

Branches complémentaires - Bachelor - Semestre d'automne 2018

MÉTHODES EMPIRIQUES DES SCIENCES SOCIALES

FLORENCE CAUHÉPÉ

Cours no 11 du 28 novembre 2018

L'ANALYSE DE CONTENU

1) LA MÉTHODE LOGICO-SÉMANTIQUE OU ANALYSE PAR CATÉGORIES

Plan

- Introduction
- **I. Présentation de l'analyse de contenu**
- 1. Antécédents historiques
- 2. Définitions et objectifs
- 3. Traitement quantitatif / traitement qualitatif
- 4. Les trois genres d'analyse de contenu
- 5. La préparation du corpus
- 6. Déroulement d'une analyse de contenu
- **II. La méthode logique et sémantique (=analyse thématique ou par catégories)**
- 1. L'analyse par catégories
- 1.1. Le choix des unités, exemple
- 1.2. La catégorisation
- 1.3. Propriété des catégories
- 2. Techniques de contraction des textes
- 2.1. La sélection selon la fréquence
- 2.2. L'art de résumer en cinq points
- 3. Les logiciels
- 4. Exemples; Alceste, Wordle, lexicométrie
- 5. Prérequis
- Bibliographie et Conclusion
- **III. Partie pratique**

Introduction

Qu'est-ce que l'analyse de contenu aujourd'hui?

- Un ensemble de méthodes, «de plus en plus raffinés et en constante amélioration, s'appliquant à des **discours** extrêmement diversifiés», Laurence Bardin, 2011, *L'analyse de contenu*, coll. Quadrige, 1^{ère} éd. 3^{ème} tirage, PUF
- Il y a **des méthodes logiques et sémantiques (logico-sémantiques)** nommées aussi **analyse thématique ou par catégories**, qui portent sur la signification et la mise en lien des thèmes (catégories-codes), elles sont souvent quantitatives
- Le facteur commun des ces méthodes multiples, «-du calcul des fréquences d'un mot fournissant des données chiffrées à l'extraction de structures se traduisant en modèles-, est une herméneutique contrôlée, fondée sur la déduction: **l'inférence**» Laurence Bardin, *idem*

I. Présentation de l'analyse de contenu

1. Antécédents historiques

- L'herméneutique, art d'interpréter les textes sacrés ou mystérieux
- A partir de 1920, aux USA, le matériel analysé est essentiellement journalistique et documentaire, l'analyse des «communications» débute avec:
 - Harold Lasswell (1902-1978) Université de Chicago: «**Propaganda Techniques in the World-War**» en 1927
 - Bernard Berelson (1912-1979) et Paul Lazarsfeld (1901-1976), Universités de Chicago et Columbia: «**The Analysis of Communication Content**» en 1948
- 1940-1950, systématisation des règles et intérêt pour la symbolique politique

2. Définitions et objectifs

Plusieurs définitions, selon que l'on insiste sur tel ou tel aspect:

- «C'est une méthode de classification ou de codification dans diverses catégories des éléments du document analysé pour en faire ressortir les différentes caractéristiques, en vue d'en mieux comprendre le sens exact et précis» René L'Ecuyer, *L'analyse de contenu: notions et étapes*, 1988, in Jean-Pierre Deslauriers, *Les méthodes de la recherche qualitative*, Sillery, Presses de l'Université du Québec
- Rechercher les informations qui se trouvent dans un matériel et dégager **le sens ou les sens** de ce qui est présenté, puis formuler et classer tout ce que contient le document ou la communication

3. Traitement quantitatif ou qualitatif

- Il est fréquent de différencier les méthodes d'analyse de contenu selon deux approches:
 - les méthodes quantitatives
 - les méthodes qualitatives

4. Les trois genres de l'analyse de contenu

- A la frontière avec la linguistique, le territoire des méthodes d'analyse de contenu comprend:
 1. **Les méthodes logiques et sémantiques** (logico-sémantiques) ou méthodes d'analyse **thématiques ou par catégories**
 2. **Les méthodes logiques, esthétiques et formelles (logico-esthétiques formelles)** = analyse du discours selon l'expression et l'énonciation (cours no 13)
 3. **Les méthodes d'analyses structurales** (cours no 13)

5. La préparation du corpus

- Les premières opérations sont :
 1. retranscrire
 2. décrire
 3. organiser
- Il y aura description du contenu du message en prenant en compte :
 - les signifiants (le lexique et la syntaxe...)
 - les signifiés (le sens)
- Puis, on va les organiser en les classant dans des sortes de casiers ou de rubriques significatives (les catégories)
 - ex. des boîtes ou des cartons de chaussures, où l'on range un contenu-des objets, selon différents critères susceptibles de faire surgir un sens

6. Déroulement de l'analyse de contenu, selon Laurence Bardin, brochure p. 29

- I. Préanalyse
 - a) Lecture flottante
 - b) Règle de l'exhaustivité pour le corpus retenu
 - c) Prendre en compte la formulation des hypothèses et des objectifs de la recherche
 - d) Le repérages des indices (=thèmes) et l'élaboration des indicateurs (comptage)
 - e) La préparation ou présentation du matériel
- II. Exploitation du matériel = administration systématique des décisions prises
- III. Le traitement des résultats obtenus et l'interprétation

II. Les méthodes logiques et sémantiques (logico-sémantiques) ou analyses thématiques ou par catégories

- Elles ne concernent pas la recherche de la structure formelle d'un texte, ni son vocabulaire, ni son style, ni sa construction
- Elles s'en tiennent au contenu explicite, manifeste; elles ne considèrent que le sens immédiatement accessible
- Elles procèdent par :
 - **Inventaire ou index**
 - **catégorisation** (=caractérisation-codification)
 - **dénombrement**

1. L'analyse par catégories

1.1. Le choix des unités

- Il faut **découper le texte en partie de sens**
- On parlera alors d'énoncés, de propositions, de phrases ou groupes de phrases :
 - l'unité de sens ou unité d'enregistrement
 - l'unité de contexte
 - l'unité de comptage = la modalité du comptage

Exemple d'unités de sens

- Q: Que retirez-vous de ce séminaire, du point de vue professionnel?
- R: «Outre une meilleure connaissance de collègues jusque-là inconnus ou méconnus, j'ai nettement amélioré ma capacité personnelle de communication, d'intervention en réunion et de travail en groupe.»
- Combien d'unités de sens ? Lesquelles ?

1.2. La catégorisation

1. Décider des unités de sens

2. Créer des catégories

Une catégorie est une notion générale. Elle peut être descriptive, déduite d'une hypothèse ou induite d'un contenu (explicative)

3. Faire l'inventaire et classer

4. Compter

■ Cf. Brochure p. 26-27

1.3. Propriétés des catégories

- Homogènes
- Exhaustives
- Exclusives
- Objectives
- Adéquates ou pertinentes

2. Techniques de contraction des textes

2.1. La sélection selon la fréquence

- **Elimination des mots vides**
- **Normalisation** = élimination des synonymes et des polysémies (redistribuées vers des synonymes sélectionnés)
- **Recherche des mots-clés et des mots associés** (avec plus ou moins de proximité)
- **Calcul des fréquences de tous ces syntagmes**
- **Ecrémage**, c'est-à-dire élimination des syntagmes à basse fréquence, (par rapport à un chiffre posé par l'analyste et donc variable)

2.2. Le résumé en cinq points

- **Simplification du texte**, par élimination des mots vides et des redondances, digressions, développements illustratifs
- **Repérage des mots-clés**, après réduction des synonymes et des polysémies
- **Index de ces mots-clés** pour appréciation de leur **fréquence**
- **Repérage des mots associés aux mots-clés** pour préciser leur champ sémantique + Evaluation de la fréquence
- **Réécriture des énoncés-clés**, par transformation syntaxique, pour éliminer les effets stylistiques inutiles et les structures de phrases complexes, sans toucher au sens

➤ Exemples

3. Les logiciels

- En 1966, le GENENRAL INQUIRER comprenait 17 index et 3521 mots
- Le Harvard THIRD PSYCHOLOGICAL DICTIONNARY: un index psychosociologique qui peut enregistrer 3564 entrées et les classer en 83 catégories ou concepts-clés
- Une analyse textuelle automatisée:
 - ✓ Atlas-Ti
 - ✓ N'Vivo
 - ✓ NUD*IST
 - ✓ Wordmapper
 - ✓ MaxQDA / Welt-QDA
 - Avec **lemmatiseur**
 - ✓ Sphinx Lexica / Sphinx Développement /Lexico 3/5
 - ✓ Alceste en France / Cassandre en Belgique
 - ✓ Trope / Tree Tagger / Cordial / Iramuteq / Spad-T
 - ✓ Le Trameur: <http://tal.univ-paris3.fr/trameur/universiteSorbonnenouvelleparis3>

Transcription et traitement d'un entretien avec le logiciel Alceste

Un enseignant, c'est d'abord un fonctionnaire c'est-à-dire quelqu'un qui a une tâche bien définie à la fois dans le cadre des [programmes_scolaires](#) à respecter, qui au passage changent assez souvent, il faut donc qu'il se tienne au courant, **GÊNE**, avec des [instructions_officielles](#) bien définies. [...] ↵

[L'informatique_à_l'école](#), ça va leur permettre de développer [l'autonomie](#). On se rend compte qu'en début d'année, pour des CE1, ils sont un peu perdus en informatique surtout dans cette école, à Tournefeuille.

RELANCE

Avec le b2i, j'arrive à leur apprendre comment se servir d'un ordinateur, mais notre [salle_informatique](#) n'est pas pratique pour recevoir la [classe_entière](#) [...] ↵

Dans ma vie privée, je m'en sers tous les jours, **SOURIRE**, essentiellement pour un aspect pratique. Je m'en sers pour rechercher une information sur internet, pour me tenir au courant de l'actualité [...] ↵

SAUT DE LIGNE OBLIGATOIRE (retranscription d'un nouvel entretien)

Exemple d'analyse lexicologique quantitative

- Analyse et distribution du lexique (du vocabulaire):

100 - (V.% + Adj. qual.% + Subs.% + Adv.%)

ex.: le quotient d'action = rapport entre le nombre de Verbes et de Substantifs

Prérequis pour les méthodes d'analyse de données textuelles (ADT)

- L'utilisation de l'outillage informatique nécessite certains prérequis:
 - S'assurer que les données collectées et constituées en corpus répondent aux questions de recherche et aux objectifs d'analyses
 - Identifier clairement les unités à traiter (mot, séquence de mots, phrases...)

Bibliographie

- Harold Lasswell, 1927, *Propaganda Techniques in the World-War*, Peter Smith
- Bernard Berelson et Paul Lazarsfeld, 1948, *The Analysis of Communication Content*, Universités de Chicago et New-York
- Roger Mucchielli, 1988, *L'analyse de contenu des documents et des communications*, Les éditions sociales françaises ESF, Paris
- Laurence Bardin, 2011, *L'analyse de contenu*, collection Quadrige, PUF, 3^{ème} tirage
- Dominique Maingueneau, 1976, *Initiation aux méthodes de l'analyse du discours*, Hachette
- Dominique Maingueneau, 2014, *Discours et analyse du discours*, Paris, Armand Colin
- Emilie Née (sous la direction de), 2017, *Méthodes et outils informatiques pour l'analyse des discours*, PUR, Coll. Didact Méthodes
- Adrien Gabouloud, Anne-Sophie Lechevallier, 2018, *Le Poids des mots du Président*, Ed. de l'Observatoire

Conclusion

- L'analyse de contenu a un champ d'application très vaste. Elle peut porter sur des communications de formes très diverses
- Elle répond à des objectifs d'analyse très variés
- L'analyse de contenu est une approche à la fois quantitative et qualitative des textes, qui décrit et décompte, pour aider, de façon objective et rigoureuse, à la compréhension du sens du message

Exercice no 1 p. 54 de la brochure: analyse des réponses à une question ouverte d'un questionnaire

- Exercice d'analyse de contenu, de type classificatoire ou catégoriel, à partir d'un matériel verbal
- Il s'agit de dépouiller les réponses suivantes concernant **les relations psychologiques que l'individu entretient avec la voiture**
- **Analysez les réponses à la question A**



Proposition d'analyse: A quoi peut-on comparer une voiture?

- Utilisation de l'analyse logico-sémantique, classer les réponses selon leur sens, mais selon quels critères?

-Il est possible de ventiler les réponses selon le critère de «l'objet» de comparaison cité: femme, animal, autre moyen de transport...

-Il est possible de ventiler les réponses selon le type de relation psychologique

-Il est possible de distinguer les attitudes positives ou négatives

- Première piste d'analyse possible:

Les relations que l'individu entretient avec sa voiture ne sont pas que fonctionnelles, mais colorées d'affectivité et chargées de symboles

Exercice optionnel no 2: A faire chez vous

- Résumez les paragraphes 1 à 4 du texte de Durkheim p. 48-49 de la brochure, en vous aidant des techniques de contraction et de résumé (slides 15 et 16)
- Ensuite «normalisez» les énoncés (= créer une phrase générique = résumer chaque paragraphe)

Consignes pour l'Atelier du 5.12.

- Prendre vos ordinateurs portables
- Constitution de groupes de 2 ou 3 pour le cadre théorique, puis travail par groupes de 5 ou 6
- Se créer un compte Gmail
- Retour auditoire habituel **G 140 : 9h15-12h**