

- ✓ Les hommes et les acteurs : *agent, équipe, personne, gens, collègue, maire, adjoint, responsable, chef...*
- ✓ Le temps : *temps, années fois, planning heure moment, jour, semaine, matin, horaire...*
- ✓ La gestion administrative et financière : *budget, contrôle, organisation, courrier, gestion, dossier, réunion, poste*

A l'inverse certain termes sont absents ou apparaissent peu, notamment tous ceux des services concrets assurés par une collectivité locale. Comme si la gestion municipale était d'abord concentrée sur ses propres acteurs ses procédures et son organisation...

Cette première approche sort les termes de leur contexte. Elle peut être complétée par un traitement plus sophistiqué mettant en évidence la manière dont les mots se trouvent associés dans une même réponse, un même paragraphe, une même phrase... L'analyse factorielle (Figure 7) des correspondances multiples permet d'effectuer ce type d'analyse. Elle donne, sous forme de carte, une représentation visuelle des associations lexicales. Les configurations ainsi mises en évidence figurent la manière dont les termes proches se trouvent associés dans un même contexte ou au contraire plutôt exclus l'un de l'autre s'ils sont éloignés.

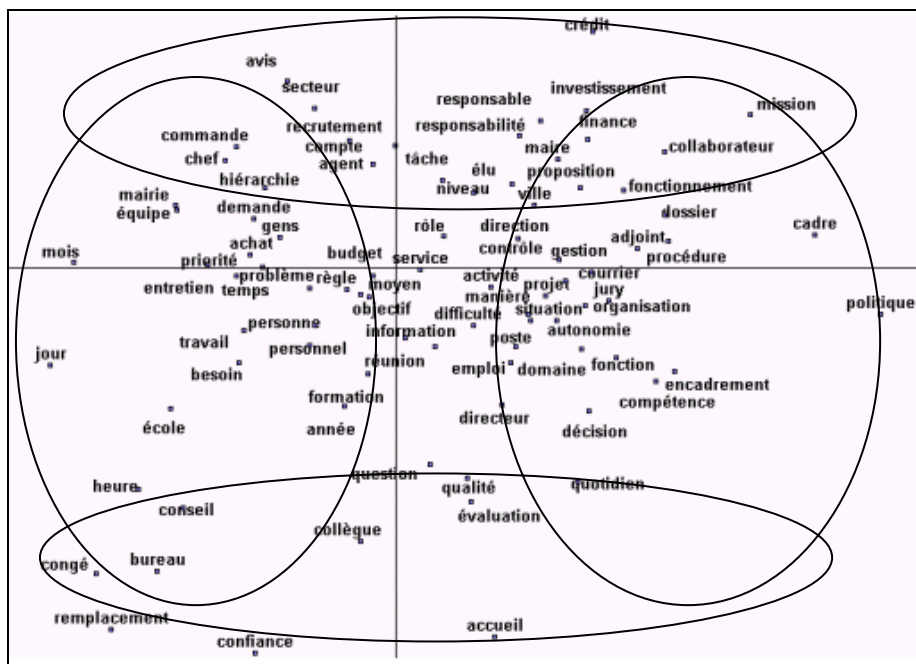


Figure 7 : Analyse factorielle des correspondances : carte des associations lexicales

A la lecture de cette carte on peut trouver 2 systèmes opposant d'une part les descriptions abstraites (à droite) et concrètes (à gauche) et renvoyant d'autre part à l'organisation (en haut) et aux acteurs (en bas). Cette interprétation offre une piste pour rendre compte de la complexité des représentations qui sous tendent les discours. Ces interprétations doivent bien sûr être vérifiées par un retour au texte afin de confronter ce que les actes de langages peuvent faire entendre et ce que le texte dit.

5.3 Analyser l'énonciation : situer les locuteurs

L'analyse lexicale permet, selon différents procédés statistiques, de réduire le texte aux éléments lexicaux. En assimilant ces éléments à des mots clé, on peut se faire une idée du contenu du texte étudié. Mais en procédant ainsi c'est le discours qu'on met en pièce pour mettre à jour sous la surface de l'énoncé les actes de langage caractéristiques de l'énonciation. Ceux-ci nous renseignent sur la posture des locuteurs. Ainsi par exemple l'examen des seuls pronoms personnels met en évidence un clivage très net entre les dirigeants et encadrants (Figure 8 carte de gauche) d'une part et les élus, responsables syndicaux et agent d'autre part. Ce clivage est celui de la posture individuelle (je, me, mon, me, leur) et du collectif, (nous, notre, on) on le retrouve à l'intérieur du seul groupe des encadrants...

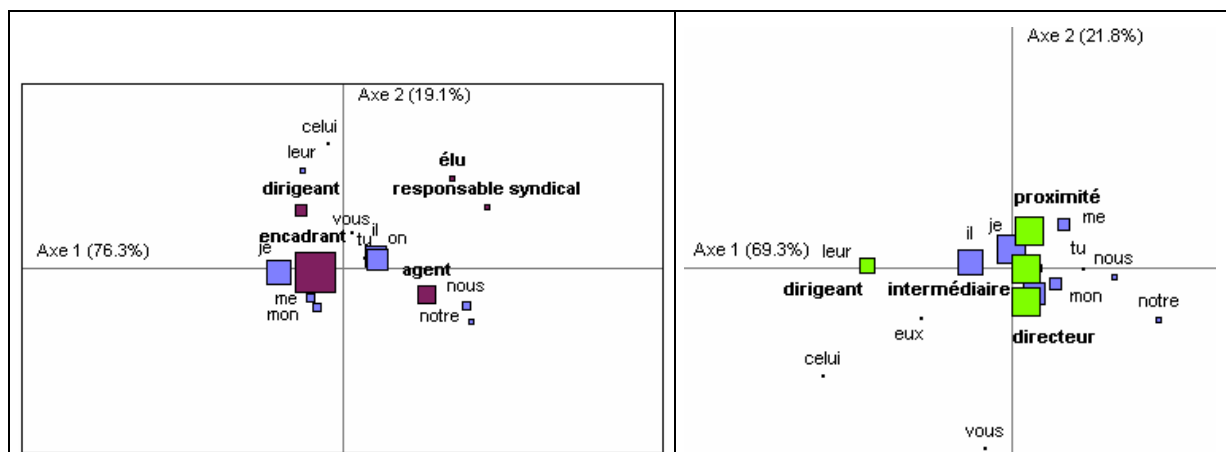


Figure 8. Les caractéristiques de l'énonciation : l'usage des pronoms personnels selon le statut

Si elle aborde le texte d'une manière radicalement différente de l'analyse de contenu, l'analyse lexicale peut être une méthode complémentaire susceptible de contrôler, valider et éclairer ou compléter les résultats de l'analyse sémantique.

6. Analyse lexicale et analyse de contenu

Procédé rapide pour une prise de connaissance des caractéristiques d'un corpus dans la tradition du commentaire, l'analyse lexicale peut également être utilisée à partir d'une réflexion modélisatrice qui précède sa mise en œuvre. On combine alors analyse de contenu et analyse lexicale. On peut le faire soit pour contrôler ou approfondir par la statistique lexicale les codifications résultant d'une analyse de contenu classique, soit pour automatiquement coder des contenus en utilisant la statistique lexicale.

6.1 Contrôler une analyse de contenu par la statistique lexicale

Une analyse de contenu menée de manière classique conduit à affecter à chaque réponse d'une interview ou fragment d'un discours les catégories thématiques que la lecture a permis de révéler. On peut alors considérer cette information comme une donnée de contexte et rechercher quel est le vocabulaire spécifique des réponses ou fragment selon qu'ils ont été affectés à telle ou telle catégorie thématique. Ceci revient en quelque sorte à mettre à jour, si ceux-ci existent, les ancrages lexicaux de l'analyse de contenu. Ce type d'analyse permet bien sûr de retrouver les règles explicites utilisées pour la codification (c'est une manière de vérifier que celles-ci ont correctement été appliquées), mais on peut également découvrir certains ancrages non explicités ou inconscients et élargir ainsi la compréhension des catégories utilisées.

La carte ci dessous (Figure 9) met en évidence le vocabulaire spécifique des réponses selon les éléments contrôlés identifiés dans la phase l'analyse de contenu décrite plus haut (cf Figure 1). Cette carte est le reflet du mode de lecture et des interprétations effectuées par le chercheur.

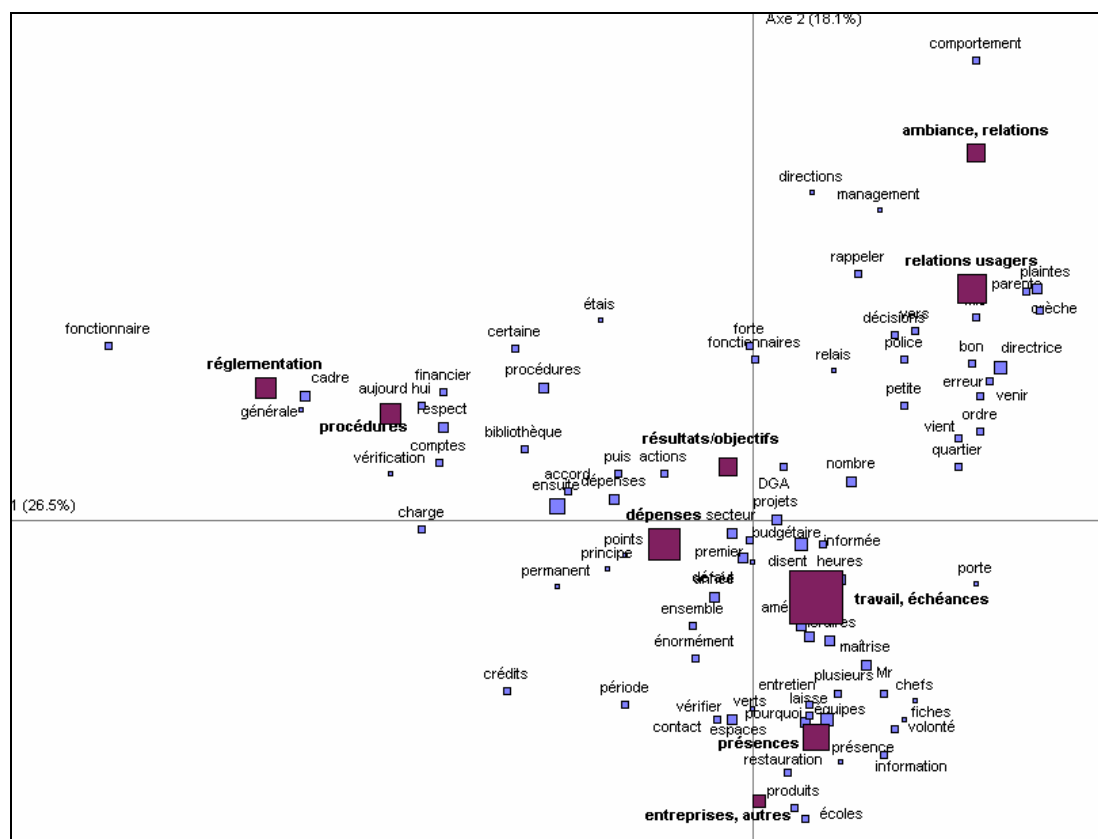


Figure 9 : Le vocabulaire spécifique des catégories utilisée pour caractériser la fonction contrôle³

6.2 Approfondir les analyses de contenu en les confrontant aux spécificités du lexique

On peut compléter cette approche en concentrant l'attention sur un thème particulier. La carte ci dessous met ainsi en évidence le fait que, selon le statut, l'expression du contrôle ne se fait pas un utilisant la même terminologie ce qui renvoie à des représentations très différenciées de la fonction.

³ Cette figure est un tri croisé entre le lexique spécifique du contrôle et la nature des acteurs. Voir annexe n°6, Tableau A.7.3. La dépendance est significative. $\chi^2 = 65,10$, $ddl = 45$, $1-p = 97,34\%$.

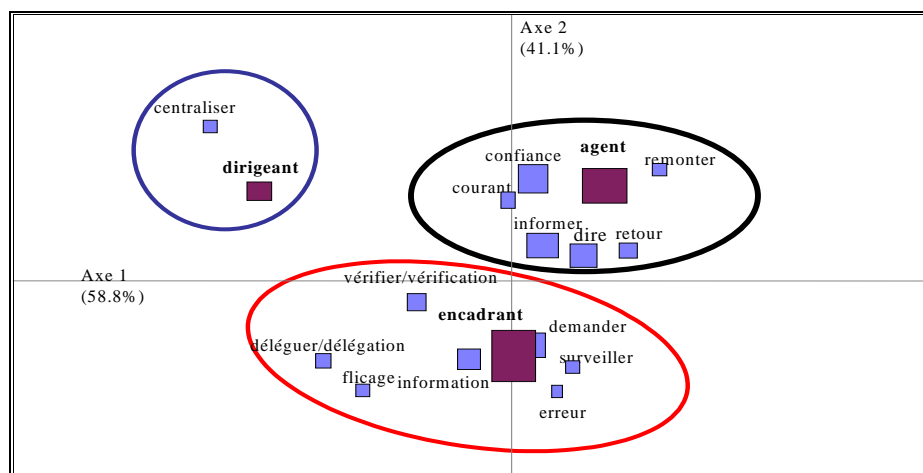


Figure 10 : Les objectifs du contrôle, de la recherche d'information à la surveillance⁴

Ce schéma montre que les représentations du contrôle varient selon la nature des acteurs.

Pour les dirigeants, l'exercice du contrôle est lié à la centralisation des procédures de décision.

Les encadrants possèdent des représentations complexes de leur rôle de contrôle. Leur lexique se rattache à la recherche d'informations tout autant qu'à la surveillance. Le contrôle est, pour eux, inextricablement mêlé avec une relation d'aide et de confiance. La conséquence de ces précautions est que **les agents** se sentent peu surveillés. Ils associent le contrôle au manque de confiance, et estiment massivement bénéficier de la confiance de leur encadrant.

6.3 Utiliser la statistique lexicale pour « mesurer des contenus » automatiquement

Cette voie n'a pas été utilisée pour la recherche que nous exposons ici et qui privilégie l'approche classique. Elle consiste à remplacer la phase d'analyse de contenu, décrite dans la première partie de cet article, par une analyse systématique du lexique et des termes utilisés pour les regrouper dans des champs lexicaux rassemblant tous les termes évoquant un des aspects du modèle retenu. Par exemple, en référence au modèle présenté au tableau 1 on aurait pu établir la liste de tous les termes évoquant la décision, l'information, les relations interpersonnelles... Pour chaque dimension du modèle, on établit ainsi un dictionnaire utilisé ensuite pour le calcul d'intensité lexicale. Cet indicateur mesure, pour chaque réponse ou fragment, la proportion des termes utilisés en provenance de chaque dictionnaire. On parvient ainsi à une mesure exprimant le poids de chacune des dimensions.

Cette manière de faire s'apparente à la lexicométrie. A partir de données textuelles on construit des mesures qui peuvent être utilisées exactement comme on le ferait à partir de données recueillies sur des échelles dans lesquelles le répondant indiquerait l'importance pour lui des dimensions du modèle. L'intérêt de procéder par le calcul d'intensités lexicales est

⁴ Cette figure est un tri croisé entre le lexique spécifique du contrôle et la nature des acteurs. Voir annexe n°6, Tableau A.7.3. La dépendance est significative. $\chi^2 = 65,10$, ddl = 45, 1-p = 97,34%.

d'éviter d'influencer le répondant en lui présentant les échelles qui révèlent le modèle. La difficulté réside dans la constitution des dictionnaires. On trouvera des exemples de cette approche dans les travaux de A. Kurk (2002) et J. Moscarola, V. Papatsiba et Y. Baulac (2002).

7. En conclusion

Les études dites qualitatives se caractérisent par la richesse du matériau textuel qu'elles permettent de récolter. Menées à partir de protocoles d'observations peu contraignants, ces recherches ont l'avantage de laisser aux répondants le libre choix de leur expression. Elles donnent au chercheur la faculté d'adapter sa quête aux terrains, en fonction des réactions qu'elle suscite. Il s'agit plus d'engager un processus de construction commun de description et de compréhension des phénomènes étudiés que de chercher à confirmer ou à infirmer une référence ou un modèle, donnés à priori. Cette attitude « respectueuse » conduit à la production d'un matériau empirique qui fait toute la difficulté de l'analyse. De la richesse mais aussi de la confusion et de l'ambiguïté des propos et expression recueillis, le chercheur doit extraire et dégager des lignes de forces pour parvenir à une construction plus générale. La modélisation se fait alors dans le temps de l'analyse.

Ainsi, le processus de recherche exposé ci avant a permis de faire émerger un nouveau modèle de représentation des rôles des cadres, alternatif aux modélisations existantes et susceptible de trouver une validité dans d'autres contextes organisationnels, publics ou privés. Ce modèle a permis de mettre en exergue la tension majeure existant entre les évolutions des différentes franges de l'encadrement des villes en contexte de modernisation. Si l'encadrement de catégorie A exerce de nouveaux rôles, les niveaux intermédiaires se caractérisent par l'exercice de rôles traditionnels amendés par l'intensité des relations. Ce résultat conduit alors à remettre en cause l'idée couramment admise selon laquelle on assisterait à une mutation de la gestion des villes [SOLDANO 00 ; BODIGUEL 98...], en faisant apparaître que les transformations observées se cantonnent essentiellement à la surface des organisations. Si la partie émergée de l'iceberg municipal donne à voir une renégociation des rapports entre l'administration et le politique, les 9/10èmes qui se trouvent sous la ligne de flottaison demeurent dans l'ombre et sont peu concernées par les transformations observées. La persistance de ces formes traditionnelles d'encadrement témoigne des blocages de l'innovation dans les villes.

Dans ce qui précède, l'émergence de nouvelles modélisations et de résultats contre-intuitifs est inséparable d'une attitude pragmatique. Celle-ci, à travers un pluralisme méthodologique, permet la triangulation des méthodes et des angles de vision et le passage permanent de l'exploration à la confrontation, de la vérification à l'élaboration, de la modélisation à l'analyse. Techniquement, ce pluralisme s'est notamment traduit par l'utilisation conjointe de démarches habituellement concurrentes :

L'analyse de contenu s'intéresse aux significations du texte, aux indications qu'il apporte sur le sujet étudié. Celles-ci sont dégagées par la lecture et organisées, catégorisées selon le modèle qui guide le chercheur. Il s'agit en quelque sorte de mettre en évidence dans quelle mesure les postulats du modèle sont présents dans le texte. La lecture permet d'identifier et de catégoriser, la statistique de dénombrer et ainsi de mesurer la proximité ou la distance entre le texte et le modèle.

L'analyse lexicale, considère un texte comme un paquet de mots dont on ignore l'organisation en phrases [LEBART 88], et permet des traitements statistiques sur la base du vocabulaire employé par les acteurs interrogés. Le vocabulaire est cependant replacé dans son contexte pour éviter les erreurs d'interprétation. Cette méthode assure une relative neutralité des traitements, ce qui lève une partie des réserves qui pèsent sur le caractère scientifique de la recherche qualitative. La rationalisation opérée par le chercheur porte sur les thèmes et sur les lexiques, mais ne porte pas directement sur le contenu du discours.

En effet, dans une perspective positiviste, la statistique lexicale met en œuvre des mécanismes indépendants de l'analyse du sens, évitant de soumettre la préparation du matériau textuel à l'interprétation du chercheur. La catégorisation intervient au regard de critères statistiques. Dans une perspective constructiviste, ces méthodes permettent la conservation de la richesse des réponses à travers le recours au langage.

Ces deux approches ne doivent pas être opposées, elles se complètent naturellement dans l'affrontement avec la complexité du langage, la confrontation aux connaissances existantes ou la production de nouvelles modélisations. Seuls le manque de temps ou la faiblesse des outils utilisés peut expliquer les clivages qui tendent encore à opposer ces deux approches.

Références

- [AUSTIN 70] AUSTIN J.L., *Quand dire c'est faire*, Paris, 1970
- [BENVENISTE 66] BENVENISTE E., *Problèmes de linguistique générale, I*, Paris, Gallimard, 1966
- [BENZECRI 81] BENZÉCRI J.P. & coll., *Pratique de l'analyse de données : linguistique et lexicologie*, Paris, Dunod, 1981
- [BERELSON 54] BERELSON B. V., « Content Analysis », *Hanbook of Social Psychology*, G. Lindsey (Ed.), Readings, Addison Wesley, 1954
- [BODIGUEL 96] BODIGUEL J. L., « Les fonctionnaires en proie au changement », in *Le service public en recherche, quelle modernisation?* sous la direction de GREMION C. et FRAISSE R., Paris, p. 91 à 101, 1996
- [BOLDEN 00] BOLDEN R. Moscarola J., « Bridging the Quantitative-Qualitative Divide. The Lexical Approach to Textual Data Analysis », *Social Science Computer Review*, Vol 19 N°4, Winter, 450-460, 2000
- [CHANAL 98] CHANAL V. Moscarola J., « Langage de la théorie et langage de l'action, analyse lexicale d'une recherche action sur l'innovation », *Acte des 4^{ème} journées JADT*, 1998
- [CULIOLI] CULIOLI, C. Fuchs, M. Pécheux, *Considérations théoriques à propos du traitement formel du langage, tentatives d'application au problème des déterminants*, Paris, Dunod
- [FAYOL 16] FAYOL H., *Administration industrielle et générale*", Paris, Dunod, 1916, réédition en 1970, 151 p.
- [FLOCH 88] FLOCH J.M., « The contribution of structural semiotics to the design of a hypermarket », *International Journal of Research in Marketing*, 4, 3, Semiotics and Marketing Communication Research, 217-232, 1988
- [GAVART-PERRET 98] GAVART-PERRET M.L. Moscarola J. (1998), « Enoncé ou énonciation ? Deux objets différents de l'analyse lexicale en marketing », *Recherche et application en marketing*, vol. 13, n°2.
- [GLASER 67] GLASER P. & STRAUSS A., *The Discovery of Grounded Theory, Strategies for Qualitative Research*, New York, Aldine Press, 1967
- [HUBERMAN 91] HUBERMAN A. M. et MILES M. B., *Analyse de données qualitatives. Recueil de nouvelles méthodes*, Bruxelles, De Boeck, 1991
- [JOHNSON-LAIRD 94] JOHNSON-LAIRD P.N., *L'ordinateur et l'esprit*, Paris, Odile Jacob, 1994
- [KURK 02] KURK A., « The importance of work values and social representations of young people with school and / or professional failure », *8yh ISSWOV Conference*, Warsaw, 2002
- [L'ÉCUYER 87] L'ÉCUYER R., « L'analyse de contenu : notion et étapes. » in *Les méthodes de la recherche qualitative sous la dir. De J-P. Deslauriers*, Montréal, Presses de l'Université du Québec, 1987
- [LEBART 88] LEBART L., SALEM A., *Analyse statistique des données textuelles*, Paris, Dunod, 1988

- [LEBART 94] LEBART L., SALEM A., *Statistique textuelle*, Paris, Dunod, 1994
- [MARCHAND 98] MARCHAND P., *L'analyse de discours assistée par ordinateur*, Paris, Armand Colin, 1998
- [MINTZBERG 84] MINTZBERG H., *The nature of managerial work*, Prentice Hall 1973, Traduction française, "Le Manager au quotidien, les dix rôles du cadre", Paris, Les Editions d'Organisation, 1984
- [MISCHLER 86] MISCHLER G., *Research Interviewing*, Havard University Press, 1986
- [MOSCAROLA 02] MOSCAROLA J., PAPATSIBA V., BAULAC Y., « Exploration sans a priori ou recherche orientée par un modèle : Contributions et limites de l'analyse lexicale pour l'étude de corpus documentaires », *Jadt*, 2002
- [MULLER 93] MULLER C., *Principes et méthodes de statistique lexicale*, Genève, Champion, 1993
- [PAPATSIBA] PAPATSIBA V. *Le séjour d'études à l'étranger : expériences et savoirs. Analyse des rapports d'étudiants français du programme Erasmus*. Sous la dir. Beillerot J., Thèse en cours.
- [PECHEUX 69] PECHEUX M., *L'analyse automatique du discours*, Paris, Dunod, 1969
- [POPPER 85] POPPER K., *Conjectures et réfutations*, Paris, Payot, 1985
- [REINERT 98] REINERT M., Quel objet pour une analyse statistique du discours ? Quelques réflexions à propos de la réponse Alceste. *Acte des 4^{ème} journées JADT*, 1998
- [REINERT 93] REINERT M., Quelques problèmes méthodologiques posés par l'analyse des tableaux énoncés x vocabulaire, *Acte des 2^{ème} journées JADT*, Université de Paris X, 1993
- [SEARLE 72] SEARLE J., *Les actes de langage*, Paris, Hermann, 1972