

I- Analyse de corpus (TD)

1- Les étapes de l'analyse des données (TD1)

- **Objectif** : Saisir les différentes étapes de l'analyse des données et leurs caractéristiques
- **Support** : Quivy, R. & Van Campenhoudt, L. et Marquet, J. (2011). *Manuel de recherche en sciences sociales*. Paris : Dunod.

1- Les étapes de l'analyse des données (Extrait 1, p.197)

- Quelles sont les étapes de l'analyse des données ?
- Quels sont les types d'analyse qui accompagnent chaque étape de l'analyse des données ?

(Extrait 1, p.197)

2. LES TROIS OPÉRATIONS DE L'ANALYSE DES INFORMATIONS

Fort différentes, les deux démarches illustrées ci-dessus impliquent de multiples opérations spécifiques tantôt à l'analyse quantitative, tantôt à l'analyse qualitative. Mais, dans tous les cas de figures, trois opérations sont toujours nécessaires et inévitables, même si, selon les méthodes, elles peuvent s'agencer les unes aux autres de manières diverses. Ces trois opérations sont : primo, la préparation des données ou informations ; secundo, la mise en relation des données ou informations ; tertio, la comparaison des résultats obtenus aux résultats attendus par hypothèse. Pour exposer chacun de ces points, on se placera alternativement dans le scénario d'une analyse quantitative et d'une analyse qualitative.

1.1. La phase de préparation des données (Extrait 2, pp.197-199)

1.1. 1. L'analyse quantitative

- En quoi consiste l'analyse quantitative liée à la préparation des données ?
- Quelles sont les exigences de l'analyse quantitative ?
- Quel concept est au cœur de la préparation des données des informations ?
- Quel traitement est accordé à ce concept dans le cadre de cette étape de l'analyse des données ?
- Quelles sont les différentes précisions apportées par l'auteur concernant le traitement de ce concept ?
- Quels sont les types de variables concernés par ce traitement lors de la phase de préparation des données ? Donnez les caractéristiques de chacune d'elle
- Quelle est l'importance des précisions à apporter aux variables ?

(Extrait 2, pp.197-199)

2.1. La préparation des données ou informations

a. Analyse quantitative : décrire et agréger

Pour tester une hypothèse, il faut d'abord exprimer chacun de ses deux termes par une mesure précise, afin de pouvoir examiner leur relation. Dans la préparation des données, la description et l'agrégation des données visent précisément à cela. Décrire les données d'une variable revient à en présenter la distribution à l'aide de tableaux ou graphiques, mais aussi à exprimer cette distribution par une mesure synthétique. Dans cette description, l'essentiel consiste donc à bien mettre en évidence les caractéristiques de la distribution de la variable.

Agréger des données ou des variables consiste à les regrouper en sous catégories ou à les exprimer par une nouvelle donnée pertinente. C'est ce que nous avons fait en construisant l'indice de croyance globale. Mais décrire une variable par une expression synthétique (la croyance moyenne chez les jeunes, par exemple), suit des procédures différentes selon le type d'information dont on dispose. Voici quelques précisions à ce sujet.

Les réponses-informations obtenues pour chaque indicateur lors de l'observation sont les données qui vont faire l'objet de l'analyse. Ces données manifestent les différentes modalités ou les différents états d'une variable. À titre d'exemples, allemand, belge et français sont des modalités ou états de la variable nationalité, comme 30 ans est une modalité ou un état de la variable âge. On appelle la variable toute caractéristique susceptible de prendre plusieurs modalités. Si ce n'est pas le cas, on a affaire à une constante et non à une variable. Lorsqu'un concept n'a qu'un seul indicateur, la variable s'identifie à l'indicateur (par exemple l'âge). Quand un concept est composé de plusieurs dimensions, le chercheur pourra souhaiter construire une variable résultant de l'agrégation des diverses dimensions (comme la croyance globale construite dans l'exemple précédent).

Les variables qualitatives sont soit nominales, soit ordinales. Une variable est dite nominale si ses modalités ne présentent pas d'ordre naturel, ou, ce qui revient au même, si tout ordonnancement des modalités reste purement arbitraire. C'est le cas de la variable nationalité pour laquelle on peut certes opérer un classement des modalités (les différentes nationalités) par ordre alphabétique, par exemple, mais dont le caractère arbitraire apparaît pour peu que l'on procède à une traduction qui change du coup l'ordre initial. Quand la variable nominale ne compte que deux modalités (par exemple le sexe), on parle de variable nominale dichotomique. Une variable est dite ordinale si ses modalités sont ordonnées mais sans que l'on ait une mesure de l'importance de l'écart entre deux modalités successives. C'est le cas d'une variable telle que le degré d'accord à l'égard d'une opinion et dont les modalités seraient par exemple : pas du tout d'accord, plutôt pas d'accord, plutôt d'accord, tout à fait d'accord. Les quatre modalités sont clairement ordonnées, mais rien n'est précisé quant aux distances qui les séparent les unes des autres et il serait sans doute hasardeux de prétendre que les modalités sont équidistantes. Enfin, il existe des variables, dites quantitatives, et dont les modalités ont une valeur numérique. Lorsque l'on mesure la taille d'un individu, que l'on dénombre le nombre d'enfants d'une famille, que l'on enregistre le pourcentage à l'examen de mathématiques d'un élève, etc., les différentes valeurs récoltées (1,80 m, 3 enfants, 71 %) n'ont rien d'arbitraire et indiquent bien plus qu'un ordre entre les modalités. Chaque réponse est ici signifiée par un nombre qui renvoie à une métrique permettant de mesurer des écarts entre les réponses des différents individus.

Ces précisions un peu techniques ne sont pas inutiles, car lors de la description et de l'agrégation des données ou des variables, il faut adopter les procédures de calcul adéquates. On ne traite pas les variables qualitatives de la même manière que les variables quantitatives. Pour décrire une variable par une expression synthétique, on utilisera par exemple les pourcentages si elle est nominale, la médiane si elle est ordinale et la moyenne si elle est

continue. Il faut y penser au moment de l'élaboration des instruments d'observation car il n'est pas indifférent que les réponses obtenues donnent à la variable un caractère nominal, ordinal ou continu. C'est à cela notamment que nous faisons allusion lorsque nous avons parlé d'anticipation des réponses lors de la formulation des questions. La description d'une variable et l'usage que l'on peut en faire varient selon qu'elle est nominale, ordinale ou continue. Ainsi, pour l'agrégation des variables, on ne peut regrouper des mesures de types différents sans passer par un dénominateur commun, ce qui conduit à une sérieuse perte d'information. Ceci est particulièrement important lorsqu'il faut agréger des variables pour reconstituer un concept et l'exprimer par une mesure synthétique. Analyser les relations entre les deux concepts d'une hypothèse devient difficile à partir du moment où l'on ne peut les exprimer par une mesure adéquate. Or c'est bien le but d'un travail scientifique

1.2. La phase de préparation des données (Extrait 3, p.199)

1.1. 2. L'analyse qualitative

- En quoi consiste l'analyse qualitative ?
- Quelles sont les opérations autour desquelles s'articule l'analyse qualitative ?
- Donnez les détails de chaque opération.
- Lors de cette phase, à quels outils le chercheur a-t-il recours ?
- Donnez des exemples de ces outils.

(Extrait 3, p.199)

b. Analyse qualitative : retranscrire et organiser

Pour tester les hypothèses à partir d'entretiens semi-directifs, il est nécessaire de commencer par les retranscrire. Sauf situation exceptionnelle, cette retranscription est intégrale, même si elle occupe des dizaines de pages, sans quoi des analyses fines sont très difficiles, surtout si elles visent à reconstituer la structure ou la dynamique d'ensemble du propos. La retranscription intégrale permet aussi d'éviter d'écarter trop vite de l'analyse des parties de l'entretien qui seraient jugées a priori inintéressantes, ce qui pourrait se révéler inexact au fil de l'analyse. Avant toute chose, il faut un matériau consistant et de qualité (en l'occurrence des entretiens) qui soit parfaitement restitué et entièrement disponible pour l'analyse. Quelles que soient les opérations auxquelles il sera procédé par la suite, il sera toujours possible de revenir à ce matériau de base et de s'y retrouver. Ensuite, il faut organiser ce matériau d'une manière qui permette son analyse. Les logiciels informatiques permettent de nombreuses manipulations, à commencer par les simples traitements de textes grâce auxquels on peut rechercher des mots et les souligner, déplacer des passages pour les mettre en parallèle avec d'autres, etc. Mais tout cela ne sera d'aucune utilité si le chercheur n'a pas une vision claire des principes selon lesquels il veut organiser son matériau.

Dans notre second exemple sur le Mouvement blanc, cette organisation est relativement sophistiquée. Reposant sur une grille d'analyse visant à reconstituer et formaliser le schéma actantiel, elle empiète déjà sur l'opération suivante.

1.3. La mise en relations des données ou informations (Extrait 4, pp.200-201)

- Autour de quelle opération s'articule l'analyse quantitative de la phase de mise en relation des données ?
- Expliquez en quoi consiste cette opération
- Est-ce que les procédures déployées dans le cadre de cette phase sont identiques pour tous les problèmes de recherche ? Justifiez votre réponse.

(Extrait 4, pp. 200-201)

2.2. La mise en relations des données ou informations

a. Analyse quantitative : l'analyse des relations entre variables

L'analyse des relations entre les variables constitue le deuxième passage obligé. Les variables à mettre en relation sont donc celles qui correspondent aux termes de l'hypothèse, c'est-à-dire soit les concepts impliqués dans les hypothèses, soit les dimensions, soit les indicateurs ou attributs qui les définissent. L'exemple ci-dessus illustre l'état de la relation entre l'âge et les croyances d'une part, entre celles-ci et le sexe d'autre part. Dans la pratique, on procède d'abord à l'examen des liens entre les variables des hypothèses principales et ensuite on passe aux hypothèses complémentaires. Celles-ci auront été élaborées dans la phase de construction, mais elles peuvent aussi naître en cours d'analyse à la suite d'informations inattendues. Rappelons que c'est ici qu'interviennent les variables-tests. Celles-ci sont introduites par les hypothèses complémentaires pour s'assurer que la relation supposée par l'hypothèse principale n'est pas fallacieuse, comme c'était le cas, dans l'exemple précédent, pour la relation entre le sexe et les croyances. En effet, grâce à l'hypothèse complémentaire introduisant l'activité professionnelle comme variable-test, on a pu découvrir que la relation entre le sexe et les croyances n'était pas fondée. En fait, elle n'est que le reflet de la relation entre l'activité professionnelle et les croyances. Ceci n'est qu'un cas particulier d'un problème général, celui de la pertinence des variables prises en considération. Si deux variables *A* et *B*, sans lien entre elles, sont étroitement dépendantes d'une autre variable *C*, toute variation de celle-ci entraînera des variations parallèles des deux premières. Si on ne connaît pas l'existence de *C*, la co-occurrence de *A* et *B* sera interprétée comme l'expression d'une relation directe entre elles, alors qu'elle n'est que le reflet de leur dépendance à l'égard de *C*. (...). Les procédures d'analyse ou d'agrégation des variables sont très différentes selon les problèmes posés et les variables en jeu. De plus, chaque méthode d'analyse des données implique des procédures techniques spécifiques et nous ne pouvons être plus précis ici sans nous engager dans des techniques trop particulières par rapport à nos objectifs.

1.4. La mise en relations des données ou informations (Extrait 5, p.201)

- Autour de quelles opérations s'articule l'analyse qualitative de la phase de mise en relation des données ?
- Quel est le principe sur lequel se base l'analyse de mise en relation des données ?
- Quel est l'effet de l'application de ce principe sur le corpus recueilli ?
- Relevez les techniques de recueil et d'analyse de corpus, citées dans le cadre de cet extrait.

- Quelles transformations sont observées au niveau des données ou du corpus recueilli ?

(Extrait 5, p.201)

b. Analyse qualitative : comparaisons et typologies

L'utilisation d'une grille d'analyse pour traiter des informations qualitatives, comme le contenu d'un entretien, permet de faire des liens à deux niveaux. Au niveau de chaque entretien, les relations entre ses éléments (par exemple entre les actants dans l'entretien de Marion) peuvent être mises en évidence, de manière à reconstituer la structure de l'entretien et, par-là, le système de représentation et d'action de la personne interviewée. Au niveau de l'ensemble des entretiens, des comparaisons peuvent être faites (entre Marion, Roland et les autres), des convergences et divergences peuvent être mises en évidence de manière à faire apparaître les logiques sociales implicites, qui, en l'occurrence, pourront être elles-mêmes saisies par l'outil de la typologie. Qu'il s'agisse de méthode quantitative ou qualitative, le principe de l'analyse est toujours de faire des liens (ou de montrer qu'il n'y en a pas), quels que soient les modalités et les termes utilisés : corrélation, co-occurrence, opposition, indépendance, convergence ou divergence, etc. Par les opérations diverses qui établissent des liens (statistiques, grilles d'analyse...), le matériau de départ change profondément de nature. À partir de réponses individuelles à un questionnaire, on construit des corrélations entre variables grâce à l'outil statistique ; à partir de propos individuels (ou de l'observation directe de comportements), on construit une structure de pensée et d'action grâce aux outils de la grille d'analyse et de la typologie. Dans l'analyse, le matériau de départ a, en quelque sorte, « fermenté » par la comparaison et la mise en relation de ses composantes, révélant des perspectives d'explication non perçues à l'origine de la recherche, jetant sur les phénomènes un regard plus éclairant, à la fois plus englobant et plus subtil. Quelle que soit la méthode utilisée, un bon chercheur est celui qui sait faire fermenter son matériau.

1.3. La comparaison des résultats et l'interprétation (Extrait 6, pp.201-202)

- Quel est l'objectif de l'observation et de l'analyse ?
- En quoi consiste la phase de comparaison et d'interprétation des résultats ?
- Peut-on considérer la recherche achevée une fois que les conjectures sont infirmées ?
- Quelles sont les conséquences de cette infirmation des conjectures sur les résultats de la recherche ?
- Commentez le schéma ci-dessous.

(Extrait 6, pp.201-202)

2.3. La comparaison des résultats observés avec les résultats attendus et l'interprétation des écarts

Chaque hypothèse élaborée lors de la phase de construction exprime les relations que l'on pense correctes et que devraient donc confirmer l'observation et l'analyse. Ainsi, dans l'étude du phénomène religieux, nous avons émis une hypothèse concernant la relation entre l'âge et

la croyance : les jeunes seraient moins croyants que les vieux. Les résultats attendus par hypothèse devaient donc être qu'aux âges les plus jeunes soient associés les taux de croyance les plus faibles et que dans les catégories les plus âgées on trouve les taux les plus élevés. Les résultats observés sont ceux qui résultent des opérations précédentes. C'est en comparant ceux-ci aux résultats attendus par hypothèse que l'on pourra tirer les conclusions. S'il y a divergence entre les résultats observés et les résultats attendus, ce qui n'est pas rare, il faudra soit examiner d'où viennent les écarts et chercher en quoi la réalité est différente de ce qui était présumé au départ, soit élaborer de nouvelles hypothèses et, à partir d'une nouvelle analyse des données disponibles, examiner dans quelle mesure elles sont confirmées. Dans certains cas, il sera même nécessaire de compléter l'observation.

Au départ de la recherche sur le Mouvement blanc, quatre hypothèses avaient été formulées. La première, sur laquelle a porté l'illustration, était celle d'une plainte complexe et diversifiée. Elle n'a pas été simplement confirmée par l'analyse ; la complexité et la diversité supposées ont été explorées et ont reçu un contenu synthétisé dans la typologie. À partir de là, de nouvelles hypothèses, plus fines et plus judicieuses encore, pourraient être formulées qui peuvent à leur tour être explorées, si l'on dispose du temps et des moyens nécessaires. Ces exemples montrent bien qu'il ne s'agit pas seulement de vérifier ou d'infirmer l'hypothèse. Qu'elle soit vérifiée ou infirmée, en tout ou en partie, si l'observation et l'analyse ont été bien menées, c'est un ensemble d'enseignements qui apparaissent, l'hypothèse n'étant que le fil conducteur qui va permettre de savoir de quel côté il est raisonnable et intéressant de chercher. Ce que l'on trouve alors est toujours bien plus nuancé que l'hypothèse qui n'en aura pas moins joué son rôle. L'interaction que nous venons d'évoquer entre l'analyse, les hypothèses et l'observation est représentée par deux boucles de rétroaction :

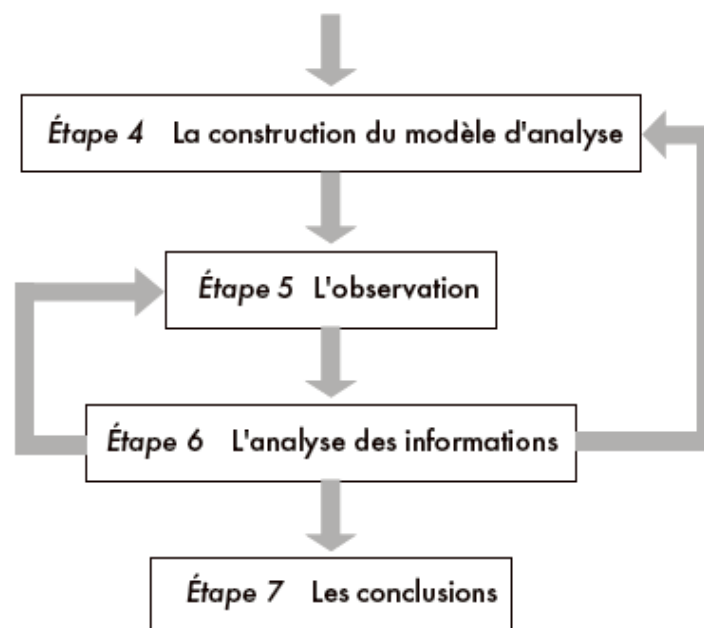


Figure 1 : Quivy, Van Campenhoudt et Marquet (2011, p.202)

2- Les méthodes d'analyse des données (TD2)

- **Objectif** : Saisir les différentes méthodes d'analyse des données et leurs caractéristiques
- **Support** : Quivy, R. & Van Campenhoudt, L. et Marquet, J. (2011). *Manuel de recherche en sciences sociales*. Paris : Dunod.

2- Les méthodes d'analyse des données (Extrait 1, p.203)

- Quelles sont les méthodes d'analyse des données présentées dans le cadre de l'extrait 1 ?
- Quel constat est fait à propos du recueil et de l'analyse des données ?

(Extrait 1, p.203)

PANORAMA DES PRINCIPALES MÉTHODES D'ANALYSE DES INFORMATIONS

La plus grande partie des méthodes d'analyse des informations relève de deux grandes catégories : l'analyse statistique des données et l'analyse de contenu. Ce sont donc elles qui seront présentées ici, avec quelques-unes de leurs variantes. Certaines méthodes présentées dans l'étape précédente comme méthodes de recueil des informations associent toutefois étroitement le recueil et l'analyse. C'est notamment le cas de certaines méthodes d'observation ethnologique et de l'entretien compréhensif. C'est également le cas de la méthode d'analyse en groupe qui associe directement les acteurs concernés à l'analyse elle-même. Les distinctions entre le recueil et l'analyse des informations ne sont donc pas forcément aussi nettes que la présente organisation des étapes peut le laisser supposer.

2. Les méthodes d'analyse des données (Extrait 2, pp.203-204)

2.1. L'analyse statistique

- Quelles sont les procédures statistiques auxquelles a recours le chercheur pour l'analyse des données ?
- Quels sont les éléments évoqués dans le cadre de l'extrait 2 et qui assurent la cohérence du processus de recherche ?
- Quelles sont les techniques d'analyse des données citées dans l'extrait 2 ? Qu'est-ce qui les caractérise ?
- Quelles sont les autres variantes de ces méthodes d'analyse présentées dans le cadre de l'extrait 2 ?
- Quels sont les objectifs fixés pour ces méthodes d'analyse des données ?
- Quels sont les avantages et les limites de ces méthodes d'analyse des données ?

(Extrait 2, pp.203-204)

3.1. L'analyse statistique des données

a. Présentation

Le développement et la démocratisation des ordinateurs et des logiciels ont profondément transformé l'analyse des données. La possibilité de manipuler rapidement des masses de données considérables a encouragé la mise au point de nouvelles procédures statistiques telles que l'analyse factorielle des correspondances qui permet de visualiser et d'étudier les liaisons entre plusieurs dizaines de variables en même temps. Parallèlement, la facilité avec laquelle les données peuvent être façonnées et présentées a incité de nombreux chercheurs à les étudier pour elles-mêmes, sans référence explicite à un cadre d'interprétation.

Présenter les mêmes données sous diverses formes favorise incontestablement la qualité des interprétations. En ce sens, la statistique descriptive et l'expression graphique des données constituent bien plus que de simples méthodes d'exposition des résultats. Mais cette présentation diversifiée des données ne peut remplacer la réflexion théorique préalable qui seule procure des critères explicites et stables pour le recueil, l'organisation et surtout l'interprétation des données et assure ainsi sa cohérence et son sens à l'ensemble du travail.

D'autre part, les chercheurs ne renoncent pas pour autant à l'usage de certaines techniques plus anciennes comme celle des tableaux croisés. Ces derniers sont souvent mal interprétés ou peu exploités en dépit ou, peut-être bien, à cause de leur apparente simplicité. Bref, les techniques les plus récentes voisinent normalement avec d'autres plus simples et plus anciennes qu'elles enrichissent mais ne remplacent pas forcément. Ces techniques graphiques, mathématiques et statistiques concernent principalement l'analyse des fréquences des phénomènes et de leur distribution ainsi que celle des liaisons entre variables ou entre modalités de variables.

b. Variantes

- Lorsque les données à analyser préexistent à la recherche et sont rassemblées par recueil de données documentaires, on parlera couramment d'analyse secondaire. Dans ce cas, le chercheur est plus ou moins limité dans ses analyses par le problème de la compatibilité des données entre elles et avec le champ de phénomènes qu'il souhaite étudier.
- Lorsque les données à analyser ont été spécialement récoltées pour les besoins de la recherche à l'aide d'une enquête par questionnaire, on parlera couramment de « traitement d'enquête ». Dans ce cas, les analyses sont généralement plus approfondies car les données sont en principe plus complètes et parfaitement standardisées au départ.
- Les méthodes d'analyse statistique des données sont également utilisées pour l'examen de textes. Il s'agit alors d'une méthode d'analyse de contenu qui est reprise plus loin sous ce titre.

c. Objectifs pour lesquels la méthode convient particulièrement

- Par définition, elle convient pour toutes les recherches axées sur l'étude des corrélations entre des phénomènes susceptibles d'être exprimés en variables quantitatives. Dès lors, elle s'applique généralement très bien aux recherches menées dans une perspective d'analyse causale. Mais ce n'est guère exclusif. Par exemple, dans le cadre d'un schème d'intelligibilité systémique, une corrélation entre deux variables sera interprétée, non comme une relation de causalité, mais comme une covariation entre composantes d'un même système qui évoluent conjointement.

• L'analyse statistique des données s'impose dans tous les cas où ces dernières sont recueillies à l'aide d'enquêtes par questionnaire où les questions sont dites « fermées », c'est-à-dire lorsque les répondants doivent choisir entre un petit nombre de réponses pré-formulées (de type oui/non, toujours/souvent/rarement/jamais). Il faut donc se référer aux objectifs pour lesquels cette méthode de recueil des données convient elle-même.

d. Principaux avantages

- La précision et la rigueur du dispositif méthodologique qui permettent aux pairs d'apprécier la démarche de la recherche.
- La puissance des moyens informatiques qui permettent de manipuler très rapidement un grand nombre de variables.
- La clarté des résultats et des rapports de recherche, notamment lorsque le chercheur met à profit les ressources de la présentation graphique des informations.

e. Limites et problèmes

• L'outil statistique a un pouvoir d'élucidation limité aux postulats (l'importance du nombre, des corrélations entre variables...) sur lesquels il repose, mais il ne dispose pas, en lui-même, d'un pouvoir explicatif. Il peut décrire des relations, des structures latentes, mais la signification de ces relations et de ces structures ne vient pas de lui. C'est le chercheur qui donne un sens à ces relations par le modèle théorique qu'il a construit au préalable et en fonction duquel il a choisi une méthode d'analyse statistique. (Cette question est bien traitée dans Franck R., 1994, *Faut-il trouver aux causes une raison ? L'explication causale dans les sciences humaines*, Lyon, Institut interdisciplinaire d'études épistémologiques.)

f. Méthodes complémentaires

En amont : l'enquête par questionnaire et le recueil de données statistiques existantes.

2. Les méthodes d'analyse des données (Extrait 3, pp.206-211)

2.2. L'analyse de contenu

- Sur quels types de données porte l'analyse de contenu ?
- Quelles sont les techniques qu'implique ce type d'analyse ?
- Quels sont les objectifs de l'analyse de contenu ?
- Quels sont les avantages de cette méthode d'analyse ?
- Quelles sont ses variantes ?
- Quelles sont les principales catégories de cette méthode et quelles sont leurs particularités ?
- Quels sont les objectifs fixés pour ces méthodes d'analyse des données ?
- Relevez les autres méthodes qui viennent compléter l'analyse de contenu.

(Extrait 3, pp.206-211)

3.2. L'analyse de contenu

a. Présentation

L'analyse de contenu porte sur des messages aussi variés que des œuvres littéraires, des articles de journaux, des documents officiels, des programmes audiovisuels, des déclarations politiques, des rapports de réunion ou des comptes rendus d'entretiens semi-directifs. Le choix des termes utilisés par le locuteur, leur fréquence et leur mode d'agencement, la construction du « discours » et son développement constituent des sources d'information à partir desquelles le chercheur tente de construire une connaissance. Celle-ci peut porter sur le locuteur lui-même (par exemple l'idéologie d'un journal, les représentations d'une personne ou les logiques de fonctionnement d'une association dont on étudierait les documents internes) ou sur les conditions sociales dans lesquelles ce discours est produit (par exemple un mode de socialisation ou une expérience conflictuelle).

Les méthodes d'analyse de contenu impliquent la mise en œuvre de procédures techniques relativement précises (comme le calcul des fréquences relatives ou des co-occurrences des termes utilisés, par exemple). Seule l'utilisation de méthodes construites et stables permet en effet au chercheur d'élaborer une interprétation qui ne prend pas pour repères ses propres valeurs et représentations. Contrairement à la linguistique, l'analyse de contenu en sciences sociales n'a pas pour objectif de comprendre le fonctionnement du langage en tant que tel. Si les aspects formels les plus divers du discours peuvent être pris en compte et examinés parfois avec une minutie et une patience de moine, ce n'est jamais que pour en retirer un enseignement qui porte sur un objet extérieur à eux-mêmes. Les aspects formels de la communication sont alors considérés comme des indicateurs de l'activité cognitive du locuteur, des significations sociales ou politiques de son discours ou de l'usage social qu'il fait de la communication.

L'analyse de contenu occupe une grande place dans la recherche sociale, notamment parce qu'elle offre la possibilité de traiter de manière méthodique des informations et des témoignages qui présentent un certain degré de profondeur et de complexité, comme par exemple les rapports d'entretiens semi-directifs. Mieux que toute autre méthode de travail, l'analyse de contenu (ou du moins certaines de ses variantes) permet, lorsqu'elle porte sur un matériau riche et pénétrant, de satisfaire harmonieusement aux exigences de la rigueur méthodologique et de la profondeur inventive qui ne sont pas toujours facilement conciliables. Les progrès récents des méthodes d'analyse de contenu ont certainement été encouragés par ce souci conjoint et largement partagé de rigueur et de profondeur. Ils ont été favorisés par les progrès de la linguistique, des sciences de la communication et de l'informatique. Pour ce qui concerne plus particulièrement la recherche sociale proprement dite, ils doivent beaucoup à Roland Barthes, Claude Lévi-Strauss et Algirdas Julien Greimas notamment.

b. Principales variantes

Il est courant de regrouper les différentes méthodes d'analyse de contenu en deux catégories : les méthodes quantitatives et les méthodes qualitatives. Les premières seraient extensives (analyse d'un grand nombre d'informations sommaires) et auraient comme information de base la fréquence d'apparition de certaines caractéristiques de contenu ou les corrélations entre elles. Les secondes seraient intensives (analyse d'un petit nombre d'informations complexes et détaillées) et auraient comme information de base la présence ou l'absence d'une caractéristique ou la manière dont les éléments du « discours » sont articulés les uns aux autres. Ces distinctions ne sont valables que très globalement : les caractéristiques propres des deux types de démarche ne sont pas aussi nette et plusieurs méthodes font aussi bien appel à l'un qu'à l'autre. Sans prétendre régler toutes les questions de démarcation entre les différentes méthodes d'analyse de contenu, nous proposons ici de distinguer trois grandes

catégories de méthodes selon que l'examen porte principalement sur certains éléments du discours, sur sa forme ou sur les relations entre ses éléments constitutifs. À l'intérieur de chaque catégorie, nous nous limiterons à l'évocation de quelques-unes des principales variantes.

■ *Les analyses thématiques*

Ce sont celles qui tentent principalement de mettre en évidence les représentations sociales ou les jugements des locuteurs à partir d'un examen de certains éléments constitutifs du discours. Parmi ces méthodes, on peut distinguer notamment :

– *l'analyse catégorielle* : la plus ancienne et la plus courante. Elle consiste à calculer et à comparer les fréquences de certaines caractéristiques (le plus souvent les thèmes évoqués) préalablement regroupées en catégories significatives. Elle se fonde sur l'hypothèse qu'une caractéristique est d'autant plus importante pour le locuteur qu'elle est fréquemment citée. La démarche est essentiellement quantitative ;

– *l'analyse de l'évaluation* : qui porte sur les jugements formulés par le locuteur. La fréquence des différents jugements (ou évaluations) est calculée mais aussi leur direction (jugement positif ou négatif) et leur intensité.

■ *Les analyses formelles*

Ce sont celles qui portent principalement sur les formes et l'enchaînement du discours. Parmi ces méthodes, on peut distinguer notamment :

– *l'analyse de l'expression* : qui porte sur la forme de la communication dont les caractéristiques (vocabulaire, longueur des phrases, ordre des mots, hésitations...) apportent une information sur l'état d'esprit du locuteur et ses dispositions idéologiques ;

– *l'analyse de l'énonciation* : qui porte sur le discours conçu comme un processus dont la dynamique propre est en elle-même révélatrice. Le chercheur est alors attentif à des données telles que le développement général du discours, l'ordre de ses séquences, les répétitions, les ruptures du rythme, etc.

■ *Les analyses structurales*

Les analyses structurales partent de l'idée que la fréquence à laquelle revient un mot ou un thème n'est pas, en soi, un indicateur suffisant pour décrypter les significations d'un texte (par exemple un entretien retranscrit). En effet, un texte n'est pas une somme d'éléments juxtaposés ; il les articule ou les « structure » d'une certaine manière et chaque élément occupe une place qui peut être centrale ou non, quelle que soit sa fréquence. Le plus important à étudier est donc la manière dont les éléments composant le texte sont agencés.

– *L'analyse des co-occurrences* : elle représente une modalité relativement simple de l'analyse structurale. Il s'agit d'examiner non plus les fréquences individuelles (ou occurrences) des thèmes mais bien leurs associations (ou co-occurrences). Savoir quels thèmes sont systématiquement associés dans les propos du répondant informe sur son univers culturel ou idéologique.

– *L'analyse structurelle proprement dite* : elle va plus loin car elle vise à mettre au jour la structure d'ensemble du texte, les principes sous-jacents d'organisation entre ses éléments, le système de relations entre eux. L'analyse structurale s'intéressera notamment aux couples d'oppositions entre éléments et à la manière dont ces couples d'oppositions s'articulent les uns aux autres pour former des structures d'oppositions plus englobantes des éléments du texte, aux règles d'enchaînement entre les éléments du texte, à sa trame et, plus largement, à tout ce qui structure le texte et y révèle un « ordre » sous des propos qui, au premier abord, peuvent sembler désordonnés voire décousus. La méthode élaborée par A.J. Greimas, utilisée dans le deuxième exemple, est une de ces variantes.

c. Objectifs pour lesquels la méthode convient particulièrement

Sous ses différentes modalités, l'analyse de contenu a un très vaste champ d'application. Elle peut porter sur des communications de formes très diverses (textes littéraires, émissions

télévisées ou radiophoniques, films, rapports d'entretiens, messages non verbaux, ensembles décoratifs, etc.). Sur le plan des objectifs de recherche, elle peut être notamment utilisée pour :

- l'analyse des idéologies, des systèmes de valeurs, des représentations et des aspirations ainsi que de leur transformation ;
- l'examen des logiques de fonctionnement d'organisations grâce aux documents qu'elles produisent ;
- l'étude des productions culturelles et artistiques ;
- l'analyse des processus de diffusion et de socialisation (manuels scolaires, journaux, publicités...);
- l'analyse de stratégies, des enjeux d'un conflit, des composantes d'une situation problématique, des interprétations d'un événement, des réactions latentes à une décision, de l'impact d'une mesure... ;
- la reconstitution de réalités passées non matérielles : mentalités, sensibilités...

d. Principaux avantages

- Toutes les méthodes d'analyse de contenu conviennent à l'étude de l'implicite.
- Elles obligent le chercheur à prendre beaucoup de recul à l'égard des interprétations spontanées et, en particulier, des siennes propres. En effet, il ne s'agit pas d'utiliser ses propres repères idéologiques ou normatifs pour juger ceux des autres, mais bien de les analyser à partir de critères qui portent davantage sur l'organisation interne du discours que sur son contenu explicite.
- Portant sur une communication reproduite sur un support matériel (habituellement un document écrit), elles permettent un contrôle ultérieur du travail de recherche.
- Plusieurs d'entre elles sont construites de manière très méthodique et systématique sans que cela ne nuise à la profondeur du travail et à la créativité du chercheur, bien au contraire.

e. Limites et problèmes

Il est difficile de généraliser ici car les limites et les problèmes posés par ces méthodes varient fortement de l'une à l'autre. Les différentes variantes ne sont guère équivalentes et ne sont donc pas interchangeables. Dans le choix de l'une d'entre elles, on sera particulièrement attentif aux points suivants.

- Certaines méthodes d'analyse de contenu reposent sur des présupposés pour le moins simplistes. Le record à cet égard appartient sans aucun doute à l'analyse catégorielle (voir plus haut). Il faut donc se demander si la recherche peut s'accommoder de ces limites. Si non, il faudra retenir une autre méthode ou en utiliser plusieurs conjointement. L'analyse catégorielle est d'ailleurs souvent mise utilement en œuvre en complément d'autres méthodes plus futées.
- Certaines méthodes, comme l'analyse évaluative, sont très lourdes et laborieuses. Avant de s'y engager, il faut être certain qu'elles conviennent parfaitement aux objectifs de la recherche et que l'on dispose du temps et des moyens nécessaires pour les mener à bien.
- Si l'analyse de contenu, prise globalement, offre un champ d'application extrêmement vaste, il n'en va pas de même pour chacune des méthodes particulières dont certaines ont, au contraire, un champ d'application très réduit. En réalité, il n'y a pas une mais des méthodes d'analyse de contenu.

f. Méthodes complémentaires

Les méthodes complémentaires sont des méthodes de recueil de données qualitatives et se situent donc normalement en amont de l'analyse de contenu qui portera sur les informations rassemblées.

Les plus couramment associées à l'analyse de contenu sont :

- surtout : les entretiens semi-directifs dont les éléments d'information conviennent particulièrement bien à un traitement par l'analyse de l'énonciation (qui en démontrera la dynamique) et l'analyse structurale ;

- le recueil de documents sur lesquels l'analyse de contenu portera ;
- plus rarement : les enquêtes par questionnaire pour le traitement des questions ouvertes.