HAL
open science

L'apport de la méthode Alceste dans l'étude des représentations sociales

Nikos Kalampalikis

To cite this version:

Nikos Kalampalikis. L'apport de la méthode Alceste dans l'étude des représentations sociales. Jean-Claude Abric. Méthodes d'étude des représentations sociales., Erès, pp.147-163, 2005, 9782749201238. 10.3917/eres.abric.2003.01.0147 . halshs-00532849

HAL Id: halshs-00532849

https://shs.hal.science/halshs-00532849

Submitted on 13 Apr 2020

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

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

CAIRN.INFO
Chercher, repérer, avancer.

L'APPORT DE LA MÉTHODE ALCESTE DANS L'ANALYSE DES REPRÉSENTATIONS SOCIALES

Nikos Kalampalikis

in Jean-Claude Abric , *Mhodes d'ude des reprentations sociales*

ERES | *Hors collection*

2003
pages 147 à 163

ISBN 9782749201238

Article disponible en ligne à l'adresse:

---

http://www.cairn.info/methodes-d-etude-des-representations-sociales--9782749201238-page-147.htm

---

Pour citer cet article :

Kalampalikis Nikos,« L'apport de la méthode Alceste dans l'analyse des représentations sociales », *in* Jean-Claude Abric , Mhodes d'ude des reprentations sociales
ERES « Hors collection », 2003 p. 147-163. DOI : 10.3917/eres.abric.2003.01.0147

Distribution électronique Cairn.info pour ERES.

© ERES. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

L'apport de la méthode Alceste dans l'analyse des représentations sociales

Nikos Kalampalikis

Si le choix d'une méthode spécifique pour étudier le fait social correspond à une façon spécifique de penser le fait social, il nous semble indispensable, avant d'entrer dans le vif du sujet, de spécifier notre propre *façon de penser*. Pour cela, nous nous servirons en la reprenant et en la développant, d'une distinction proposée par Moscovici (1991) quant aux deux catégories de *faits* – on pourrait également parler de *phénomènes* – dont s'occupent les sciences de l'homme : les faits « anonymes », et les faits « nominés ».

Les faits « nominés » sont étayés par une théorie conçue pour et par des spécialistes (p. ex. la science, la religion) ; ils sont objectivés par les institutions qui les incarnent et imprégnés par un langage qui leur est propre. Ils sont parlés et classés dans des catégories bien distinctes et visibles, ils se déroulent en pleine lumière pour ainsi dire.

Les faits « anonymes » consistent, en revanche, en actions réciproques, une sorte de miroir de notre société, imprégnés de rencontres, de conversations, d'échanges, de souvenirs, de jugements, de rumeurs qui mettent en scène et parfois déterminent notre existence concrète. Ces faits anonymes se déroulent souvent dans le « silence », un silence partagé qui, loin d'être vide, passe inaperçu tout en influençant notre praxis quotidienne.

Nikos Kalampalikis, Laboratoire de psychologie sociale et de psychologie de l'action, EHESS, *Paris.*

Nous tenons à remercier Max Reinert, auteur d'Alceste, pour sa relecture attentive et ses remarques constructives.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

Ce silence remplit d'images, de gestes, d'émotions et, bien sûr, de mots, cette cognition ordinaire, contenant toute une *phénoménologie de notre sociabilité vécue*, fournit un terrain privilégié où l'on peut saisir la façon – *les façons* – dont les membres d'une société interagissent et communiquent entre eux, interprètent leur vie, nomment et définissent leur réalité collective, anticipent et jugent autrui, se souviennent de leur passé partagé.

Voici donc, en bref, une façon de penser *le fait psychosocial* qui s'inspire largement de la tradition théorique issue du courant des représentations sociales (cf. en dernier lieu, Jodelet, 2002). Qui dit fait, dit souvent objets d'étude, et ceux des représentations sociales portent en eux des caractéristiques et des qualités spécifiques liées à leur relation inextricable au savoir du sens commun et à sa mise en œuvre dans l'expérience quotidienne. Rappelons que les représentations sociales sont :

« – des *programmes de perception*, des constructs à statut de *théorie naïve*, servant de *guide d'action* et de *grille de lecture* de la réalité, des *systèmes de significations* permettant d'interpréter le cours des événements et des relations sociales ;

– elles expriment le *rapport* que les individus et les groupes entretiennent avec le monde et les autres ;

– elles sont forgées dans *l'interaction* et au contact des *discours* circulant dans l'espace public ;

– elles sont inscrites *dans* le *langage* et les *pratiques* et fonctionnement *comme* un *langage*, en raison de leur fonction symbolique et des *cadres* qu'elles fournissent pour coder et catégoriser ce qui peuple l'univers de vie. » (Jodelet, 2002, c'est nous qui soulignons)

Cette riche axiomatique contient en soi un programme de travail pour les études à venir dans le domaine. Elle nous intéresse tout particulièrement ici dans sa double assertion des représentations, *dans* le langage – *comme* un langage, ainsi que dans l'attention portée à leur fonction symbolique, expression et interprétation d'une réalité sociale.

Le matériel discursif est souvent considéré comme une *via regia* pour l'étude de cette dynamique psychosociale des représentations, qu'il soit spontané (associations verbales), exprimé dans des entretiens (individuels ou collectifs), induit par des questions ou cristallisé dans des corpus littéraires, documentaires, filmiques, iconographiques ou même épistolaires. Ceci étant, une fois le matériel de recherche recueilli, et notamment quand celui-ci est particulièrement volumineux, nous nous rendons compte que son analyse risque de devenir une tâche particulièrement laborieuse et que la seule analyse de contenu classique ne nous suffira pas, malgré son induscitable efficacité. D'autant plus que depuis quelques

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

années, avec la multiplication des sources informatiques, la massification des données numériques et l'automatisation des procédures de traitement de données sur ordinateur nous sommes tentés de soumettre notre matériel à plus d'une analyse. D'où l'appel aux différents logiciels d'analyse textuelle existant (Alceste, Prospéro, Tropes etc.) afin d'avoir une vision plus globale et plus homogène de notre matériel (linguistique, grammaticale, sémantique, thématique), mais tout aussi bien une subtilité et une finesse du détail qui ne sont pas forcément visibles « à l'œil nu ».

Cet appel à l'informatique, sœur jumelle dans ce cas spécifique de la linguistique, ne peut pas ne pas avoir d'incidence sur la façon dont nous traitons et interprétons nos matériaux psychosociaux, sur nos *façons de penser* notre matériel discursif. Force est de constater que jusqu'à présent, deux « écoles », officieuses mais visibles, se dessinent plus au moins clairement dans le paysage des utilisateurs de ce type d'outils pour leurs recherches : soit une confiance « aveugle » en l'outil informatique, devenant désormais un emblème indispensable de scientificité, accompagnée d'une réflexion timide sur le langage ; soit un travail analytique proche de celui des linguistes qui, de par sa complexité minutieuse, évacue la pertinence sociale et l'aspect communicationnel de l'objet d'étude et qui, de plus, fait fuir les novices.

Indubitablement, il ne s'agit pas pour les psychologues sociaux de devenir linguistes (ni, bien évidemment, pour les linguistes de devenir psychologues sociaux). Il s'agit de prendre en considération les apports de chaque discipline et de les incorporer aux besoins de recherche mutuels, afin de dépasser ce clivage décrit avec pertinence par Faucheux et Moscovici, déjà en 1966 (cf. Moscovici, 1967), mais qui, selon nous, perdure encore aujourd'hui. D'une part, un *langage sans communication*, celui des linguistes pour qui la communication est souvent une abstraction, et de l'autre, une *communication sans langage*, celle des psychologues sociaux, pour qui les phénomènes étudiés (changement d'attitude, influence, innovation etc.) semblent souvent indépendants du principal système de signes qui les rend possibles. Nous pouvons éventuellement ajouter à cette dernière distinction une plus récente, issue du développement technologique et de l'informatique, à savoir *le béhaviorisme numérique* (cf. Buschini et Kalampalikis, 2002), c'est-à-dire la transformation de l'ordinateur en une nouvelle boîte noire, pour ce qui concerne les logiciels statistiques et plus particulièrement les logiciels d'analyse textuelle qui nous intéressent ici.

Le but de ce chapitre est de présenter une méthode et un outil de statistique textuelle dont l'apport pour le domaine d'étude des représentations sociales se mesure déjà au nombre de recherches l'ayant utilisé avec

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

pertinence et succès. C'est une méthode qui, selon nous, peut stimuler le dialogue au sein des disciplines du social quant à la nature de nos corpus textuels, nos hypothèses, traitements et interprétations, bref quant à la *mesure du pouvoir symbolique du langage des représentations*. Nous n'avons pas là, la prétention de présenter le détail algorithmique du logiciel, ni même de suivre à la lettre les théorisations qui ont inspiré sa conception. Cette méthode et ce logiciel nous intéressent dans la mesure où ils contiennent un potentiel important d'acclimatation à « notre[1] » logique de recherche contribuant à une approche herméneutique du contenu des discours et tentant plus généralement de rendre intelligible la communication sociale.

ALCESTE : UNE MÉTHODE D'ANALYSE TEXTUELLE

Comme son nom l'indique – Analyse des Lexèmes Cooccurrents dans les Énoncés Simplifiés d'un Texte – Alceste est un logiciel d'analyse textuelle qui, comme tant d'autres[2], propose aux chercheurs en sciences sociales une aide informatique face à des corpus textuels numérisés. La particularité de ce logiciel réside dans le fait que derrière sa matérialité informatique se cache une méthode de statistique textuelle, d'inspiration harissienne et benzécriste[3], dont l'intérêt n'est pas moindre pour la psychologie sociale et, notamment, pour les études qui utilisent l'approche des représentations sociales.

Loin de nous la volonté de baptiser Alceste comme la seule ou l'unique méthode d'analyse de textes. C'est une méthode et un outil parmi tant d'autres qui peuvent nous aider à mieux comprendre notre matériel discursif et à prendre en considération l'axe du langage dans nos études ; de plus, des recherches pertinentes et fines lui ont permit de faire ses preuves dans le domaine (cf. p. ex. Lahlou, 1998 ; Masson et Moscovici, 1997 ; Kronberger et Wagner, 2000). Néanmoins, elle n'est en aucun cas la seule, et ne peut aucunement substituer ni l'intuition du chercheur, ni la finesse herméneutique issue de l'analyse de contenu classique, ni le

1. Logique par définition partagée, depuis longtemps, par d'autres disciplines du social travaillant avec la notion de la représentation et/ou l'analyse du discours, notamment la sociologie, mais aussi l'anthropologie sociale et l'histoire.

2. Le lecteur peut trouver une description fine du paysage des logiciels d'analyse textuelle dans un article de Jenny (1997).

3. Benzécri lui-même revendiquait une double affiliation de ces travaux : d'une part la théorie distributionnelle de Harris, de l'autre, la « loi de Zipf » qu'il mérite de rappeler ici : *le produit du rang (selon l'ordre de fréquence croissante) et du nombre d'occurrences de chaque élément d'un texte est à peu près constant* (Jenny, 1999).

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

travail d'interprétation nécessaire à toute réflexion dans les sciences du social. Nous présenterons donc quelques principes fondamentaux de la philosophie de cette méthode, ainsi que son utilité à l'égard de notre discipline. Nous illustrerons nos propos à l'aide d'une recherche empirique récente portant sur les processus symboliques des représentations sociales à travers l'exemple des pratiques alimentaires.

L'approche des « mondes lexicaux »

L'objectif de cette méthode n'est pas le calcul du sens, mais l'organisation topique du discours à travers la mise en évidence des « mondes lexicaux ». Dans Alceste, le vocabulaire d'un énoncé constitut une trace, une référence, une activité, bref une intention de sens du sujet-énonciateur. Pour cette méthode, qui ne cache en rien son fondement lexicométrique [4], les traces les plus prégnantes de ces activités dans le lexique sont les « mondes lexicaux ».

La traduction statistique de ce présupposé théorique se retrouve dans la redondance des traces lexicales. Nous parlons de traduction étant donné que la statistique nécessite forcément de la « matière » pour circonscrire son analyse. Cette matière dans la méthode Alceste n'est autre que le lexique [5]. Nous pouvons d'ores et déjà dire que la trace lexicale d'un énoncé n'est pas forcément égale à la somme de tous les mots constituant cet énoncé, mais à ceux qui servent, de par leur potentiel sémantique et grammatical, d'attracteurs sémantiques. Une redondance de traces lexicales circonscrit un lieu usuel du discours, un lieu donc où un *certain* vocabulaire est utilisé *fréquemment*, un lieu référentiel où le sujet-énonciateur semble avoir déposé du *sens*. La méthode Alceste permet donc de mettre en lumière les traces lexicales les plus prégnantes de ces espaces de référence, les « mondes lexicaux », là où une *empreinte du sens* attend l'œil vigilant du chercheur.

Restons un instant sur ce présupposé de base dans la mesure où il constitue, selon nous, une des clefs de compréhension des résultats obtenus par le biais de cette méthode. Jusqu'à présent, nous n'avons guère évoqué, ni le « contenu », ni la « signification », ni « l'interprétation ». Nous sommes au stade – modeste, prudent, mais important – de la loca-

4. La lexicométrie est une méthodologie d'étude du discours dont l'acuité dépend, au moins, de trois principes : l'invariabilité de l'unité de comptage, des quantités importantes et équilibrées d'occurrences, la comparabilité et l'interprétabilité des résultats (Charaudeau et Maingueneau, 2002, p. 342).

5. Le terme de trace lexicale désigne une forme de lexique spécifique d'un énoncé, les « mots-pleins » – nous y reviendrons.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

lisation des mondes lexicaux, de la topographie du discours. À la manière des archéologues qui utilisent des vues aériennes de l'espace pour cartographier la région des fouilles leur permettant de découvrir des fragments significatifs d'une vie collective passée, nous sommes en train de circonscrire l'espace de notre corpus lexical et de regrouper des objets et des lieux usuels, avant de tenter d'en donner une description précise et une interprétation fine.

Selon Max Reinert (1999), auteur de la méthode, Alceste permet de cartographier les principaux lieux communs d'un discours, « communs » dans le sens où ils sont plus souvent « habités » par les énonciateurs. « Ces mondes lexicaux sont donc des traces purement sémiotiques inscrites dans la matérialité même du texte. En eux-mêmes, ils sont indépendants de toute interprétation. Mais, ils ne prennent sens, pour un lecteur, qu'à travers une activité interprétative particulière en fonction de son propre "vouloir-lire" » (*ibid.*, 1997).

Les « mondes lexicaux » ont un sens et ne se prêtent à l'interprétation *que* dans un rapport d'opposition, un système de relations antithétique. Ainsi, leur autonomie est proportionnelle à leurs liens et, de surcroît, aux rapports de force qui les rendent différentes. Autrement dit, les classes lexicales bénéficient d'un degré d'autarcie égal à zéro. Nous pouvons ajouter à cet axiome, une caractéristique supplémentaire à savoir que les mondes lexicaux fuient la solitude ; ils n'ont de sens et d'existence que dans la pluralité. Aucune analyse à l'aide d'Alceste ne donnera une seule classe lexicale pour tout le corpus analysé (d'ailleurs, il n'y aurait pas besoin d'en faire une) ; il y en aura au moins deux, sinon plus (l'exemple que nous allons présenter par la suite est tiré d'une analyse donnant un réseau de neuf classes lexicales[6]). La méthode Alceste s'atttache ainsi à saisir la dynamique de la production verbale à travers sa nature « conflictuelle », principe à ne pas omettre lors de l'interprétation des classes lexicales obtenues.

Le fond associatif

Afin de mieux comprendre la précédente axiomatisation nous pouvons nous référer à une technique « populaire » en psychologie sociale, les associations verbales. Lorsque l'on demande à des sujets de donner les mots qui leur viennent à l'esprit à partir de tel ou tel mot-stimulus, ils produisent une série de mots, souvent des noms, des verbes, parfois des

6. Le nombre des classes lexicales n'est pas une preuve de richesse du matériel mais de diversité.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

adverbes, moins fréquemment des phrases entières. Les mots associés ont un statut sémantique particulier pour le chercheur, disons pour faire bref, *différent* de celui des énoncés d'un entretien ou d'une réponse à une question ouverte. Nous avons tendance à les voir, à les considérer et à les analyser comme étant des productions verbales isolées quoique reliées entre elles aussi bien dans l'espace que dans le temps. Mises à part les diverses méthodes d'analyse que nous pouvons leur appliquer, ce qui demeure vrai c'est que nous les considérons *grosso modo* comme étant plus « pleins » d'un point de vue sémantique que ceux d'une phrase, même lorsque il s'agit des mêmes mots. Les conditions spontanées de leur production sont certainement une des raisons de cet effet de « plénitude », il faut également ajouter l'absence de leur ossature logico-formelle habituelle, la phrase.

Un des effets de la méthode Alceste c'est justement cette traduction/transformation du discours formel en un grand ensemble de type associatif où, justement, l'attention est plus orientée, donc concentrée, sur les mots qui font sens, « déshabillés » de leur mise en forme grammaticale. Non pas que les différents connecteurs logico-sémantiques (prénoms, prépositions, conjonctions etc.), ceux que l'on appelle souvent dans le langage de certains logiciels : les « mots-outils » (ou morphèmes), n'ont pas leur place dans l'analyse. Il s'agit de focaliser l'attention plus sur des mots qui fonctionnent comme des prédicats dans le discours, les « mots-pleins » (ou léxèmes [7]). Nous sommes au cœur d'une approche des contenus à l'aide de « contenants textuels » (Reinert, 2001).

François Jodelet dans son article sur l'association verbale estime que « l'ensemble des liaisons associatives d'un locuteur peut être conçu comme un système doué de propriétés structurales […], support d'un processus de signification » (1972, p. 99). Ces liaisons qui restent dans un état latent au sein d'un discours normal, constituent, selon le même auteur, un *schème opératoire du comportement verbal*, un *fond associatif ou topique* selon Reinert [8], opérant à travers la cooccurrence des mots pleins. La compréhension de la spécificité des mots-pleins est donc essentielle, car elle nous permet de comprendre le mécanisme de constitution d'une grande partie des classes lexicales mais aussi elle nous fournit une clef d'interprétation à travers l'idée sous-jacente d'un *fond associatif* qui les sous-tend.

7. Par exemple, dans la phrase « ceci est un chapitre » le seul mot plein sera le mot « chapitre ».
8. L'exemple des associations de mots n'est guère choisit au hasard, Max Reinert ayant lui-même consacré sa thèse de doctorat sur le sujet.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

ALCESTE : LE LOGICIEL

La méthode Alceste repose sur un découpage du corpus en fragments de taille relativement analogues, dits « unités de contexte[9] ». Ces mêmes fragments sont ensuite classés statistiquement selon une procédure descendante hiérarchique. La fragmentation du corpus est effectuée de manière algorithmique. À partir d'une première distribution des fragments en deux classes les plus différenciées possible – au niveau de leur vocabulaire spécifique – la plus grande d'entre elles se voit re-distribuée à nouveau jusqu'à ce qu'elle se divise en deux. Cette double opération de distribution/classification continue jusqu'à obtenir un nombre stable de classes. L'objectif de cette classification descendante hiérarchique étant la répartition des énoncés en classes marquées par le contraste de leur vocabulaire (Reinert, 1993 ; 2001).

L'ensemble des calculs est planifié en trois grandes étapes :
– le calcul du tableau de données avec en lignes, les énoncés, et en colonnes, les lexèmes retenus ;
– le classement des énoncés en fonction de la ressemblance ou dissemblance de leur vocabulaire ;
– la description des classes obtenues grâce à différentes procédures, telle que la sélection du vocabulaire spécifique de chaque classe ; la sélection des énoncés de chaque classe les plus représentatifs de ce vocabulaire ; le calcul des segments répétés par classe ; etc. (Reinert, 2001).

D'un point de vue plus pratique, Alceste est un logiciel informatique pouvant traiter tout document (sauvegardé sous format « texte seulement avec saut de ligne ») dépassant une taille minimum de dix mille mots (environ mille lignes de soixante-dix caractères).

Quel type de matériel ?

Le matériel à traiter *(corpus)* doit être enregistré dans un fichier unique et peut être composé d'entretiens, associations de mots, articles de presse, réponses à des questions ouvertes, textes littéraires, en fait tout document écrit à l'aide de l'alphabet latin, pouvant comprendre les dix chiffres et les signes usuels de ponctuation, pourvu que ce matériel présente une cer-

9. Dans le vocabulaire d'Alceste nous distinguons deux types d'unités de contexte ; les « unités de contexte initiales » (uci) qui correspondent aux divisions naturelles du corpus (par exemple les chapitres pour un livre, les réponses à une même question ouverte etc.) et les « unités de contexte élémentaires » (uce) qui ressemblent à la « phrase » grammaticale habituelle. La ressemblance est plutôt d'ordre métaphorique étant donné qu'une uce peut contenir des segments de texte de quelques lignes.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

taine homogénéité (par exemple, il ne sera pas recommandé de mélanger dans une même analyse, les associations de mots avec les entretiens et/ou avec les réponses à des questions ouvertes).

Quelle préparation ?

Une préparation du corpus est nécessaire si nous sommes intéressés par le rôle que certaines *variables externes* au contenu, telle que l'âge, le sexe, l'orientation politique, le statut marital, la date, le titre, etc. peuvent jouer. Il s'agit d'une procédure de codification permettant de calculer par la suite ces variables dans l'analyse, de manière à ce qu'on puisse mesurer leur poids et leur distribution dans l'ensemble du corpus. De même, notamment par rapport à des corpus d'entretiens, nous pouvons introduire des *variables internes thématiques* (c'est le cas dans l'exemple présenté). Ces variables sont introduites par une *ligne étoilée* (d'où le besoin préalable de supprimer le symbole étoile (*) si il existe dans le corpus).

Dictionnaires et réduction des formes

Le logiciel possède différents dictionnaires de reconnaissance des locutions potentiellement modifiables par l'utilisateur. De plus, il établit un dictionnaire du vocabulaire du corpus, puis un dictionnaire des « formes réduites » (p. ex. il rassemble les formes « congélation », « congelés », « congeler », sous la même racine « congel+ » dont la fréquence sera prise en compte [10]).

Expressions clés

Si le corpus contient des formes verbales (par exemple des expressions telles que « noyau central » ou « représentation sociale ») que nous souhaitons traiter comme des expressions il faut les unir avec un tiret bas (« noyau_central », « représentation_sociale »).

Analyse et rapport d'analyse

L'analyse peut être lancée à partir des paramètres par défaut du logiciel, modifiables selon le degré de familiarisation de chacun. Un rapport est automatiquement préparé à la fin de l'analyse contenant l'ensemble

10. Ces formes réduites sont issues d'une opération de « lemmatisation » consistant à remplacer une forme textuelle par son lemme, c'est-à-dire, la forme standardisée des dictionnaires de langue.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

des opérations statistiques mises en œuvre (classification descendante hiérarchique, profil des classes, vocabulaire spécifique, analyse factorielle des correspondances, segments répétés etc.).

Un exemple : les mutations dans les pratiques alimentaires

Plusieurs études en psychologie sociale utilisent la méthode Alceste pour l'analyse de leurs données discursives. Nous avons pris pour exemple une étude menée par Estelle Masson et Serge Moscovici [11] (1997) au laboratoire de psychologie sociale de l'École des hautes études en sciences sociales [12], dans le cadre d'un contrat de recherche ministériel, autour des mutations dans la pratique alimentaire. Notre choix a été dicté par le fait que ces auteurs ont su faire un usage intelligent d'Alceste, combiné aussi bien à une analyse thématique préalable qu'à un travail réflexif et herméneutique a posteriori qui permet d'aller au-delà des simples impressions sémantiques et de saisir des véritables processus symboliques qui sous-tendent les pratiques des mangeurs contemporains que nous sommes.

Cette recherche a eu pour objectif d'étudier la place des conserves et plus spécifiquement des plats préparés industriels dans les comportements alimentaires des mangeurs d'aujourd'hui. D'un point de vue méthodologique, elle repose sur l'analyse de cinquante entretiens en profondeur recueillis auprès d'une population mixte [13]. Elle nous paraît exemplaire dans la manière d'utiliser intelligemment la méthode Alceste dans la mesure où ses auteurs ne se limitent pas à ses résultats bruts, mais forgent leur interprétation tant grâce à une analyse de contenu thématique qu'aux apports de la littérature issue des sciences sociales (notamment l'anthropologie, la sociologie et l'histoire). Nous allons justement essayer, succinctement, de suivre cette logique d'analyse en présentant une partie de leurs résultats.

---

11. Nous tenons à remercier les auteurs de nous avoir permis de citer largement leur recherche.
12. Il mérite de rappeler que les premiers travaux assistés par Alceste dans le domaine des représentations sociales ont été initiés et réalisés par plusieurs membres de ce laboratoire.
13. Un premier répertoire de l'argumentation et des grandes familles thématiques relatives à l'objet de l'étude a été effectué à travers une première série de dix entretiens exploratoires qui a permis l'élaboration de la grille d'entretien. La dimension « convivialité » apparaissant comme discriminante des pratiques et des représentations, les auteurs se sont appliqués à constituer un échantillon composé de deux groupes spécifiques de vingt personnes. Un groupe de sujets mangeant fréquemment seuls et un autre mangeant majoritairement en compagnie de commensaux. À l'intérieur de ces deux groupes ils ont contrôlé diverses caractéristiques sociologiques (âge, profession, lieu et type d'habitation, sexe, situation matrimoniale, avec ou sans enfant etc.).

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

Topographie du discours sur les aliments : les classes lexicales

Tout d'abord, les auteurs ont regroupé des séquences de leurs entretiens où il était question d'aliments dans un corpus du même nom qui a été soumis à une analyse Alceste. Leur objectif était triple :
– la mise en évidence des *contenus* du discours des interviewés sur les aliments ;
– l'isolement des *thématiques* autour desquelles s'articule ce discours ;
– l'étude des *liens* entre classes.

L'analyse Alceste a fait apparaître neuf classes lexicales distinctes dans le corpus « aliments » que nous présentons dans le tableau ci-dessous accompagnées de certains traits sémantiques typiques pour chacune d'entre elles et le thème auquel chacune renvoie. Notons qu'il s'agit donc d'ores et déjà d'un premier niveau d'interprétation tenant à donner un sens aux « mondes lexicaux » isolés par Alceste.

Ce premier degré d'interprétation se base sur une série d'indices apportés par Alceste que le chercheur doit combiner, comparer et interpréter à partir d'un critère de pertinence par rapport à son objet d'étude.

Ces indices peuvent être :
– le poids de la classe lexicale à l'égard de l'ensemble du corpus (en termes de pourcentage) ;
– le chi2 d'appartenance d'un trait lexical (mot, lemme) à une classe et/ou à l'ensemble du corpus [14] ;
– les traits lexicaux les plus/les moins typiques [15] de chaque classe ;
– le vocabulaire spécifique de chaque classe ;
– l'incidence des variables externes (signalétiques) ou internes (thématiques) sur la constitution de chaque classe etc.

Nous pouvons par exemple prendre la première classe lexicale du tableau précédent (« congélation ») et suivre le raisonnement des auteurs de cette étude, afin de saisir véritablement *comment* et *pourquoi* ils lui ont attribué ce thème. Cette classe, selon eux, « renvoie très clairement à une technique industrielle de conservation qui est la *congélation* (congel+ chi2 = 162). Les traits lexicaux qui caractérisent cette classe constituent non seulement une description de la technique de congélation, mais éga-

14. La valeur du chi2 (à un degré de liberté) exprime le coefficient d'association d'une forme lexicale à une classe calculé sur un tableau de contingence qui croise la présence/absence de la forme dans une uce et l'appartenance ou non de cette uce à la classe lexicale.
15. La typicité d'un trait peut résulter d'une combinaison de plusieurs indices, comme sa fréquence, son chi2, sa projection sur l'analyse factorielle (AFC), ses variations (lemmatisation), sa répétition sous forme de segment répété etc.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

Tableau 1. – Les neuf classes lexicales issues de l'analyse Alceste.

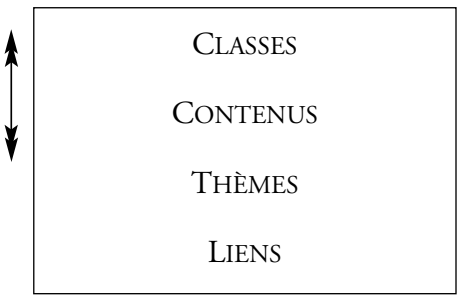
	Thème de chaque classe	Traits lexicaux typiques par classe
1	congélation	congel+, conserv+, decongel+, surgele+, produit+, glace+, boit_de_conserve, froid+, etat+, temperature+, ressembl+, gard+, technique+, industriel+, saveur+, tuer, dependre, pourri+, cuire, mourir, concentr+, frais+, durer, perdre, preserv+, suite+, excell+, epinard+, dat+, dur+, transform+, mieux, beaucoup, chaud+, conflit+, longtemps, princip+, eau, vie, consomm+, mode+, texture, brut+, constat+, procede+, plusieurs, traite+, haricot_vert
2	conserve	conserv+, bocal+, boite+, verre+, gout+, metal, fer, enlev+, eau, huile+, cuire, rinc+, sale<, ferraille+, haricot+, flotte, artisanal+, baign+, gras, en, mode+, ils, jus, sel+, condition+, haricot_vert, prefer+, ajout+, car, espece+, mou+, change+, laver, differ+, cassoulet+, propre+, frais+, proche+, cuiss+, mettre, bouill+, regle+
3	aliments de base civilisés	sauc+, thon+, beurre, boite+, flageolets, poivr+, jus, ail+, ingredient+, tomate+, epice+, mettre, melange+, champignon+, sardine+, nouilles, rajout+, dedans, ravioli+, petit_pois, sel+, fad+, avec, haricot_vert, pas_bon+, degueulasse+, sal+, truc+, cuisin+, vin, petit+, dans, choucroute+, haricot+, creme+, cassoulet+, concentr+, servir, graisse+, systematique+, blanc+, donn+, autant, four, precis+, ouvrir, plat+
4	dépannage	depann+, soir+, heure<, ouvrir, dessert+, seul<, machi+, puree+, invit+, pret+, vite+, finir, avantage+, mage+, sachet+, frit+, prepar+, minute<, truc+, facil+, paquet+, gens, tarte+, te, attendre, plat_tout_pret, accompagne+, filet+, temps, rapid+, passer, tu, faire, hop, global+, rechauff+, maison, pratique+, blanc+, rest+, chou_fleur+, ami+, permettre, cuisin+, envi+, revenir, genre+, table, usine+, stock+, voir, style+, mois, mar+, sol+, limite+, travail+, repas+, as, sortir, base, besoin, quelqu-un, pomme_de_terre, assiette, cantine, pizza+, notion
5	plats préparés industriels	plat+, surgele+, ravioli+, achet+, plat_prepare, paella+, en, sous_vide, findus, prepar+, arriv+, rare+, couscous, poisson+, soupe+, carotte+, cassoulet+, j, chili, aim+, rayon+, trouv<, pane+, ajout+, hiver, varie+, rechauff+, souvenir, aime_pas_trop, carre+, restaura+, semoule, entier+, petit_pois, hache+, invit+, congel+, genial+, exist+, genre+, repas+, balle+, boite+, petit+, fraiche+, plat_tout_pret, precis+, essai+, quantite+, pas_mal, mere, souvent, truc+, jamais, prefer+
6	fruit - légume poisson	saison+, fruit+, mur+, cueillir, tomate+, soleil, pousser, jardin+, odeur+, holland+, peche+, frais+, mer, venir, touch+, pys<, couleur+, franc+, gout+, abime+, aspect+, hiver, aubergine+, legum+, possible+, exemple+, ont, amener, product+, huitre+, pomme+, parents, nom+, origin+, poisson+, normal+, direct<, raison+, table, faux<, marchand+, orange+, belle+, etal+, difficile+, coca, pere, labas, marche+
7	lieux d'approvision-nement	surface+, semaine+, aller, grand+, chez, primeur+, magasin+, casino, bouche+, course+, achet+, commercant+, boulanger+, legum+, regard+, moment+, saint, pain+, connaitre, choisi+, import+, cherch+, tes, jour+, vendre, fruit+, confi+, kilo+, rentr+, endroit+, matin+, franc+, attent+, emmen+, provenance, lait+, parl+, arriv+, fois, mon, prix, argent, aim+, rayon+, lieu+, marchand+, viande+, monde+, loin, agreable, cher+
8	viande	poul+, elev+, ferme+, fermier<, batterie+, oeuf+, mort<, veau+, tendre<, sang, boeuf+, rendre, rouge+, compte+, vivant<, vivre, hormones, bet+, mouton+, campagne, nouvelle+, viande+, chair+, morceau+, agneau+, couche+, courir, blanch+, steak+, evident, vieux, cout+, bouchers, beau+, vie, cent<, rond+, enorme+
9	la problématique de l'assimilation	aliment+, sante, nutritiv+, mauvais+, corps, haut+, sucre+, sens, naturel+, ingenieur+, force+, gamme, professionnel+, apport+, associ+, critere+, monoprix, echelle, risque+, alcool+, consci+, appreci+, chimique+, equilibre+, bon+, porc, pate<, esprit, savoir, plait, riche+, manque+, lourd+, controle+, nourriture, joue<, sain+, bois, uniquement, provenance, element+, complet+, yaourt+, poids, graisse+, image+, nourri+, evite+, special+, type+, engrais, malad+, qualite+, princip+, trop, complique+

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

lement son principe de fonctionnement. C'est un procédé dont la finalité est : « conserver » (conserv+ chi2 = 150) « garder », « préserver », c'est-à-dire d'enrayer temporairement le processus naturel de putréfaction (« pourrir », « mourir »). Cette finalité est atteinte par le recours au « froid », « glace », à savoir par une action sur la « température ». Les effets de ce procédé sur les produits sont appréhendés au travers de deux logiques : la logique de la ressemblance et de la dissemblance par rapport aux produits frais (qui s'exprime par la présence des traits : ressemble, frais, saveur, perdre etc.) et la logique du vivant et du mort (« tuer », « mourir », « préserver »). À un niveau pragmatique le processus est simple et résumé par les trois traits les plus typiques de la classe : « congelé » → « conservé » → « décongelé ». Dans cette classe il n'est pas question d'aliments mais de produits (produit+ chi2 = 82). Ce n'est qu'au seizième rang avec « épinard » (chi2 = 15) qu'apparaît un trait renvoyant à quelque chose pouvant être mangé. Selon toute évidence il ne s'agit pas dans cette classe d'un discours portant sur des denrées comestibles, mais véritablement d'un discours relatif à une technique de conservation appliquée à des produits alimentaires. La présence de « boîtes de conserves » (chi2 = 59) suggère une évaluation comparative du procédé de congélation à celui de l'appertisation. Nous baptiserons cette classe *congélation* du fait que son trait le plus typique est « congèl » qui est la forme réduite de congélation. » (*op. cit.*, p. 78).

Et ainsi de suite pour les huit autres classes dans un raisonnement interprétatif que nous avons essayé de résumer à l'aide du schéma suivant :

Figure 1. – Schéma d'interprétation possible des classes lexicales issues d'Alceste.



Voici donc une première cartographie du discours obtenue grâce à l'analyse Alceste. Elle fournit un ensemble relativement organisé quoique hétérogène (notamment dans le cadre de l'exemple que nous avons choisi étant donné le nombre élevé de classes obtenues). Précisons d'emblée que cette impression d'hétérogénéité est dûe, mise à part la nature du corpus

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

analysé, à la recherche de différences, critère de la constitution même des classes dans Alceste. D'où un besoin de faire un pas en avant et d'essayer d'homogénéiser les thèmes obtenus.

Des mondes lexicaux aux univers thématiques référentiels

À partir d'un premier découpage en neuf classes lexicales issues de l'analyse Alceste, les auteurs de cette recherche ont regroupé dans des catégories thématiques plus globales les classes lexicales dont les thèmes étaient proches. Par exemple, sur le tableau suivant, les thèmes des quatres classes lexicales (aliments de bases civilisés, plats préparés industriels, fruits-légumes-poissons, viandes) ont été assemblés en une thématique référentielle plus globale : les « nourritures ». Le principe opératoire de ce regroupement est issu de l'analyse de contenu classique et tient à prendre comme unité de signification de base, le thème ; il n'est pas sans rapport avec la logique d'Alceste dans la mesure où le découpage initial du corpus cherche à illustrer le noyau conflictuel de la source de la production verbale, en deux mots les différences.

Ce type de rassemblement permet d'avançer à petits pas dans l'interprétation de notre corpus en passant des *mondes lexicaux* aux *univers thématiques référentiels.* Cela serait une illusion positiviste que d'affirmer que les frontières entre les deux opérations sont clairement définies et délimitées. Nous pouvons néanmoins soutenir, avec un peu plus de certitude, qu'à partir de cette étape de l'analyse nous passons d'une *démographie sémiotique*[16] à la formation d'un *réseau de significations*.

Le tableau page suivante résume ces regroupements, ainsi que la représentativité de chaque thématique (pourcentage des uce des classes lexicales) par rapport à l'ensemble du corpus.

Les variations autour du thème de la nourriture, des techniques de congélation, des contextes de consommation, de l'accès aux aliments et *in fine* de leur destin, l'assimilation, donnent à voir les principaux pôles du réseau de significations qui ont été activés lorsque les sujets de cette recherche ont été invités à se prononcer sur les aliments. Ce qui frappe les auteurs de cette recherche à la lumière de ce regroupement c'est la quasi absence du contexte de célébration du manger. Une explication possible, selon eux, pourrait être les conditions de la vie moderne qui imposent un

16. Nous empruntons la notion de « démographie » à Pêcheux (1967) et nous nous permettons de l'utiliser dans ce contexte car le logiciel fonctionne en mettant en œuvre des mécanismes d'analyse indépendants du sens à la manière de la démographie recensant des populations d'individus indépendamment de leur traits de personnalité.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

Tableau 2. – De l'interprétation des classes lexicales à leur thématisation.

Thèmes des classes	*Thématiques référentielles (% uce)*
– aliments de base civilisés – plats préparés indutriels ⇒ – fruits, légumes, poissons – viandes	NOURRITURES 37 %
– congélation ⇒ – appertisation	TECHNOLOGIES 21 %
– dépannage ⇒	CONTEXTES DE CONSOMMATION 12 %
– lieux d'approvisionnement ⇒	ACCÈS AUX ALIMENTS 15 %
– problématique d'assimilation ⇒	CORPORÉITÉ/MANGER 15 %

rythme contraignant, obligeant les mangeurs à se satisfaire de repas rapides quasiment préparés à l'avance. Pour l'instant, restons simplement sur le constat que le discours sur les aliments participe à un discours sur les comportements alimentaires en général; il s'intègre dans les représentations de ce qu'est manger. Les aliments, toujours selon les auteurs, n'ont d'intérêt pour le mangeur que dans la mesure où ils peuvent être consommés.

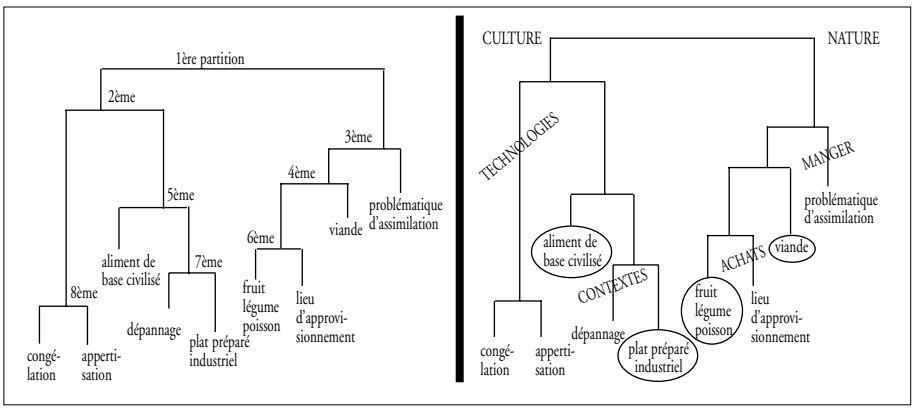
Le sens de la classification

Nous avons mentionné la façon dont Alceste fragmente le corpus en segments de texte, classifiés ensuite selon la technique de la classification descendante hiérarchique. La première représentation graphique de l'ordre de ce classement donne à voir une sorte de *généalogie de la classification*. Une généalogie implique *de facto* le facteur temporel, mais aussi, dans un sens, le facteur spatial. Etudier la distribution des classes permet de retracer leur partition successive par étapes à partir de l'arborescence du corpus. Pour les résultats de cette étude, les étapes sont au nombre de huit et se trouvent symbolisées graphiquement par les traits horizontaux. Les « morceaux » de corpus (uce) restant solidaires à l'issu de chaque étape de partition sont eux symbolisés graphiquement par les traits verticaux.

Nous avons opté pour une présentation en « face à face » de l'arborescence (à gauche) et de celui de la répartition des classes et des thématiques (à droite), afin d'illustrer le cheminement interprétatif allant de l'un à l'autre.

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

Figure 2. – La structure de l'arborescence (à gauche) et la répartition des classes et des thématiques (à droite) (Masson et Moscovici, 1997).



D'emblée, la structuration bipolaire de l'arborescence que l'on voit à gauche du graphique laisse apparaître une dichotomie de taille. D'un côté, nous retrouvons le schéma basique et primaire de ce que c'est manger (nourritures, accès, assimilation). De l'autre, elle rassemble les classes appartenant aux thématiques de la technologie, des contextes et, partiellement, des classes traitant des nourritures dans un univers fortement marqué par la civilisation. À la vue du contenu des classes terminales, les auteurs concluent que cette première partition correspond à l'opposition *nature-culture*, illustrée par la partie droite de la figure.

L'interprétation du contenu des classes lexicales et l'analyse de l'arborescence de la classification descendante hiérarchique ont mis en lumière l'ancienne opposition nature-culture, qui se traduit dans ce corpus à travers une différenciation entre « aliments naturels » et « aliments culturels ». Les premiers sont issus d'une nature domestiquée (produits animaux issus de l'élevage), ou encore organique (fruit-légume-poisson) toutes deux dénouées de tout aspect « sauvage », tout comme les « authentiques nouveaux territoires » (grandes surfaces, boucheries etc.) qui en tant que lieux d'approvisionnement remplacent les anciens territoires de chasse. Les seconds (aliments de base civilisés, plats industriels) représentent une culture « austère et pauvre. Il n'y est question ni d'art ni de convivialité, mais essentiellement de techniques et de solitude. Ainsi à une nature domestiquée s'oppose une culture industrielle et technologique. Les nourritures qui s'y trouvent portent la trace de cette culture : elles sont conditionnées, normées, calibrées. »

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES

En guise de conclusion

L'apport démonstratif que les méthodes hypothético-déductives autorisent reste tributaire d'un raisonnement interprétatif d'ensemble (Passeron, 2002). Alceste, et surtout les utilisateurs d'Alceste, n'échappent pas à cette règle. L'apport de cette méthode – à travers les classes lexicales et l'étude de leurs liens, le fond associatif ou topique qui les sous-tend, les impressions et la structure sémantiques qu'elles donnent à voir, le dictionnaire du vocabulaire du corpus qu'on obtient –, est considérable et il l'est d'autant plus qu'il est considéré par son propre auteur comme une approche complémentaire à l'analyse de contenu et une aide à l'interprétation des corpus. Mais au-delà des calculs automatiques, il ne faut oublier ni le travail réflexif qui précède l'analyse (intuitions, hypothèses de recherche), ni le cadre théorique dans lequel l'usage de telle ou telle méthode doit trouver sa légitimité et son adéquation avec *la façon de penser le fait psychosocial*, ni, enfin, la démarche finale consistant à interpréter et à donner à comprendre aux autres.

Nous avons essayé de montrer à travers un exemple empirique, l'apport analytique d'Alceste pour l'étude des représentations, notamment, par la symétrie que nous pouvons établir entre l'approche des mondes lexicaux et l'exploration du sens commun. Des mondes où les mots forment des univers de référence « communs » pour les locuteurs et les lecteurs dans la dynamique temporelle de leur production et l'assise culturelle de leur signification, autrement dit, des lieux où les mots forment et opèrent ce *langage de représentations* auquel nous faisions allusion en débutant ce chapitre.

Cette réflexion va effectivement au-delà de la mise en mots des représentations, de leur inscription *dans* le langage et nous permet de les considérer *comme* un langage. Ce langage contient et combine mots et idées, images et concepts, perceptions et cognitions, observations et interprétations, remplissant à la fois une fonction de connaissance et de communication. Comme les images pour les historiens médiévistes – et pour paraphraser Jean-Claude Schmitt (2002, p. 37) –, ce langage sait plus d'une fois nous rappeler que sa fonction est moins de *représenter* une réalité extérieure à lui que de *faire être le réel* sur un mode qui lui est propre. En prenant en considération cette mission du langage, nous pouvons atteindre une explication plurielle qui s'attache à l'interprétation des expressions culturelles dans leur apparence énigmatique (Geertz, 1986).

Document téléchargé depuis www.cairn.info - Université Louis Lumière Lyon 2 - - 159.84.125.51 - 27/04/2015 18h37. © ERES