

## Analyse de contenu

### 1. Vers une caractérisation de l'analyse de contenu.

#### 1.1. L'analyse de contenu : qu'est-ce que c'est ? qu'est-ce que ce n'est pas ?

L'**analyse** renvoie à l'idée de **décomposition en éléments**.

Dans le *Dictionnaire de l'Académie Française* (1822) l'article *Analyse* renvoie à celui d'*Analyse* : « terme didactique. La réduction, la résolution d'un corps dans ses principes. *Faire l'analyse d'une plante. Analyse chimique. Analyse par voie de Chimie.*

On dit aussi, *Faire l'analyse d'un discours*, pour le réduire dans ses parties principales, pour en mieux connoître l'ordre et la suite.

En mathématiques, on appelle *Analyse*, l'art de résoudre les problèmes par l'algèbre. *Être versé dans l'analyse.*

En logique, on appelle *Analyse*, la méthode de résolution qui remonte des conséquences aux principes, des effets aux causes. »

Dans *Le Robert*, il s'agit en premier sens « d'une opération intellectuelle consistant à décomposer un texte en ses éléments essentiels afin d'en saisir les rapports et d'en donner un schéma de l'ensemble »

Le **contenu** renvoie à ce qui se trouve dans un **contenant**.

L'**analyse de contenu** a principalement comme objet : le **contenu des contenus de communication humaine**. L'ensemble des **signifiants** utilisés par les êtres humains pour communiquer est susceptible d'être l'objet de l'analyse de contenu. Les signifiants les plus usuellement considérés sont les signifiants langagiers verbaux écrits ou oraux mis en œuvre dans les *textes* et dans le *discours*. Mais il importe de considérer aussi les signifiants non verbaux tels que des signifiants gestuels, iconiques, sonores, olfactifs, tactiles et bien d'autres codes sémiotiques qu'il conviendrait d'identifier. Les notions de phrase musicale, de texte musical sont usitées. Deux signifiants synonymes ont même *contenu*. La question du rapport "*contenu*" et "*sens*" apparaît à ce niveau. Tout *contenu* a-t-il un *sens* ? Le *contenu* se réduit-il au *sens* ?

La notion de **discours** renvoie à l'idée de production verbale, écrite ou orale, rapportée au contexte dans lequel elle se réalise. Si ce produit de l'action humaine qu'est le discours a été jusqu'à une époque encore récente l'objet de la **stylistique** ou de la **rhétorique**, il est devenu un objet de l'analyse de contenu actuelle. L'ouvrage de P. Marchand [20] *L'Analyse du Discours Assistée par Ordinateur* en constitue un exemple. **Analyser le contenu** consiste donc à ne s'intéresser non pas aux modes d'expression ou de manifestation dont les *textes* et les *discours* sont des formes particulières, mais essentiellement à *ce que veulent dire* ces manifestations expressives, *aux manières de voir le monde* qu'elles transportent. De là, la notion de contenu peut être étendue aux résultats du faire. L'analyse de contenu peut alors être appliquée à des *choses faites*, des *dispositions d'objets*. C'est s'intéresser en premier lieu au **sens manifeste** et puis au-delà à un **sens latent**.

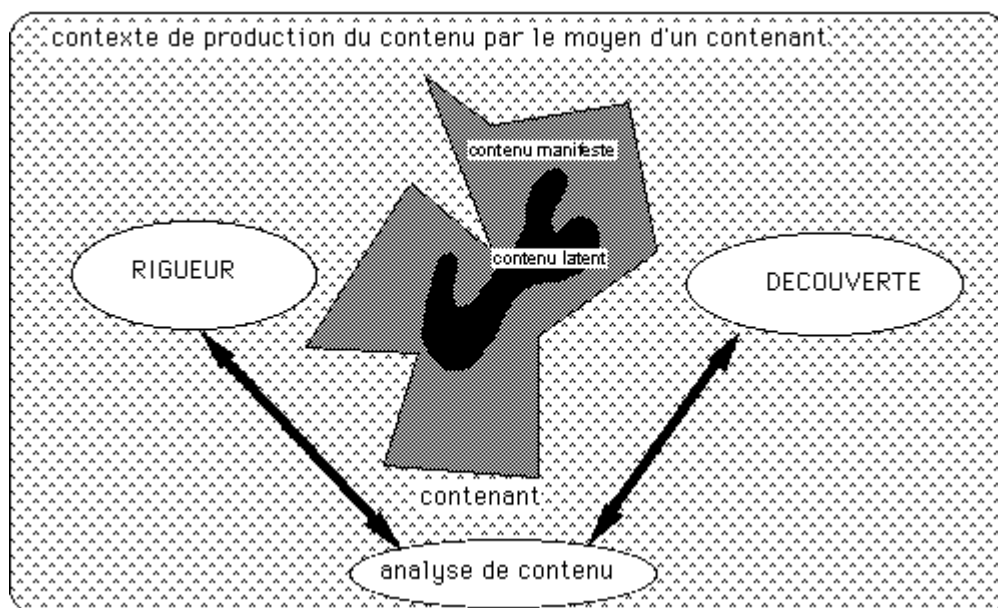
En puisant dans l'histoire de l'analyse de contenu, nous prendrons la définition<sup>1</sup> due à B. Berelson dans les années 50, pour fixer cette notion : « **L'analyse de contenu est une technique de recherche pour la description objective, systématique et quantitative du contenu manifeste de la communication** ».

Pour resituer le propos, nous rappelons que « . Dans l'**analyse quantitative** ce qui sert d'information est la **fréquence** d'apparition de certaines caractéristiques de contenu. Dans l'**analyse qualitative** c'est la **présence\_ou l'absence\_** d'une caractéristique de contenu donnée ou d'un ensemble de caractéristiques, dans un fragment de message, qui est prise en considération. »<sup>2</sup>

Aujourd'hui l'**analyse de contenu** désigne « **un ensemble de techniques d'analyse des communications visant, par des procédures systématiques et objectives des descriptions de contenu des messages, à obtenir des indicateurs (quantitatifs ou qualitatifs) permettant l'inférence de connaissances relatives aux conditions de production/réception (variables inférées) de ces messages** »<sup>3</sup>

En souhaitant ne pas s'en tenir à la *simple lecture du réel toujours séduisante*, le détour par des méthodes (chemin) d'analyse de contenu traduit deux visées :

- *Dépasser l'incertitude issue de la première lecture* : ce que je crois percevoir dans le contenant y est-il effectivement contenu ? Le résultat de cette perception est-il partagé par d'autre ? Dans quelle mesure la *lecture* du contenant est-elle valide et généralisable ?
- *Enrichir la compréhension par une seconde lecture* : n'y a-t-il pas un *sens latent* par delà la surface du contenant qui donne à percevoir le *sens manifeste* ?



Deux fonctions majeures peuvent être attribuées à l'analyse de contenu, et qui la soumettent à une tension entre les deux pôles DECOUVERTE et RIGUEUR :

- *la fonction heuristique* (découverte) de l'analyse de contenu exploratoire *pour voir* .
- *la fonction d'administration de la preuve* (validation, confirmation ou infirmation) de l'analyse *pour prouver* .

<sup>1</sup> Bardin [ 6 ] p 21

<sup>2</sup> Bardin [ 6 ] p 24

<sup>3</sup> Bardin [ 6 ] p 47

L'analyse de contenu : du sur-mesure et non du prêt-à-porter, reste un ensemble de techniques et outils mis à la disposition du chercheur à partir desquels il doit, à chaque circonstance de travail, élaborer sa propre méthode d'analyse adaptée aux matériaux, à la problématique à laquelle il se confronte, à la nature des contenants qu'il a recueillis, à ses visées, à ses compétences dans le domaine des théories de la communication, de la sémiotique et des N.T.I.C. appliquées à la recherche et à la formation, et même à sa propre capacité d'empathie.

Citons quelques exemples qui illustreront la mise à l'œuvre de cette capacité d'invention du chercheur :

- démasquer l'axiologie sous-jacente à des manuels scolaires
- démasquer la place et le rôle du genre dans les manuels scolaires de mathématiques
- retrouver l'inconscient collectif derrière l'incohérence apparente des graffiti tracés sur les murs dans les lieux
- radiographier les réseaux de communications formelles et informelles d'un établissement scolaire à partir des notes de services
- démasquer les conceptions pédagogiques des enseignants à partir des appréciations écrites dans les bulletins scolaires trimestriels.
- identifier les conceptions des parcours diversifiés en collège à partir des réponses à des questions ouvertes adressées aux principaux.
- construire une typologie des collèges à partir des réponses à un questionnaire comportant aussi des questions ouvertes, adressé aux principaux.
- expliciter le sens donné aux parcours diversifiés par les professeurs à partir d'un entretien semi-directif
- identifier les effets d'une formation pour des personnels dans le domaine de la gérontologie à partir de réponses à des questions ouvertes et de comptes rendus réalisés par les formateurs.

Les domaines d'application restent très vastes et variés. Le développement de l'analyse est basé à la fois sur l'effort de rigueur et sur l'inventivité de chaque chercheur.

## 1.2. Organisation globale de l'analyse de contenu

Elle se déroule globalement selon trois grandes phases :

- *phase n°1* : période où l'intuition est le processus le plus sollicité. Durant celle-ci le chercheur s'efforce d'opérationnaliser et à systématiser ses idées initiales afin de parvenir à établir un plan d'analyse. Ce programme établi doit être précis et flexible, c'est à dire susceptible d'intégrer de nouvelles procédures tout au long du parcours. Il s'agit pour le chercheur de procéder à :

- *choix des documents à soumettre à l'analyse*

A l'issue du choix des documents, le chercheur aboutit à un **corpus**, ensemble des documents retenus pour être analysés. La **constitution de ce corpus** est régie par des règles dont les principales sont :

-Règles de l'exhaustivité et de la non-sélectivité  
prendre en compte tous les éléments présents dans le champ du corpus

-Règle de la représentativité  
recourir à des procédures d'échantillonnage

-Règle de l'homogénéité

retenir des documents qui obéissent à des critères précis sans présenter une singularité excessive pour des caractéristiques non considérées par les critères de choix

-Règle de pertinence

retenir des documents adéquats en tant que source d'information à l'objectif de recherche qui suscite l'analyse de leur contenu.

- *formulation des conjectures et des objectifs*
  - *élaboration des indicateurs en vue de l'interprétation finale.*
- *phase n°2* : période d'exploitation du matériel,
- *phase n°3*: période du traitement des résultats obtenus et de l'interprétation

### 1.3. Exploitation du matériel

#### 1.3.1. *Le codage*

Cette procédure de traitement correspond à une transformation des données brutes fondée sur le découpage et l'agrégation. " Le codage est le processus par lequel les données brutes sont transformées systématiquement et agrégées dans des unités qui permettent une description précise des caractéristiques pertinentes du contenu " O.R. Holsti<sup>4</sup>

- choix des unités pour/par le découpage
- choix des règles de dénombrement pour/par l'énumération
- choix des catégories pour/par la classification et l'agrégation

Quels sont les unités d'enregistrement, c'est à dire les unités de signification à coder ?

Quels sont les unités de contexte qui servent d'unités de compréhension pour coder les unités d'enregistrement ?

#### 1.3.2. *La catégorisation*

Cette procédure est une classification des unités d'enregistrement constitutives de l'ensemble étudié préalablement différenciées puis regroupées par attribut selon des critères préalablement définis.

Deux étapes :

- inventaire des unités d'enregistrement (unité statistique)
- classification selon des critères explicites et précis posé a priori ou a posteriori

Les critères théoriques de qualité d'une catégorisation sont :

- exclusion mutuelle
- homogénéité
- pertinence
- objectivité
- fidélité
- productivité

#### 1.3.3. *L'inférence*

Sur quoi peut porter l'inférence (raisonnement contrôlé) en analyse de contenu ?

Déjà par la prise en compte d'un certain nombre de facteurs impliqués dans une situation de communication :

- l'émetteur
- le récepteur
- le message
- le code
- la signification
- le médium

---

<sup>4</sup> cité par Bardin [ 6] p 134

Exemple de caractérisation de l'analyse de contenu :

Dans *l'Encyclopaedia Universalis*, nous avons trouvé ces caractéristiques :

### ANALYSE DE CONTENU

© 1998 Encyclopædia Universalis France S.A

Méthode destinée à permettre un traitement scientifique des documents et définie par B. Berelson, qui en a été l'instigateur avec H. Lasswell, comme «une technique de recherche pour la description objective, systématique et quantitative du contenu manifeste des communications». Dans cette technique d'analyse des textes, on peut distinguer les trois phases suivantes: d'abord, une fragmentation du texte, chaque fragment devenant une suite d'analyse; en deuxième lieu, le codage de chaque unité d'analyse à l'aide d'une grille d'analyse, laquelle définit un certain nombre de catégories. Le codage consiste à faire entrer chaque unité d'analyse dans l'une des catégories. L'ensemble des catégories définissant la grille d'analyse doit être construit de telle façon que chaque unité d'analyse puisse être associée à une catégorie et une seule. Pour prendre un exemple simple, dans une analyse de contenu de documents de presse, on peut choisir comme unité d'analyse l'article de journal et comme ensemble de catégories un ensemble de rubriques telles que: politique intérieure, politique étrangère, éducation, religion, sports, spectacles, critique littéraire ou de spectacles, faits divers. La troisième phase consiste en une analyse statistique des données ainsi recueillies. Les deux premières phases ne sont pas indépendantes: le choix de l'unité d'analyse et la construction de la grille d'analyse vont de pair: selon le type d'unité choisie la grille sera différente. La fragmentation d'un texte en unités d'analyse est un problème particulièrement délicat. On choisit parfois des unités d'analyse définies objectivement: par exemple, la phrase, la proposition. Cela peut se justifier dans certains cas, mais ne résout pas le problème de façon générale. Deux unités objectivement équivalentes peuvent présenter des degrés de complexité fort différents du point de vue de la signification et devraient relever de traitements différents. Idéalement, l'unité d'analyse devrait coïncider avec l'unité de signification. Bien que les recherches théoriques sur l'analyse sémantique aient fait beaucoup de progrès ces dernières années, il n'existe pas actuellement de théorie générale d'analyse de la signification qui permette de donner à ce terme un statut précis et une définition opérationnelle.

On peut facilement soutenir que l'analyse de contenu n'est pas fondée théoriquement. Il reste qu'elle n'a pas attendu les progrès de la linguistique pour se développer et que ses résultats apparaissent relativement cohérents. En fait, des solutions empiriques, difficiles à justifier dans le cadre d'une théorie générale, ont fait dans la pratique la preuve de leur intérêt. Une solution souvent adoptée consiste à distinguer un ensemble de thèmes et à repérer dans le texte le passage d'un thème à un autre. Il y a, bien entendu, beaucoup de latitude relativement à ce que l'on pourra appeler un thème: ce pourra être un ensemble de significations très complexes ou, au contraire, quelque chose d'assez élémentaire. En fait, on ne peut pas définir à un niveau général ce qu'est un thème. Il importe seulement de souligner qu'il existe de multiples analyses possibles d'un texte; cela dépend du point de vue où l'on se place. Pour qu'une analyse de contenu soit possible, il faut définir préalablement une hypothèse de travail, prévoir un objectif d'analyse, qui permet de déterminer ce qui, de ce point de vue, est important et intéressant dans le texte. Alors seulement on peut se poser le problème du découpage du texte et de la définition des thèmes. On peut, pour des textes présentant des caractéristiques bien définies, adopter une analyse plus structurale: dans le cas de récits par exemple, on pourra s'intéresser au type d'action réalisée, à l'auteur de l'action, à l'objet de celle-ci, au type de résultat de l'action, etc.

La forme que prend l'analyse de contenu varie considérablement en fonction du type de texte considéré et de la problématique choisie. L'analyse statistique appliquée aux

données obtenues après codage peut aussi varier considérablement du point de vue de la complexité des techniques utilisées. Elle peut se réduire à un simple décompte des fréquences d'apparition de chaque catégorie. Elle peut rechercher la structure des thèmes en partant par exemple de la probabilité de passer d'un thème à un autre. Il n'existe pas de techniques statistiques propres à l'analyse de contenu; en fait, on en utilise diverses sortes, qui vont des plus élémentaires aux plus complexes.

L'analyse de contenu est utilisée dans des domaines très divers: analyse de documents de presse, de récits, d'entretiens, de questionnaires d'enquêtes, de réponses de sujets à des tests psychologiques, de réponses d'élèves à des épreuves de contrôle des connaissances. On notera que l'analyse de contenu ne diffère pas, dans son principe, des techniques d'observation systématique, qui supposent, elles aussi, une délimitation des éléments à observer, les constructions d'une grille d'observation et une analyse statistique des données ainsi recueillies.

## 2. Quelques types d'analyse de contenu

### 2.1. Analyse de l'évaluation [6]

Voir [ 6] page 208-222

*The evaluative assertion analysis* (EAA), l'**analyse d'assertion évaluative** a été mise au point en 1956 par Osgood, Saporta et Nunnally. Son but est de mesurer les attitudes du locuteur à l'égard d'objets au sujet desquels il s'exprime.

Cette perspective s'appuie sur le présupposé que le langage représente, reflète directement celui qui l'utilise. En conséquence, le chercheur peut se contenter d'indicateurs manifestes explicitement contenus dans la communication pour faire des inférences au sujet de la source d'émission.

### 2.2. Analyse de l'énonciation [6]

Voir aussi analyses pragmatiques.

Cette analyse :

- s'appuie sur une conception de la communication comme processus et non produit,
- fonctionne par le détour des structures et des éléments formels.

Cette analyse examine l'acte de parole en considérant que c'est au moment de la production du discours que le sens s'élabore, le langage ne constitue pas le reflet intégral de l'opinion des sujets interrogés. Ce processus d'élaboration se réalise sur un fond d'idéologies, d'incohérences, de contradictions, de défenses, de résistances, de rationalisations, d'ambivalences, etc.

Cette approche prend comme objets : les agencements et la logique du discours, le repérage des éléments formels atypiques, des ruptures telles que les omissions, les répétitions, les silences, les lapsus, les figures de rhétorique, etc.

### 2.3. Analyse propositionnelle du discours (A.D.P.) [6]

L'objectif est l'identification de l'univers de références des acteurs sociaux. Comment et au travers de quelle structure argumentative s'expriment les enjeux et les actions des acteurs ?

Cette analyse vise à établir des schémas argumentatifs qui symbolisent le rôle des acteurs. Ses auteurs Ghiglione et alii [30], proposent une méthode en plusieurs étapes :

- repérages des *référents-noyaux* (RN) en tant qu'objets thématiques à forte valeur référentielle
- découpage du texte en *propositions*, phrases qui qualifient expliquent les RN
- description des *propositions* par les *schémas argumentatifs* en tant traduction simplifiée des *propositions*
- délinéarisation du texte par regroupement des *propositions* caractérisées par un RN pour parvenir à identifier l'organisation cognitive et affective sous-jacente.
- Interprétation

La proposition est définie comme « un segment de texte à forme généralement prédicative : sujet, verbe, complément. »



## 2.4. Analyses de l'expression [6]

Ces analyses se fondent sur des indicateurs formels non sémantiques tels que des indicateurs lexicaux, stylistiques ou syntaxiques. Par exemple<sup>5</sup> : richesse ou pauvreté du vocabulaire (TTR = L/O), longueur des phrases, nombre de mots exprimant une caractéristique, (DRQ), quotient de genre grammatical, PNAQ, etc.

Elles se réfèrent au présupposé théorique qu'il existe une correspondance entre le type de discours et les caractéristiques du locuteur ou de son environnement. Ainsi des traits personnels, l'état du locuteur ou sa réaction à une situation sont considérés comme des variables transformatrices du discours dans sa forme comme dans son contenu.

## 2.5. Analyse des relations [6]

- **Analyse des cooccurrences** : « L'analyse de contingence ne se demande pas quelle est la fréquence d'apparition d'une forme signifiante donnée dans chacune des diverses parties d'un texte, mais quelle est sa fréquence d'apparition avec d'autres unités signifiantes » écrit Osgood.
- **Analyse structurale**
- **Analyse du récit**
- **Analyse du discours**

## 2.6. Dans le cadre de l'analyse du discours [20]

### 2.6.1. *Analyses para-verbales*

Il s'agit d'un premier niveau d'analyse lié à la mesure des aspects connexes à la verbalisation. Le chercheur s'appuie sur le fait que quelque chose est dit, et cela d'une certaine manière, mais il ne considère pas effectivement ce qui est dit.

Ces analyses interviennent dans les études de l'impact de la verbalisation sur les processus mentaux. ( étude sur la dissonance cognitive, les conversations, ...)

Les *analyses phonologiques* vont prendre en compte ce qui intervient dans l'activité complexe et coordonnée de parler. Parler met au moins sept systèmes musculaires en jeu :

- Respiratoire ( expulsion de l'air des poumons)
- Phonatoire (action des cordes vocales pour produire un son)
- Pharyngal (contrôle de l'articulation)
- Vélopharyngal (contrôle de la nasalité)
- Lingual (articulation des consonnes et des voyelles)
- Labial (contrôle des lèvres)
- Mandibulaire (contrôle de la mâchoire)

La prise peut alors être décrite à l'aide de la qualité, de la fréquence, de l'intensité et de la durée des sons émis.

Les performances vocales sont à considérer sous trois facettes :

- la *voix quasi permanente* du locuteur qui permet de le reconnaître
- la *tonalité* de la voix
- les *réalisations phonétiques d'unités linguistiques*

---

<sup>5</sup> voir Bardin [ 6] pp 256-259

Ainsi dans le cadre d'entretiens, le chercheur peut travailler sur l'analyse des intonations, des prononciations, des liens que nous avons tendance à établir entre la voix et les traits physiques.

*Par exemple :*

Des études sur l'évaluation des élèves par leurs enseignants ont montré qu'une voix lente et grave était jugée prédictive d'un échec scolaire.

### **2.6.2. Analyses lexicales**

Analyse fondée sur l'analyse du lexique

### **2.6.3. Analyses morpho-syntaxiques**

Analyse fondée sur les catégories morpho-syntaxiques a priori et sur les relations logiques des formes entre elles et non simplement sur les occurrences ou les cooccurrences. Ce sont les relations syntagmatiques des formes dépendant des fonctions des différentes catégories de termes (articles, noms, verbes, adjectifs, adverbes, plus généralement les connecteurs, etc. ) qui sont étudiées pour rechercher les significations et soutenir les interprétations. Par exemple le travail avec les conjonctions : mais, ou, et, donc, or, ni, car; les adverbes : ainsi, aussi, donc.

### **2.6.4. Analyses sémantiques**

Analyse qui se fonde sur le sens (sémantique). Deux approches peuvent être repérées :

- classifier les énoncés d'un texte pour en synthétiser le contenu et en livrer une compréhension a posteriori sur la base de conjectures initiales en amont du traitement,
- établir des catégories générales dans la langue pour les appliquer ensuite à plusieurs corpus.

### **2.6.5. Analyses pragmatiques**

La pragmatique du discours constitue une branche de la linguistique concernant le « rapport du discours aux circonstances les plus générales de production d'un sens communicable »<sup>6</sup> (Jacques, 1979) [ 20 ]

La pragmatique considère que tout discours permet d'accomplir trois fonctions essentielles :

- dire quelque chose, ou dire comment sont les choses
- faire quelque chose ou dire ce qu'on pense des choses
- induire une conduite ou altérer l'état de l'auditeur.

Ces analyses prennent alors les actes de paroles comme objet.

## **3. Les produits de l'entretien et des questions ouvertes soumis à l'analyse de contenu.**

### **3.1. Analyse de contenu de réponses à une ou plusieurs questions ouvertes d'un questionnaire**

#### **3.1.1. Techniques manuelles**

Il s'agit de techniques s'appuyant des procédures de *post-codage* consistant à :

- construire une batterie de *modalités* à partir d'un sous-échantillon de réponses,
- codifier ces modalités pour les utiliser comme résultats d'une ou plusieurs questions fermées.

---

<sup>6</sup> *Dialogiques: recherches logiques sur le dialogue*, Paris, PUF

L'idée est de transformer une question ouverte en une ou plusieurs questions fermées. Elle s'avère utile, efficace et sans inconvénient majeur, lorsqu'il s'agit de questions sollicitant des réponses simples, typées et couvrant un éventail restreint de possibilités. Dans le cas contraire, les obstacles qui se dressent, deviennent plus contraignants. En tout état de cause, nous pouvons noter quelques-uns des obstacles auxquels le chercheur doit tenter de faire face.

#### **3.1.1.1. La subjectivité dans la procédure de post-codage :**

L'attribution d'une modalité à une réponse soulève des questions et la décision d'affectation peut s'avérer très discutable.

#### **3.1.1.2. La mutilation de la forme**

La qualité de l'expression, le registre du vocabulaire disparaissent

#### **3.1.1.3. L'appauvrissement du contenu**

Les réponses composites, floues, complexes, diverses sont écrasées. En conséquence de quoi, le chercheur perd la valeur heuristique qu'apportait l'ouverture des réponses libres.

Le traitement des réponses rares

Ces réponses rares, originales parfois, souvent peu claires en première lecture sont affectées à la modalité « autres réponses ». De là elles se retrouvent dans un univers d'une grande hétérogénéité et perdent complètement leur valeur opératoire. Cependant il est remarquable que ces réponses rares peuvent parfois comporter des significations d'un grand intérêt pour les interprétations des résultats. Le traitement de type *a priori* auquel conduit le post-codage s'avère inadéquat pour ce genre d'information dont la richesse se révèlera *a posteriori*.

### **3.1.2. Technique assistée par ordinateur**

Ces techniques s'appuient sur l'analyse lexicale. Les outils statistiques assistés des outils informatiques sont alors sollicités au maximum de leur possibilité. Certes il ne s'agit d'investir dans des attentes qui les dépassent. Ainsi l'ordinateur traite à partir d'un programme (logiciel) des quantités astronomiques d'unités informationnelles qui le rendent plus performant que l'être humain dans des tâches répétitives, mais il ne possède nullement d'esprit-cerveau qui lui permette de **donner du sens**. Il peut aider efficacement l'être humain à conduire son processus d'interprétation mais il n'interprète pas de lui-même. Il reste une machine au service des êtres humains. Il convient de ne pas se laisser piéger par une pensée magique qui investit dans l'ordinateur des capacités qu'il n'a pas. Certes d'aucuns voudraient nous le laisser croire, surtout quand cela leur sert pour justifier leurs propres intérêts en profitant de l'ignorance d'autrui. Mais ici nous n'utiliserons l'ordinateur que pour nous aider dans des tâches de repérage, de dénombrement et de réalisation d'algorithme de calcul, très fastidieuses et parfois même impossible pour l'être humain. Ce n'est nullement l'ordinateur qui nous fournira l'interprétation que nous choisirons.

#### **3.1.2.1. Les unités de base.**

L'unité de base est la **forme graphique**, c'est à dire une suite de caractères graphiques non-délimiteurs, entourée par des caractères graphiques délimiteurs. Pour ce qui nous concerne dans les traitements de réponses ouvertes : une forme graphique pourra être un mot constitué de lettres séparé du mot précédent par un espace et du mot suivant par un point : le chat.

Cette idée de travailler sur la forme graphique pour réductrice qu'elle paraît puisque dans un premier temps, elle met de côté le sens, apporte un formidable outil de traitement de

données. Ainsi le traitement ne se limite pas à un idiome particulier ou à un lexique spécifique.

Le chercheur va alors décider de ce à quoi renvoie une forme graphique selon les objectifs de sa recherche. Certes cette approche se paie par deux conséquences qu'il convient de contrôler au moment même de la saisie des formes graphiques dans un document textuel.

**L'ambiguïté** issue du fait qu'une même forme graphique peut renvoyer à deux ou plus mots dans un lexique d'une langue donnée : il livre un lit / il lit un livre.

La "**redondance non pertinente**" car les mots LIVRE, Livre, LiVre vont être considérés comme trois formes graphiques distinctes.

Les **segments répétés** constitués d'une **suite de formes graphiques** qui apparaissent au moins deux fois dans le **corpus** étudié.

Partant ces deux types d'unités statistiques, nous nous intéressons à leurs profils de fréquences, c'est à dire aux vecteurs dont les composantes sont les fréquences de chacune des unités utilisées par un individu ou par un groupe.

### **3.1.3. Les traitements du corpus des formes graphiques**

- La **numérisation** qui consiste à attribuer un numéro à chaque forme graphique et le stocker dans un dictionnaire de formes propres au corpus de l'étude qui est conduite.
- La **lemmatisation** qui consiste à déclarer équivalentes des formes graphiques distinctes.
- La constitution de **tableaux lexicaux** :
  - un premier tableau R ( tableau des pointeurs) est construit de la façon suivante : les lignes correspondent chacune à une réponse à une question ouverte et les colonnes, en nombre égal aux nombres de formes graphiques de la plus longue réponse, indique les formes graphiques utilisées, chaque cellule contient le numéro de la forme graphique dans le dictionnaire
  - un second tableau T est le tableau de contingence [ Individus x Formes graphiques]
- L'analyse factorielle des correspondances des tableaux lexicaux
- La recherche des **spécificités** c'est à dire des **formes graphiques** ou des **segments répétés caractéristiques** : "anormalement" fréquente dans les réponses d'un groupe d'individus
- La recherche des **réponses modales** :
  - Critère du Chi-2 : le choix est fondé sur l'écart (distance du Chi-2) entre le profil d'une réponse et le profil moyen de la classe à laquelle cette réponse a été affectée
  - Critère de la fréquence : choix fondé sur le rang moyen des formes caractéristiques que contient la réponse
- Le **regroupement des individus** en groupes les plus homogènes possibles vis à vis des caractéristiques en jeu au travers des **réponses** sans en privilégier a priori.
- La recherche des **concordances** qui permet de replacer une forme graphique dans son contexte. Une forme graphique est donnée avec les formes qui la précèdent et celles qui la suivent

## **3.2. Analyse de contenu d'entretiens**

### **3.2.1. Technique manuelle**

Voir analyse de contenu non assistée par informatique

### **3.2.2. *Technique assistée par ordinateur***

Le même schéma que dans 3.2.1 peut être repris. La différence est que le corpus va être constitué d'entretien, c'est à dire des textes longs dont la redondance est moins forte que dans les réponses aux questions ouvertes. L'équivalent de la réponse est à choisir entre différentes possibilités :

- ligne du texte
- phrase
- paragraphe
- autres découpages

### **3.3. conclusion**

L'ensemble de ces techniques fondées sur une approche différentielle apporte un complément puissant à l'analyse de contenu pratiquée dans une perspective plus traditionnelle.

Des sources pour aborder théoriquement l'analyse de contenu :

n°	Titre ou citation	ouvrage	auteur
6	<p>« Sur le plan méthodologique la querelle approche quantitative et approche qualitative occupe certains esprits. Dans l'analyse quantitative ce qui sert d'information est la <u>fréquence</u> d'apparition de certaines caractéristiques de contenu. Dans l'analyse qualitative c'est la <u>présence</u> ou l'<u>absence</u> d'une caractéristique de contenu donnée ou d'un ensemble de caractéristiques, dans un fragment de message , qui est prise en considération. »</p>	<p>p. 24</p> <p>L'Analyse de contenu PUF 6ème éd. 1991 (291 p)</p>	<p>Bardin, L.,</p>
7	<p>L'analyse de contenu est une méthode objective, systématique et, à l'occasion, quantitative d'étude des textes, en vue d'en classer et d'en interpréter les éléments constitutifs qui ne sont pas totalement accessibles à la lecture naïve.</p> <p>Texte = toute production verbale orale ou écrite (ce qui renvoie aux problèmes soulevés par le langage et les situations d'énonciation)</p> <p>Extension de l'analyse de contenu aux textes musicaux, à l'image fixe ou mobile.</p> <p>Interpréter = inférer, accomplir une opération logique par laquelle on tire d'une ou plusieurs propositions (en l'occurrence les données établies au terme de l'application des grilles d'analyse) une ou plusieurs des conséquences qui en résultent nécessairement. I s'agit de justifier la validité de ce qu'on avance à propos de l'objet étudié en exposant les raisons de la preuve.</p> <p>Le recours à la quantification ...ne doit pas faire oublier le rôle de l'esprit de finesse qui reste indispensable à la pénétration d'un univers mental particulier.</p>	<p>L'analyse de contenu Que sais-je ? n°3271 PUF 1997</p> <p>p 4</p> <p>p 32</p>	<p>Robert, A., Bouillaguet , A.,</p>

8	<p><i>L'introduction vise à argumenter à propos de l'importance à donner aux questions :</i>  <i>comment peut-on établir des conclusions fiables à partir de données qualitatives ?</i>  <i>quelles méthodes d'analyse peut-on utiliser qui soient à la fois pratiques, communicables et objectives,... scientifiques dans le meilleur sens du terme ?</i></p>	<p>pp 21-26</p> <p>Analyse des données qualitatives</p> <p>Recueil de nouvelles méthodes</p> <p>De Boeck Université 1991 (480 p)</p>	<p>A. M. Huberman  M. B. Miles  traduit de l'anglais par C. De Backer et V. Lamongie</p>
9	<p><i>Analyse des données textuelles</i>  <i>Des textes particuliers : les réponses aux questions ouvertes (pp 23-32)</i></p>	<p>Statistique textuelle  Dunod 1994 ( 342 p)</p>	<p>Lebart, L.,  Salem, A.,</p>
18	<p><i>Traitement des questions ouvertes (pp 227-246)</i></p>	<p>Traitements statistiques des enquêtes  Dunod 1994</p>	<p>Grangé, D.,  Lebart, L.,</p>
20	<ul style="list-style-type: none"> <li>- <i>Analyses para-verbales (p 21)</i></li> <li>- <i>Analyses lexicales (p 29)</i></li> <li>- <i>Analyses morpho-syntaxiques (p 89)</i></li> <li>- <i>Analyses sémantiques (p 121)</i></li> <li>- <i>Analyses pragmatiques (p 129)</i></li> </ul>	<p>L'analyse du discours assistée par ordinateur  A. Colin 1998 (217 p)</p>	<p>Marchand ,  P.,</p>
28	<p><i>Analyse structurale de contenus et modèles culturels (pp 111-144)</i></p>	<p>Pratiques et méthodes de recherche en sciences sociales  Albarello, L., &amp; alii.  A. Colin 1995 (179 p)</p>	<p>Hiernaux  J.-P.,</p>
30	<p><i>Les direx analysés : l'analyse propositionnelle du discours</i></p>	<p>PUV, 1985</p>	<p>Ghiglione,  R.  Matalon,  B.  Bacrin,</p>