

CRÉDOC

CAHIER DE RECHERCHE

ÉVOLUTION DES REPRÉSENTATIONS SOCIALES DU BIEN MANGER

Thierry MATHE
Diane BELDAME
Pascale HEBEL

■ DÉCEMBRE 2014



Sommaire

SYNTHESE	4
1.1 Classification descendante d'Alceste – meilleure méthode.....	5
1.2 Le « bien manger » entre 1988 et 2013 : du plat à partager au manger sain fait soi-même 6	6
1.2.1 Femme-nourriture : un lien persistant.....	6
1.2.2 2013 : le « fait maison » allie produits sains et plaisir de cuisiner	7
1.2.3 L'approche restrictive de l'alimentation apparue en 2007 s'infléchit en 2013	7
INTRODUCTION	10
2 ANALYSE LEXICALE POUR L'ETUDE DES REPRESENTATIONS SOCIALES	13
2.1 Représentations sociales	13
2.2 La méthodologie de l'analyse lexicale.....	14
2.3 Diversité de logiciels d'analyse textuelle	16
2.3.1 Analyses statistiques développées par les logiciels d'analyse textuelle.....	21
2.3.1.1 Faire apparaître les proximités du langage.....	21
2.3.1.2 Créer des classes de représentations sociales	21
2.3.1.2.1 L'analyse factorielle des correspondances	21
2.3.1.2.1 L'analyse factorielle multiple sur tableau de contingence	22
2.3.1.2.2 La classification descendante hiérarchique au sens de Reinert.....	22
2.4 Choix d'Alceste comme logiciel adapté à l'analyse des représentations sociales	23
2.4.1 Clefs de choix logiciels	23
2.4.2 Les logiciels adaptés pour traiter ce type de corpus	23
3 LES REPRESENTATIONS SOCIALES DU « BIEN MANGER » EN 2013	25
3.1 Matériel et méthodes	25
3.1.1 Eléments d'analyse.....	25
3.1.2 Recueil de données.....	26
3.1.3 Pré-traitements du corpus	27
3.1.4 Analyses statistiques descriptives	27
3.1.5 Classification et caractérisation des classes.....	28
3.2 Résultats.....	29
3.2.1 Description du corpus	29
3.2.2 Catégorisation du vocabulaire selon le sexe, l'âge, la CSP et l'IMC	29
3.2.3 Six classes de représentations du bien manger	36
3.2.3.1 Classe 1 : Exemples de plats et d'aliments, 20%.....	38
3.2.3.2 Classe 2 : Plaisir de manger à satiété, 11%	39
3.2.3.3 Classe 3 : Plaisir de manger et de l'acte social, 19%	40
3.2.3.4 Classe 4 : Manger équilibré, 26%	41
3.2.3.5 Classe 5 : discours nutritionnel, 9%.....	42
3.2.3.6 Classe 6 : Fait maison, 15%	43
4 ETUDE EN EVOLUTION DES REPRESENTATIONS DU BIEN MANGER DE 1988 A 2013	46

4.1	Enjeux autour de l'évolution des représentations sociales du « bien manger »	46
4.2	Matériel et Méthodes.....	48
4.3	Résultats.....	49
4.3.1	Des corpora différents.....	49
4.3.2	D'importantes variations de vocabulaire	52
4.3.3	1988-1995 et 2007-2013 : deux univers contextualisés	56
4.3.4	Evolution des représentations du « bien manger » par analyse factorielle des correspondances : l'espace des représentations.....	58
4.3.5	Evolution des représentations du « bien manger » par comparaison de classifications	62
4.3.5.1	Les proximités entre 2013 et 1995	64
4.3.5.2	L'année 2013 en recentrage par rapport aux préoccupations nutritionnelles de 2007 .	65
CONCLUSION		67
INDEX DES TABLEAUX		69
INDEX DES FIGURES		70
BIBLIOGRAPHIE		71

SYNTHESE

Depuis plus de trente ans, le CREDOC réalise des analyses des représentations sociales en utilisant l'analyse lexicale. Le principe consiste par une formulation spécifique à ce que la personne interrogée propose les notions sur lesquelles reposent sa conception et ses émotions de l'objet étudié. Chaque individu donne sa représentation mentale de l'objet. L'ensemble des représentations mentales de la population française permet ensuite d'établir les représentations sociales de l'objet. Les premiers travaux du département « Consommation » sur l'alimentation au début des années 1990 ont été menés en utilisant des formulations permettant d'accéder aux représentations mentales. L'alimentation, parce qu'elle a de tous temps occupé une place vitale dans la vie humaine, est profondément liée à la culture. Elle est le support de nombreux rites, légendes, récits. L'alimentation pose la question de l'incorporation (l'homme est ce qu'il mange). L'analyse de cet objet complexe qu'est l'alimentation est d'autant plus intéressante qu'elle est délaissée par la sociologie française. Après le développement de SPAD-T qui analysait les mots comme des variables classiques pour l'analyse des questions ouvertes de l'enquête « Aspirations et Conditions de Vie », l'équipe Consommation du CREDOC a développé avec Max Reinert un outil beaucoup plus stable et adapté aux représentations sociales, notamment celles de l'alimentation. Les travaux de Saadi Lahlou pour sa thèse de doctorat en psychologie sociale (Lahlou, 1995) ont permis grâce à cet outil de modéliser les représentations du « bien manger », en faisant émerger des noyaux de sens stables et reproductibles de la pensée humaine. Il a ainsi été démontré que la représentation sociale sert à créer de la réalité. Cet outil est essentiel pour comprendre les changements sociaux et culturels. Les recherches en analyse lexicale menées au CREDOC n'ont pas été renouvelées depuis 1998. Il était temps de faire un état des lieux des avancées de la recherche sur cet outil, d'autant qu'avec l'avènement du Big Data, il est très utilisé. L'application de l'outil sur le même objet avec la même formulation était possible puisque l'enquête existe toujours dans le système « Comportements et Consommation Alimentaire en France » (CCAF). Nous avons souhaité analyser le même objet, un quart de siècle après, pour faire émerger les changements sociaux et culturels autour du fait alimentaire.

1.1 Classification descendante d'Alceste – meilleure méthode

Les principales évolutions concernent les possibilités de traiter des textes de plus en plus importants avec des outils qui en font l'analyse. Les outils les plus récents utilisés dans le cadre du Big Data sont les suivants :

- la recherche d'information (RI) s'attache à sélectionner automatiquement les mots-clés de documents volumineux en vue de les retrouver grâce à des requêtes. C'est l'art et la manière de retrouver une information dont on sait qu'elle est présente, mais parfois difficilement accessible, dans une collection de texte.
- le Traitement Automatique des Langues (TAL) ou Traitement Automatique de la Langue Naturelle (TALN), à la frontière avec l'intelligence artificielle, identifie des syntagmes pour les reconstruire ensuite. Les domaines d'applications sont d'abord la traduction automatique, la correction orthographique ou la génération automatique de texte, comme par exemple la saisie intuitive T9© des téléphones mobiles. Elle trouve aussi des applications en traitement du signal (ex : reconnaissance vocale), en RI et en Text-Mining.
- l'analyse sémantique latente (LSA) construit des concepts avec les termes des documents pour les relier entre eux. A la croisée du TAL et de la RI, elle est utilisée pour analyser des processus cognitifs ou étudier synonymie et polysémie des mots.

Ces outils permettent de faire de l'analyse textuelle, afin d'en extraire des noyaux de sens qui correspondent à des représentations sociales. Seuls six logiciels traitant du texte pour l'analyser ont été identifiés. Les méthodes d'agrégation de réponses proches n'ont pas évolué en 25 ans, il s'agit soit d'analyse des correspondances soit de méthodes de classification. Les apports des nouvelles technologies pour l'analyse lexicale mise au point par Max Reinert en font la méthode la plus stable (comparativement à Cohéris Spad). Cette stabilité est obtenue grâce à l'enchaînement de plusieurs outils d'analyse de données classiques. La méthode ALCESTE commence par une classification descendante, ce qui permet de ne pas inclure les petits noyaux très typés qui provoquent dans une analyse factorielle la création d'axes liés à des artefacts n'ayant pas de sens lorsque l'on travaille sur du sens et des mots. Ensuite le logiciel reclasse les points peu typiques et crée un groupe de non classés. Une analyse factorielle est ensuite réalisée sur l'appartenance aux classes. De nouveaux logiciels sont apparus, très complets comme IraMuteQ basés sur le même algorithme qu'Alceste et plus complets avec l'analyse des similitudes, mais leur manque de convivialité nous a conduits à conserver Alceste pour la classification. L'analyse en évolution des réponses à la question « Pour vous, qu'est-ce que bien manger ? », a combiné plusieurs approches, les comparaisons des mots et des classes entre les différentes années, mais aussi des analyses factorielles classiques. Comme l'ont montré tous les travaux des 25 dernières années sur l'analyse

lexicale au CREDOC, ces outils peuvent être appliqués à de très nombreux corpus (presse, textes littéraires, objets complexes, ...) et fondent de nombreux espoirs pour l'avenir du Big Data.

1.2 Le « bien manger » entre 1988 et 2013 : du plat à partager au manger sain fait soi-même

1.2.1 Femme-nourriture : un lien persistant

Le bien manger ne se conçoit pas de la même façon selon que l'on soit un homme ou une femme. La mère, et ce dans toutes les cultures, sécurise son enfant en le nourrissant et notamment dès la naissance en l'allaitant. La mère garde ce rôle de prévention en santé au sein de la famille. Les femmes sont alors nettement plus sensibles aux liens entre alimentation et santé. Les mots caractéristiques des femmes appartiennent au champ lexical de la nutrition et de la diététique. On retrouve ici l'image du corps particulière à la femme qui craint de prendre du poids, concomitante à celle de suivre un régime. Chez les femmes les représentations du corps, les conceptions de la santé et les représentations ancestrales se télescopent et mettent en évidence un mélange confus sur l'alimentation.

Pour l'homme, le bien-manger est associé à la femme. Dans les cultures patriarcales, la femme ne peut être que mère et nourricière. Ces représentations anciennes témoignent d'une approche attachée à la satiété, au plaisir du ventre, au goût, et au plaisir du repas partagé (« amis »). Chez les 15-24 ans, on retrouve la place de la « mère » nourricière, puisqu'ils associent bien-manger avec mère.

La distinction sociale, chère à Bourdieu, est toujours présente. Les ouvriers associent le bien manger au copieux alors que les cadres l'associent au goût, à la variété.

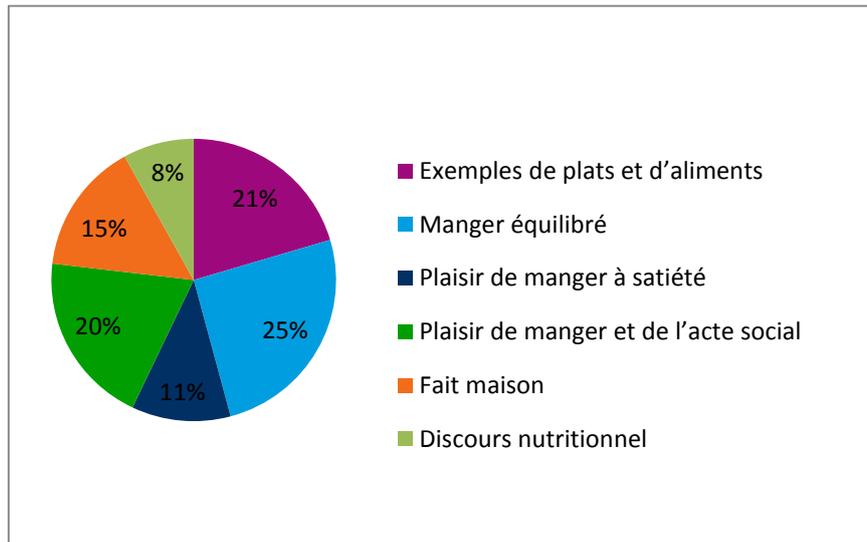
Figure 1 - Nuage de mots des formes graphiques caractérisant le sexe



Source : Enquête CCAF 2013, CREDOC

1.2.2 2013 : le « fait maison » allie produits sains et plaisir de cuisiner

Figure 2 - Répartition des classes lexicales du corpus "bien manger" en 2013 :



L'apparition d'une classe du « *fait maison* » en 2013 est significative d'un retour du plaisir dans l'alimentation, tant du point de vue de la préparation que du point de vue du goût attendu. Elle illustre aussi un déficit de confiance vis-à-vis des produits industriels tels que les produits transformés ou les plats préparés, faisant percevoir le « *fait maison* » comme garantie d'un « *bien manger* ». La crise de la fraude à la viande de cheval vendue comme du bœuf en 2013 est passée par là et à contribuer à créer un climat de défiance marqué vis-à-vis des produits industriels. D'où le désir de revenir au local, comme garantie plus forte de produits sains. A ce titre, le « fait maison » apparaît aux yeux des consommateurs comme une garantie de manger plus sainement qu'avec des plats tout-prêts. Cuisiner est aussi une source de plaisir en soi, et en perspective du point de vue du goût et de la convivialité, car on ne cuisine guère pour soi uniquement : la manière de s'alimenter des personnes seules en témoigne.

1.2.3 L'approche restrictive de l'alimentation apparue en 2007 s'infléchit en 2013

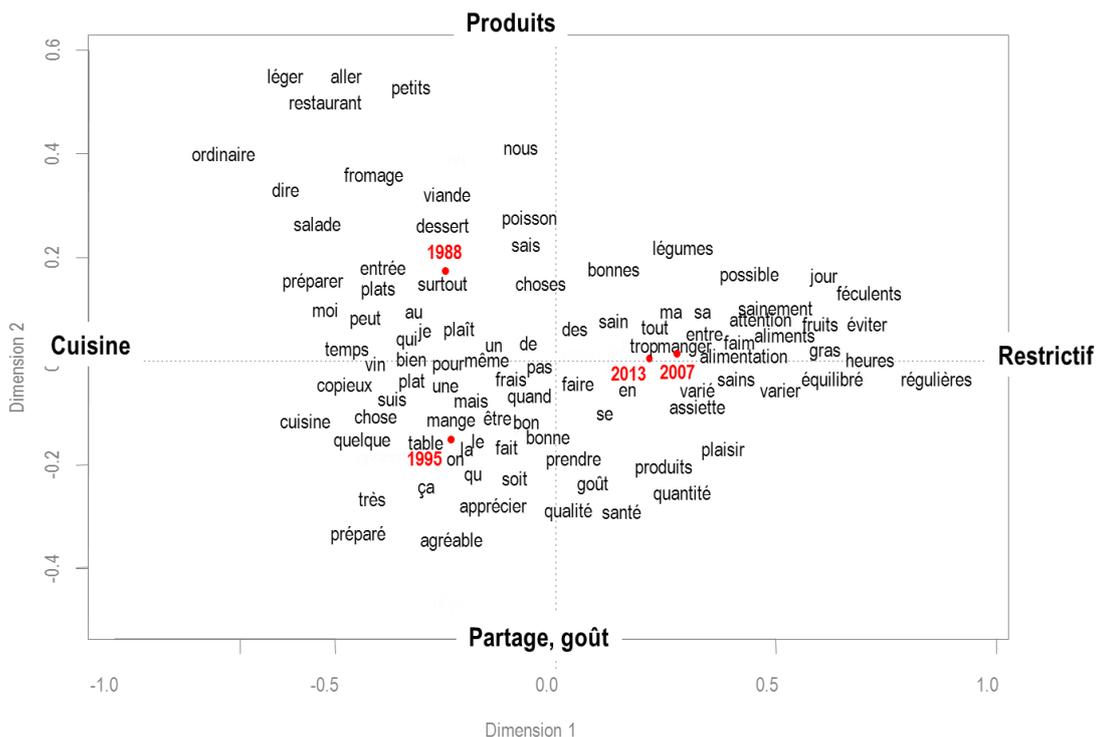
Le champ lexical de la restriction et de la surveillance ne s'est pas démenti depuis son apparition en 2007. En 2013, les formes « *raisonnables* » et « *raisonnable* » sont principalement assorties du mot « quantité » : il faut manger (de tout) en quantité raisonnable. « *Grignoter* » (+1125), forme nouvelle en 2013, est complétée par l'idée que l'on doit éviter de grignoter « entre » (+73) les repas.

La prise en compte de l'année 2007 montre cependant une inflexion en 2013 relativisant la progression de l'approche nutritionnelle que dessine la courbe 1988-2013. Les préoccupations des Français semblent s'éloigner des injonctions nutritionnelles (classe du discours nutritionnel : 9% en

2013) sans cependant renier le souci de manger « équilibré » (26%). L'engouement pour le fait « maison » (+54) et la ferveur pour la gastronomie à la française des « belles assiettes » ou des « bonnes assiettes » est fortement présent en 2013, traduisant le retour du plaisir, même si cela n'efface pas les inquiétudes sur les risques alimentaires. En 2013 en effet, la santé n'est pas perçue comme mise en danger par une mauvaise alimentation qui serait le fait de son propre comportement (et dont on se sentirait coupable) mais du fait des produits eux-mêmes (tromperie, contaminations-dont OGM, défaut de traçabilité, intermédiaires trop nombreux).

Viennent ensuite les composants du repas qui semblent ne plus correspondre à la vision qu'ont les Français du bien manger en 2013 : exit le « café » (-998), les « sauces » (-920) et la « sauce » (-316), « la charcuterie » (-228) et dans une moindre mesure « le vin ». Le recul de « hors d'œuvre » (-936) et « mets » (-378) est peut-être autant dû à l'évolution du vocabulaire qu'à la déconstruction et la simplification des repas.

Figure 3 - Plan factoriel de l'analyse des correspondances du tableau lexical agrégé par année étudiée



Source : Enquêtes CCAF 1988, 1995, 2007 et 2013, CREDOC

L'analyse est riche d'enseignements sur l'évolution des représentations mentales de l'alimentation. Le lien fait entre l'alimentation et la santé en 2007, dont la progression est considérable, a pu être analysé au moins en deux courants : la recherche de l'équilibre et l'alimentation vue comme un problème. Cette dernière modalité constitue alors une nouveauté qui vient contredire l'approche traditionnelle de l'alimentation en France (plus familière, en particulier, aux aspects gustatifs et conviviaux). Les Français, qui ont été l'objet d'une diffusion massive de messages nutritionnels

induisant une approche plus utilitariste de l'alimentation, à partir du premier PNNS en 2001, se trouvent tiraillés entre la tradition culinaire nationale, une vision plus restrictive de l'alimentation et l'augmentation des pathologies du comportement alimentaire.

Il apparaît aujourd'hui que les injonctions nutritionnelles ne sont pas nécessairement productrices d'améliorations en termes de santé publique. Les nouvelles orientations prises par le PNA témoignent de cette prise de conscience du terreau culturel sur lesquels sont bâtis les modèles alimentaires. Il est à cet égard significatif que le programme national « nutrition-santé » (PNNS) ait été remplacé en 2011 par un programme national « alimentation » (PNA). Celui-ci met explicitement en avant les règles du « modèle alimentaire français » comme permettant d'expliquer en partie « *qu'en France, seulement 14% des adultes ont des problèmes d'obésité contre plus de 30% aux Etats-Unis* ». L'appel aux caractères essentiels du modèle alimentaire français laisse penser que c'est en s'appuyant sur ceux-ci qu'il devient possible d'introduire des principes de nutrition et d'équilibre, sans quoi tout enracinement est voué à l'échec.

INTRODUCTION

Chaque société est le résultat d'une synthèse historique présentant un ensemble de manières d'agir, de sentir et de penser exprimant son unité et sa singularité. Parmi ces manières, celles relatives à l'alimentation, parce qu'elles occupent une place vitale dans l'existence des individus et des groupes, constituent de profonds marqueurs culturels. La culture alimentaire est d'abord une production des représentations sociales propre à une population. Face au même impératif de se nourrir, les sociétés ont répondu et répondent différemment. Si une culture alimentaire correspond à un ensemble de pratiques observées par telle population, ces pratiques traduisent une interprétation de ses représentations alimentaires, qui vont s'incarner dans un certain nombre de règles : que faut-il manger ou ne pas manger, comment faut-il manger, etc. ?

En d'autres termes, l'acte alimentaire répond tout autant à des besoins symboliques qu'à des besoins biologiques, ce que signifie précisément le « paradoxe de l'*hominivore* » (Fischler, 1990), biologiquement capable de manger de tout mais symboliquement structuré par des prescriptions et des interdits. Le paradoxe de l'alimentation réside aussi dans son caractère salvateur (apporter de l'énergie, des nutriments) en même temps que maléfique (empoisonner), que désigne le « paradigme d'incorporation » (Fischler, 1990).

En France, les manières de s'alimenter reposent traditionnellement sur un certain nombre de règles qui favorisent à la fois le vivre-ensemble, le goût des aliments et leur variété. Le rythme et la structuration des prises alimentaires, fortement inscrits dans les habitudes des Français, n'ont été affectés qu'à la marge par le développement du grignotage en dehors des repas et par une tendance à la simplification des repas (2 séquences au lieu de 3 ou 4 par exemple). C'est en ce sens que l'on parle de modèle alimentaire français, en faisant référence à une culture alimentaire particulière, c'est-à-dire une culture alimentaire nationale définissant les principes d'alimentation dominants de la population française.

A partir des années 1990 cependant, un discours nutritionnel a commencé à être diffusé auprès d'un large public. Les différentes crises sanitaires (depuis celle liée à l'ESB, en 1996 et en 2000, jusqu'au scandale à la viande de cheval vendue comme étant du bœuf en 2013), fortement médiatisées, ont contribué à faire naître dans la population une préoccupation vis-à-vis de son alimentation du point de vue de la santé. L'augmentation du nombre de personnes en surpoids ou obèses, des cancers et des maladies cardio-vasculaires ont contribué à une telle prise de conscience. Parallèlement, le Programme National Nutrition Santé (PNNS) de 2001 a largement contribué à diffuser des messages sanitaires, participant à faire évoluer les représentations de l'alimentation. On a observé sur l'enquête de 2007 que la dimension « santé » du fait alimentaire dominait les représentations, bousculant le rapport traditionnel des Français à leur alimentation

(Pilorin et Hébel, 2012). S'agissait-il d'un phénomène transitoire ou de l'épisode significatif d'un processus en cours ?

Le CREDOC observe les évolutions des consommations et des comportements alimentaires des Français depuis plus d'une vingtaine d'années. Les représentations que se font les Français de l'alimentation, à travers la question « *Pour vous, qu'est-ce que bien manger ?* », sont un facteur explicatif de ces évolutions. Nous avons donc souhaité analyser les réponses induites par cette question, avec des méthodes d'analyse lexicale historiquement mises en œuvre au CREDOC, pour appréhender les évolutions des représentations sociales de l'alimentation entre 1988 et 2013.

Les premiers travaux en analyse lexicale au CREDOC datent du début des années 80 (Brian, 1983, 1984, 1986), Ludovic Lebart et Monique Bécue développent le logiciel SPAD.text. Fortement influencé par l'analyse des données à la française, l'outil proposé considère chaque mot comme une variable binaire et chaque individu est décrit par autant de variables qu'il existe de mots dans l'ensemble du corpus. Le tableau est ensuite traité avec les méthodes classiques d'analyse des données. Les tableaux étant « hypercreux », c'est-à-dire comportant principalement des zéros, les résultats sont très instables. De nouveaux développements étaient nécessaires. Le département consommation du CREDOC développe alors de nouvelles méthodologies pour modéliser les représentations (notamment celles de l'alimentation) avec l'émergence d'une tendance psychosociale dans l'analyse lexicale (Beaudouin et al, 1993). C'est à cette époque qu'est conçu le logiciel ALCESTE débutant par une lemmatisation du corpus afin de d'obtenir des tableaux moins creux. Des corpus de représentation très importants ont été, depuis, constitués en historique, en particulier autour de la notion de « bien manger ». La question « Pour vous, qu'est-ce que 'bien manger' ? » a été choisie par Saadi Lahlou au CRÉDOC en 1988. Cette question ouverte a été posée par la suite en face à face lors des enquêtes Comportements alimentaires en France (CAF) réalisées en 1995, 2000, 2003, 2007, 2010 et 2013.

La question choisie incite la personne interrogée à chercher les notions sur lesquelles reposent sa conception d'une « bonne » alimentation, indépendamment de son comportement alimentaire effectif. A partir de la question ouverte, elle est invitée à formuler des comportements alimentaires idéaux. Il ne s'agit donc pas de lui demander de décrire ses pratiques mais de formuler des conceptions, qui pourraient dans un second temps permettre de comprendre le sens de ses pratiques. On comprend alors comment « certains comportements qui paraissent irrationnels suivent en fait une rationalité, celle de l'enchaînement des représentations » (Lahlou, 1992). En d'autres termes, il s'agit de saisir les représentations d'une alimentation « bonne » selon la personne, et non ses habitudes alimentaires. Dans cette optique, l'objet de l'analyse lexicale est de travailler sur la matière même du discours spontané des personnes ayant répondu au questionnement sur le « *bien manger* ».

Les derniers travaux de recherche sur l'analyse lexicale datant de 1998 (Collerie de Borély, 1998), nous commençons par un point sur les nouveaux logiciels de traitements de données lexicales, en identifiant ceux qui permettent de mieux analyser les représentations sociales. Puis en partant des résultats de l'enquête de 2013 sur les représentations alimentaires des Français, nous observons et tentons de comprendre leur évolution depuis 1988, grâce à la méthodologie Alceste combinée à des outils de statistiques lexicales descriptives. Enfin, il s'agira d'examiner l'évolution du vocabulaire au moyen de l'analyse factorielle, ainsi que d'effectuer une classification des représentations dans le temps, et des statistiques lexicales, pour illustrer les mutations à l'œuvre dans les représentations du bien manger en 1988, 1995, 2007 et 2013, pour faire suite aux travaux précédents (Mathé, Pilorin et Hébel, 2008).

2 ANALYSE LEXICALE POUR L'ETUDE DES REPRESENTATIONS SOCIALES

2.1 Représentations sociales

Emile Durkheim fut le premier à évoquer la notion de représentations qu'il appelait « collectives ». La manière dont nous nous représentons le monde détermine notre comportement, les façons dont nous agissons et réagissons dans une situation particulière. Cette manière dépend de notre caractère, de notre histoire de vie, mais aussi de l'univers dans lequel nous évoluons, du milieu culturel, social, national. La construction de la réalité est celle d'une réalité (personnelle) enchâssée dans un regard culturel, social, économique, familial, etc.

Le concept fut formalisé plus tard par Moscovici (1984), qui décrit les représentations sociales comme imbriquées dans un système plus global que l'on appelle le « sens commun ». C'est en effet « la construction d'une réalité commune à un ensemble social », que Jodelet conçoit comme « des modalités de pensée pratique orientées vers la communication, la compréhension et la maîtrise de l'environnement social, matériel et idéal ». Cela correspondrait à « une forme de connaissance spécifique, le savoir de sens commun » (Jodelet, 1997).

Une représentation sociale est donc une interprétation de la réalité croisant l'individu, l'objet et le groupe. C'est une représentation mentale impactée par les éléments extérieurs (milieu d'origine, situation d'existence, interactions sociales, expériences vécues, contexte psychosocial de l'individu, etc.). Cette opération procède par catégorisation (Lahlou, 1992) ou, si l'on se réfère à la sociologie du « sens commun », par typifications (Schütz, 1987), qui est une manière ordinaire d'accéder à la connaissance du monde vécu. Schütz avance en effet qu'il y a analogie des manières de connaître tant du point de vue de la connaissance commune, celle de l'homme ordinaire dans sa vie quotidienne, que du point de vue de la connaissance scientifique. Il y aurait ainsi une différence de degré mais non de nature dans la manière de procéder.

Si une représentation est individuelle, elle se construit nécessairement à l'intérieur d'une société, d'une culture, d'un substrat collectif (composé de la participation de l'individu à différents groupes à l'intérieur d'un ensemble (famille, école, quartier, groupes de pairs, communauté religieuse, association, etc.)). C'est ce qui permet que se forment des manières communes de se représenter la réalité, sans qu'il s'agisse d'y voir une uniformisation de la pensée. La possibilité qu'il y ait formation de représentations communes est ce qui permet à un groupe, à une culture, à une société d'exister. Ceci non seulement de manière synchronique (horizontale) mais aussi dans le

temps (transmission de valeurs, de manières de voir, de conceptions, etc.), inscrivant le groupe, la société ou la culture dans la durée, et plus précisément une durée porteuse d'évolution.

Ce sont donc des conceptions partagées de ce qu'est le monde, de ce que sont les objets qui nous entourent, et des manières de se comporter par rapport à ces réalités figurées.

Lorsqu'il a mis en place son protocole d'enquête sur les comportements alimentaires en 1988, Saadi Lahlou (1992) souhaitait expérimenter une méthode d'analyse lexicale appliquée aux représentations sociales du « bien manger », à partir d'un corpus constitué des associations libres de 2000 individus, recueilli lors d'une enquête. Les réponses, données oralement puis transcrites littéralement par l'enquêteur, avaient été traitées selon une nouvelle méthode d'analyse lexicale fondée sur les occurrences de mots mais aussi de sens. Elle permettait de quantifier les réponses obtenues par cette question ouverte, puis de regrouper les phrases contenant des mots ayant la même racine, ainsi que les phrases n'ayant aucun mot en commun mais dont le sens était proche. Cette méthode va donc au-delà d'une analyse par comptage de fréquences des termes du discours spontané des enquêtés en mettant au jour les champs sémantiques des individus.

Quand Saadi Lahlou en 1988 pose la question ouverte « Qu'est-ce que pour vous « bien manger » ? » il pousse l'enquêté non pas à donner son opinion ou son attitude par rapport à l'alimentation, il oblige à formuler des comportements alimentaires idéaux, recommandables, des « *modes d'emploi du monde* ». L'analyse lexicale est un moyen de capter ces représentations, les synthétiser et les retranscrire.

Le corpus, constitué par la réponse à cette question est donc particulier en ce sens que :

- Le vocabulaire est redondant, en écho à la question, un phénomène que (Lahlou, 1992) appelle « écholalie » ;
- Il y a autant d'auteurs que de textes dans le corpus ;
- Les textes sont particulièrement courts et leur taille varie peu ;
- Le discours est explicatif et/ou descriptif (plutôt que narratif ou argumentatif).

2.2 La méthodologie de l'analyse lexicale

Dans l'analyse lexicale nous cherchons à extraire et synthétiser les concepts abordés dans un ensemble de réponses à une question d'enquête, matériau textuel informatisé formant un corpus (*pl.* corpora). L'outil statistique utilisé pour traiter cette information est l'analyse de données « à la française » qui s'inscrit dans l'approche fréquentiste issue de l'école benzécriste explicitée dans *Pratique de l'analyse des données : linguistique et lexicologie* (Benzécri, 1981). Le principe de la lexicométrie consiste à structurer les données textuelles en un tableau « individus x variables » (les variables étant les mots) sur lequel seront appliquées des techniques d'analyse des données -

ou d'« analyse factorielle » (école française). Ces méthodes de la statistique multivariée sont exploratoires, c'est-à-dire que les données sont observées sans *a priori* : on cherche à représenter graphiquement un tableau de données d'après ses plus grandes dimensions de variabilité, sans les soumettre. On ne cherche pas à prévoir des comportements grâce à un modèle statistique qui permettrait d'inférer la structure des données. Une série de transformations sont mises en œuvre sur le tableau à analyser afin de pouvoir ensuite le diagonaliser. Ce procédé conduit à l'obtention de vecteurs propres et valeurs propres qui traduisent la nature et l'intensité du lien entre les lignes (individus) et les colonnes (en analyse lexicale : les mots) de la matrice diagonalisée. Benzécri explique d'ailleurs simplement qu' « *une analyse des données n'est, en bonne mathématique, qu'une recherche de vecteurs propres ; toute la science, ou tout l'art... étant de savoir quelle matrice traiter* » (Benzécri, 1973) cité dans (Rouanet & Lépine, 1976). La méthode de prédilection pour l'analyse des données textuelles est l'Analyse Factorielle des Correspondances (AFC). Dans le cadre de cette analyse, les dimensions et le contenu du tableau « individus x mots » sont sujets à interrogations (Beaudouin, 2000). En effet, ce dernier peut être amputé du vocabulaire rare, allongé par segmentation du texte ou réduit par agrégation en lemmatisant les formes graphiques. Autant de choix méthodologiques qui impliquent une réflexion en amont lors du choix du logiciel qui exécutera les calculs.

L'analyse des données textuelles n'est pas l'apanage des traitements qu'il est possible d'expérimenter sur le texte. Il existe d'autres façons, statistiques ou non, de traiter le matériau textuel. Quelques exemples sont présentés ci-dessous :

- la recherche d'information (RI) s'attache à sélectionner automatiquement les mots-clés de documents volumineux en vue de les retrouver grâce à des requêtes. C'est l'art et la manière de retrouver une information dont on sait qu'elle est présente, mais parfois difficilement accessible, dans une collection de texte.
- le Traitement Automatique des Langues (TAL) ou Traitement Automatique de la Langue Naturelle (TALN), à la frontière avec l'intelligence artificielle, identifie des syntagmes pour les reconstruire ensuite. Les domaines d'applications sont d'abord la traduction automatique, la correction orthographique ou la génération automatique de texte, comme par exemple la saisie intuitive T9© des téléphones mobiles. Elle trouve aussi des applications en traitement du signal (ex : reconnaissance vocale), en RI et en Text-Mining.
- l'analyse sémantique latente (LSA) construit des concepts avec les termes des documents pour les relier entre eux. A la croisée du TAL et de la RI, elle est utilisée pour analyser des processus cognitifs ou étudier synonymie et polysémie des mots.

On ne s'intéressera pas ici à ces autres cadres de travail, qui formalisent des théories issues de disciplines telles que la linguistique ou les sciences cognitives (Fallery & Rodhain, 2007).

La spécificité du traitement de mots par rapport à des variables numériques est que le vocabulaire étant très large les variables analysées sont très nombreuses. Pour rechercher du sens, plutôt que le mot, c'est bien la racine du mot qui est la plus intéressante. L'analyse lexicale quand elle cherche à analyser les représentations mentales est alors réaliser après une lemmatisation du texte. Cette opération a l'avantage statistique, en plus de l'avantage analytique, de réduire le nombre de variables et de créer une matrice moins creuse. Les logiciels utilisés doivent avant tout être conçu de cette façon. La lemmatisation est une opération de modification du corpus qui consiste à réduire le nombre de formes graphiques en procédant à des associations de telle sorte que :

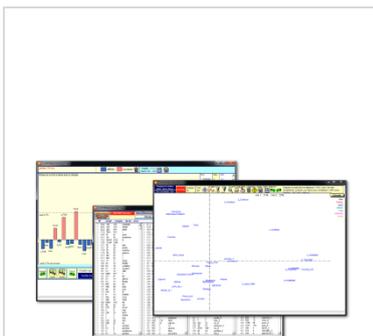
- Les formes verbales sont ramenées à l'infinitif : « *mangera* », « *mangeons* », « *manger* » deviennent « *manger+* » ;
- Les substantifs au singulier : « *fruits* » devient « *fruit+* » ;
- Les adjectifs au masculin singulier : « *pleines* », « *pleins* » deviennent « *plein+* » ;
- Les formes élidées à la forme sans élision : « *l'équilibre* » devient « *le* » et « *équilibre* » (Lebart & Salem, 1994).

La qualité de la lemmatisation est essentielle mais à fortement progressée dans les nouvelles versions des logiciels existants.

2.3 Diversité de logiciels d'analyse textuelle

L'analyse statistique des données textuelles, comme les autres disciplines de la Statistique, a pu bénéficier de l'augmentation de la puissance de calcul des micro-ordinateurs ainsi que de la multiplication et l'accessibilité des outils de programmation. Les premiers travaux de recherches et expérimentations sur le matériau textuel initiés dans les années 80 au CREDOC ont ainsi laissé place à un paysage logiciel d'analyse des données textuelles étoffé et amélioré.

Le paysage logiciel français de l'analyse des données textuelles reflète la façon dont l'équipe qui a développé l'outil appréhende la question textuelle. Les possibilités et fonctionnalités permises par les logiciels illustrent les convictions et les approches retenues par leurs concepteurs. On peut brosser un rapide portrait historique et fonctionnel des plus connus et usités.



Hyperbase : né en 1989 sous l'impulsion d'Etienne Brunet, à l'occasion du Bicentenaire de la Révolution française et à la demande du Centre Pompidou. Il répondait au souhait du commanditaire de mettre à disposition du public les textes ayant trait à la Révolution. Les fonctionnalités documentaires et

	<p>statistiques qui sont mises en œuvre dans ce logiciel reprenaient des programmes déjà écrits par Etienne Brunet à l'époque. Actuellement développé à l'unité de recherche « Bases, corpus, langages » de l'université de Nice et du CNRS sous l'égide de son créateur, Hyperbase est le seul logiciel à offrir une représentation sous forme d'analyse arborée des proximités textuelles (LUONG, 1988).</p>
	<p><i>COHERIS SPAD (anciennement SPAD-T)</i> : ou « Système Portable d'Analyse des Données » avec son module spécifique d'analyse Textuelle. Né dans le milieu des années 70 à l'initiative de Ludovic Lebart, Alain Morineau et Nicole Tabard, les programmes étaient initialement distribués librement et accompagnaient l'ouvrage « <i>Techniques de la Description Statistique</i> » des mêmes auteurs (MORINEAU, 2012). Dès le début des années 1990, avec l'engouement pour la statistique exploratoire, le logiciel connaît un développement quasi professionnel avant d'être cédé à une société à caractère commercial en 2000 pour devenir COHERIS SPAD. Il se positionne comme le premier logiciel exploratoire pour le traitement des gros jeux de données (MORINEAU, 2012).</p>
	<p><i>DTM-Vic</i> : DTM – Vic est l'acronyme de « Data and Text Mining – Visualization, Inference, Classification ». Ce programme trouve ses racines dans COHERIS SPAD jusqu'au moment où ce dernier est distribué à visée commerciale. Toujours développé et maintenu par Ludovic Lebart et son équipe, le logiciel est implémenté au fur et à mesure des publications de ses travaux. Il est le seul logiciel à proposer des techniques de validation des représentations par ré-échantillonnage (bootstrap), permettant ainsi d'évaluer leur sensibilité.</p>

	<p><i>Alceste</i> : Alceste est l'acronyme de « Analyse des Lexèmes Co-occurents dans les Enoncés Simples d'un Texte ». Il a été développé dans les années 90 au CNRS sur la base des travaux de Max Reinert (Reinert, 1983). Il était jusqu'il y a peu le seul logiciel à pouvoir exécuter une classification descendante hiérarchique sur des segments de texte, méthode originale et très utilisée en sociologie.</p>
	<p><i>Lexico</i> : Lexico est un logiciel initié par André Salem dont les premières versions (Lexico 1 et Lexico 2) mises à disposition dès 1995 étaient gratuites. La dernière version développée par l'équipe universitaire SYLED-CLA2T (Université Sorbonne Nouvelle – Paris 3), Lexico 3, fait l'objet d'une distribution commerciale depuis la fin des années 1990. Il se démarque par ses outils d'observation d'accroissement du vocabulaire.</p>
	<p><i>R, le package tm() et l'interface R.Temis</i> : R est à la fois un langage de programmation et un logiciel de statistique fondé par Ross Ihaka et Robert Gentleman au début des années 1990. Dans sa version de base, il contient toutes les fonctionnalités nécessaires à la statistique courante. Parce qu'il entre dans le spectre des logiciels libres (c'est-à-dire « dont l'utilisation, l'étude, la modification et la duplication en vue de sa diffusion sont permises, techniquement et légalement » (Wikipedia, 2014)), des bibliothèques dédiées à l'analyse des données textuelles ont été développées spécifiquement. Un des « packages » spécifique s'intitule tm() pour textmining (Feinerer, Hornik, & Meyer, March 2008) et son utilisation est facilitée par l'intégration sous forme de menus à cliquer grâce au package R.Temis (Bouchet-Valat & Bastin, 2013). Il est possible de développer et de partager ses propres développements pour en faire bénéficier la communauté d'utilisateurs.</p>

	<p>IRaMuTeQ : IRaMuTeQ est l'acronyme pour « Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires ». C'est un logiciel libre composé d'une interface programmée en python couplée au moteur statistique R. Il est développé par l'équipe REPERE au sein du LERASS (Laboratoire d'Etude et de Recherches Appliquées en Sciences Sociales) à Toulouse depuis 2008. Il propose une classification au sens de (Reinert, 1983) et une analyse des similitudes (Degenne & Vergès, 1973).</p>
---	--

L'ensemble des fonctionnalités et spécificités des logiciels décrits ci-avant est détaillé dans le Tableau 1. Ce tableau n'a pas fait l'objet d'entretiens avec leurs concepteurs : il est le fruit d'une lecture attentive de la documentation existante et de tests logiciels lorsqu'ils étaient possibles.

La moitié des logiciels sont payants avec un coût variant du simple au sextuple. Même si certains sont développés dans des laboratoires de recherche publics, les coûts de développements et de maintenance justifient le prix des licences. Seuls les logiciels R et IRaMuTeQ sont libres, c'est-à-dire modifiables pour des développements spécifiques si l'utilisateur dispose de connaissances en programmation suffisantes dans le langage de programmation. La plupart nécessitent une mise en forme particulière des données, plus ou moins laborieuse selon le format de données de départ et les encodages des caractères supportés. Une fois les données importées, la plupart des logiciels offrent la possibilité de lemmatiser.

Tous les logiciels permettent le comptage des formes retenues (lemmatisées ou non) et proposent des tris par occurrence. On peut ainsi examiner plus précisément le lexique, en particulier les hapax : les mots qui ne sont cités qu'une seule fois. Il en est de même pour les segments répétés : les logiciels détectent les « motifs » de texte et renseignent leur fréquence et leur position. Formes et segments répétés sont toujours repérables dans le texte grâce au concordancier, *id est* un renvoi au contexte dans le corpus dans lequel on trouve le mot choisi.

La différence essentielle est basée sur les outils d'analyses de données proposés.

Tableau 1 - Fonctionnalités des principaux logiciels d'analyse textuelle français

	Alceste	DTM-Vic	Hyperbase	Lexico 3	R & R.TemiS	Coheris SPAD	IraMuTeQ
Editeur commercial	Image			Université de la Sorbonne Nouvelle, Paris 3		Coheris	
Dernière version connue	Alceste plus	5.7	9	3.41	3.1.1	8	0.6 alpha 3
Derniers développements documentés en ligne	2012	2013	janvier 2011	février 2003	2013	2013	février 2013
Auteur(s)	Reinert Max	Lebart Ludovic	Etienne Brunet	Cédric Lamalle, William Martinez, Serge Fleury, André Salem	collectif	collectif privé	Pierre Ratinaud
Propriétaire/libre	Propriétaire	Propriétaire	Propriétaire	Propriétaire	Libre	Propriétaire	Libre
Prix (€) d'une licence simple	1745	Gratuit	Gratuit	600	Gratuit	3800	Gratuit
Lien web	http://www.image-zafar.com/fr/logiciel-alceste	http://www.dtmvic.com/	http://www.unice.fr/bcl/spip.php?rubrique38	http://www.tal.univ-paris3.fr/lexico/	http://www.r-project.org	http://www.coheris.com/produits/analytics/logiciel-data-mining/	http://www.iramuteq.org/
Format d'entrée spécifique des données	oui, format Alceste (**** texte)	non	oui, format Hyperbase (&&& texte &&&)	oui, format Lexico (<Variable=X> texte)	non	non	oui, format Alceste (**** texte)
Lemmatiseur intégré	oui	oui	oui	non	oui	oui	oui
Fréquence des formes	oui	oui	oui	oui	oui	oui	oui
Comptage des segments répétés	oui	oui	oui	oui	oui	oui	oui
Concordancier	oui	oui	oui	oui	oui	oui	oui
Co-occurrences	oui	oui	oui	oui	oui	oui	oui
Mesure d'accroissement du vocabulaire	oui	non	non	oui	non	non	non
AFC	oui	oui	oui	oui	oui	oui	oui
Classification descendante (au sens de Reinert)	oui	non	oui, couplage avec licence Alceste	non	non	non	oui
Analyse des similitudes	non	non	oui	oui	non	non	oui
Réseau de formes (représentation des proximités entre mots)	oui	non	non	non	non	non	non
Analyse arborée	non	non	oui	non	non	non	oui
Cartes auto-organisées de Kohonen	non	oui	non	non	oui, package Kohonen()	non	non
Techniques de validations des représentations factorielles	non	oui	non	non	non	non	non
Etude des marqueurs du discours	oui	non	oui	non	non	non	oui
Analyse Factorielle Multiple sur Tableau de Contingence (AFMTC)	non	non	non	non	oui, package FactoMineR	non	non

2.3.1 *Analyses statistiques développées par les logiciels d'analyse textuelle*

Les outils proposés sont différents d'un logiciel à l'autre. Avant de séparer le discours en classes, certains logiciels permettent de faire apparaître les proximités de langage, c'est le cas d'IraMuteQ qui ensuite permet d'utiliser différentes classifications. Ce logiciel est le plus complet.

2.3.1.1 Faire apparaître les proximités du langage

a. *L'analyse des similitudes*

C'est une technique de représentation issue de la théorie des graphes qui permet de synthétiser, en un arbre ramifié, la proximité et les relations d'un ensemble d'objets.

b. *Les réseaux de formes*

Le réseau de forme est une représentation graphique qui se focalise sur une seule forme et modélise, à la manière d'un arbre tronqué, les formes graphiques les plus proches sur la base de leurs cooccurrences. Plus une forme est citée de façon concomitante à la forme étudiée, plus elle est proche de cette dernière.

c. *L'analyse arborée*

Les analyses arborées se présentent également sous forme d'arbres. Comme pour l'analyse des similitudes, ils sont connexes (d'un seul tenant) et sans cycles. Ils représentent la distance, choisie par l'utilisateur, entre des textes.

d. *Les cartes auto-organisées de Kohonen*

Cette méthode de classification et de représentation est issue de la méthode de classification des nuées dynamiques (ou k-means). Elle se représente par une grille ou un ruban où les classes obtenues sont d'autant plus proches dans la grille qu'elles se ressemblent.

2.3.1.2 Créer des classes de représentations sociales

2.3.1.2.1 L'analyse factorielle des correspondances

C'est la méthode historique et la plus usitée. Elle permet de représenter les dépendances qui existent entre les formes graphiques et les individus. La métrique utilisée est la distance du χ^2 . Cette dépendance est illustrée sur un plan factoriel où deux mots sont d'autant plus proches qu'ils ont été employés simultanément par un individu et deux individus d'autant plus proches qu'ils ont utilisé les mêmes mots. On peut éventuellement utiliser cette méthode sur un tableau lexical agrégé – si le logiciel le permet – selon une variable d'intérêt, par exemple l'âge. Deux mots seront alors d'autant plus proches qu'ils ont été utilisés par les mêmes tranches d'âge et deux tranches d'âge se ressemblent d'autant plus qu'elles utilisent le même vocabulaire. Travailler sur un tableau

lexical agrégé améliore la qualité de représentation du plan factoriel mais cela nécessite de faire de fortes hypothèses sur le corpus.

L'inconvénient de cette méthode (Beaudouin et Lahlou, 1993) est qu'elle est peu adaptée aux réponses à une question posée dans une enquête. En effet, la question posée génère des réponses stéréotypées (notamment reprise des mots de la question) qui créent des artéfacts. Les premiers axes reposent alors sur ces réponses identiques et si l'analyse est suivie d'une classification ascendante comme cela est souvent pratiqué, les classes sont construites autour de ces artéfacts qui n'apportent rien à l'analyse des représentations. Les résultats sont en plus instables.

2.3.1.2.1 L'analyse factorielle multiple sur tableau de contingence

Cette méthode permet l'étude simultanée de plusieurs tableaux de contingence comme ceux soumis à l'analyse factorielle des correspondances. Elle permet la comparaison de plusieurs questions ouvertes lors d'une même enquête (Becue-Bertaut & Pagès, 2000) ou encore de mesurer l'impact des traitements (correction orthographique, lemmatisation) apportés au texte (Bécue-Bertaut, 2003). Dans ce type d'analyse, l'objet statistique étudié est un individu dont le profil lexical peut s'enrichir d'autres variables, quantitatives ou qualitatives.

Cette méthode a les mêmes inconvénients que la méthode précédente.

2.3.1.2.2 La classification descendante hiérarchique au sens de Reinert

Cette méthode est très utilisée en sciences humaines et sociales (entre autres). Le corpus découpé en segments de textes - les unités de contexte - permet de construire un tableau à double entrée qui croise les segments en ligne et les mots pleins (exclusion des mots outils et des mots d'une fréquence inférieure à un seuil défini) en colonne. Dans un premier temps, les unités de contexte forment une première classe. Elles sont ensuite séparées par classification descendante hiérarchique sur la base des cooccurrences des formes qui les composent. Le critère maximisé pour discriminer et séparer les unités de contexte est calculé comme le χ^2 du croisement des marges d'un découpage aléatoire en 2 sous-tableaux. Ce ne sont pas les unités de contexte qui se ressemblent qui sont assemblées mais les plus différentes qui sont divisées. Cette méthode permet de cibler la rupture dans le discours et évite les agrégations autour de noyaux lexicaux artéfactuels. Ensuite, un algorithme reclasse les unités de contexte en recherchant les réponses atypiques. Les éléments n'ayant peu de sens sont mis dans un groupe de non classés. Comme l'explique (Bastin, 2002) « *la ressemblance entre les formes n'entre pas en compte dans la méthode, la seule chose qui compte, c'est la rupture entre les classes, pas ce qu'il y a dans les*

classes ». Cette analyse fine des microstructures intrinsèques des énoncés permet de mettre en évidence des « mondes lexicaux » et des noyaux de sens (Lahlou, 1995) en alignement avec des modes de représentations issus d'un processus cognitif. Une analyse factorielle est ensuite réalisée sur l'appartenance aux classes. Cette analyse n'apporte pas d'information supplémentaire mais permet de visualiser les proximités des classes.

2.4 Choix d'Alceste comme logiciel adapté à l'analyse des représentations sociales

2.4.1 Clefs de choix logiciels

Pour choisir le logiciel le plus adapté à l'analyse des représentations sociales dans le cadre d'une enquête en face à face, ce sont les hypothèses formulées sur le matériau et les résultats escomptés qui doivent présider (Jenny, 1999) (Reinert, 1999) (Lung & Vendassi, 2012).

Plusieurs clefs permettent de choisir le logiciel adapté :

- le matériau à étudier : des articles de journaux, des discours, des tweets ;
- l'homogénéité : le corpus est constitué de matériau textuel de même nature, de collections de textes d'origine et de longueur différentes ;
- l'unité statistique choisie : les formes graphiques (les « mots bruts » du texte), les lemmes (leur forme réduite et racinisée), des segments de texte (des portions plus ou moins longues de texte) ;
- les objectifs de l'étude : post-codification, étude du vocabulaire, étude des représentations, classification, analyse du discours...

Les objectifs peuvent être pluriels, les méthodes à utiliser multiples. Il est à noter que l'utilisation de plusieurs logiciels, souvent conditionnée par les socles théoriques des utilisateurs, peut produire des résultats qui se complètent et s'additionnent dès lors que leur utilisation se fait de façon conjointe (Brugidou, et al., 2000).

2.4.2 Les logiciels adaptés pour traiter ce type de corpus

Les programmes originellement construits pour des corpus de grande taille, qui étudient le style ou la distance entre les textes, ne seront donc pas retenus.

On pourrait retenir des logiciels comme DTM-Vic car particulièrement adapté aux questions d'enquêtes (Lung & Vendassi, 2012), mais il n'y a pas de développements spécifiques adaptés à la spécificité de l'objet de la présente analyse: la représentation sociale. En effet, pour approcher au

mieux le sens des représentations (Reinert, 2000) propose une décomposition en 3 modes formant ainsi « la tresse du sens »:

- Le sens comme sensation, imaginaire ;
- Le sens comme direction, ancré dans le réel ;
- Le sens comme signification, symbolique.

Il nous invite à ne pas considérer uniquement sur le troisième mode pour étudier le sens et propose la modélisation du discours comme la trace d'une activité temporelle :

« [...] le modèle que nous proposons est donc très simple : on cherche à étudier la distribution des mots pleins dans les différents moments d'un discours pour voir en quoi l'activité discursive se stabilise. » (Reinert, 2001)

Il suggère ainsi une méthode originale, la méthode dite ALCESTE ® qui puisse tenir compte de la scansion d'un locuteur et approcher plus finement le sens. Cette méthode est implémentée dans le logiciel éponyme et également dans le logiciel libre IRaMuTeQ.

Depuis les années 80, les logiciels « historiques » se sont diversifiés en proposant de nouveaux développements, de nouveaux logiciels ont vu le jour et s'inscrivent dans des modèles de Recherche et Développement collaboratif et non propriétaires. Pourtant, Alceste® est encore un logiciel de premier choix pour l'analyse des représentations sociales. Son équivalent open source IraMuTeQ pourrait bien voir son développement s'accélérer du fait du format libre choisi par son concepteur et de la communauté d'utilisateurs qui gravite autour, mais l'interface utilisateur d'Alceste reste plus conviviale et ergonomique.

3 LES REPRESENTATIONS SOCIALES DU « BIEN MANGER » EN 2013

3.1 Matériel et méthodes

3.1.1 *Éléments d'analyse*

La croyance que « *l'on est ce que l'on mange* » (Fischler, 1990) pose la question de la confiance de l'homme en sa nourriture. Cette croyance, de l'ordre de la pensée magique (Fischler, 1994), soumet la rationalité supposée à des détours passant par la métaphore, les analogies, les approximations dont témoigne l'expression des représentations. Toutefois, les comportements paraissant irrationnels peuvent s'appuyer sur des représentations parfaitement rationnelles. Lahlou, se référant à Rozin, rappelait que la catégorisation entre objets « bons à manger » et « autres objets » résultait pour une part importante d'une transmission culturelle sans qu'il soit nécessaire d'avoir fait des expériences en la matière (Lahlou, 1998). Cette catégorisation, comme « forme d'actualisation d'une représentation », n'est donc pas tributaire des comportements réels. Elle adopte dans tous les cas un mode opératoire, au sens où, par exemple, un aliment est catégorisé comme bon ou mauvais afin de déterminer son achat (Lahlou, 1998). Les représentations servent à reconnaître ce qui est mangeable de ce qui ne l'est pas, même si encore une fois toutes les représentations ne trouvent pas de prolongement logique dans les actes des individus.

Lahlou (1998) rappelle cependant que s'il est impossible d'observer une représentation sociale, on peut en revanche observer ses conséquences sur les individus. Ainsi, on peut déduire d'un type de comportement alimentaire qu'il repose sur telle ou telle représentation, même si celle-ci peut se trouver fragilisée par des comportements sujets à la contrainte (de prix, de temps, etc.). De la même façon, on peut analyser un corpus de mots utilisés par une personne sur la représentation qu'elle se fait d'un objet, par exemple de l'idée qu'elle se fait d'une « bonne alimentation ».

Ce sont donc les représentations sociales comme manières de saisir le monde de la part d'un individu situé à un carrefour de différents milieux, carrefour faisant système, qui intéressent l'analyse lexicale. A travers le langage, un individu donne à voir l'idée qu'il se fait de quelque chose. A la fois évocation d'expériences personnelles, et expression d'un savoir commun, ce qu'exprime l'individu n'est pas représentatif mais constitue un reflet de ses caractéristiques d'appartenance (sexe, âge, profession, zone d'habitation, etc.).

3.1.2 *Recueil de données*

La question « Pour vous, qu'est-ce que bien manger ? » est la première question posée en face à face par des enquêteurs à tous les individus d'un ménage répondant au volet « Comportements et attitudes » de l'enquête CCAF (Comportements et Consommations Alimentaires des Français) réalisée par le CREDOC. Les enquêteurs ont reçu une formation : ils ont pour consigne de relancer si nécessaire et de noter textuellement la réponse fournie. Sur la dernière enquête, le terrain s'est étendu sur dix mois d'octobre 2012 à juillet 2013 et s'est effectué en quatre vagues totalisant chacune un quart de l'échantillon.

La méthode d'échantillonnage retenue pour l'enquête CCAF est la méthode des quotas avec les variables de contrôle suivantes : âge, catégorie socio-professionnelle, région, taille d'agglomération et nombre d'enfants dans le foyer. On suppose que l'ensemble des représentations sociales que l'on souhaite observer dépend uniquement de ces variables de calage. 1 200 ménages français sont interrogés : tous les individus du ménage de 3 ans et plus répondent au questionnaire. Un sur-échantillon national de 850 individus de 3 à 19 ans vient compléter la base de données soit 2 300 adultes de 15 ans et plus et 1000 enfants de 3 à 14 ans.

Les informations sur les comportements alimentaires, des ménages et des individus, sont recueillies en face à face, au domicile des personnes interrogées. Des questions relatives aux comportements alimentaires des ménages français (circuits et stratégies d'achat, choix des produits, préparation, composition et prise des repas...) sont posées. Le questionnaire décrit également les comportements alimentaires des individus, enfants (de 9 à 14 ans) et adultes (15 ans et plus), vis-à-vis de leurs opinions relatives à la santé et à la nutrition, les critères de qualité, la perception des risques et des OGM, le snacking, les consommations à domicile et hors foyer, les produits bio, etc.

Initié pour la première fois en 1988, le volet comportements de CCAF 2013 est la 8ème vague de cette enquête (1988, 1995, 1997, 2000, 2003, 2007, 2010). Le CREDOC étudie donc depuis 25 ans les comportements alimentaires des Français et propose une analyse des changements de comportements et des tendances qui s'annoncent ou se confirment.

En analyse lexicale, la constitution du corpus tout comme le choix logiciel n'est pas neutre. Ici, le corpus est bien constitué pour répondre à la question de recherche posée à savoir : appréhender les représentations sociales du bien manger dans la société française actuelle.

3.1.3 Pré-traitements du corpus

Le corpus est soumis à des prétraitements destinés à améliorer la stabilité des résultats. Ils consistent à :

- vérifier et corriger l'orthographe le cas échéant ;
- corriger les fautes de frappes ;
- développer les abréviations ;
- harmoniser les orthographes ;
- identifier et lier les noms composés (Beaudouin, et al., 1993).

Cette étape, dont il semble difficile de s'affranchir pour garantir la qualité des données, impose la lecture intégrale du corpus et peut s'avérer chronophage. Pour contourner cet inconvénient, une solution consisterait à appliquer à l'aveugle une correction orthographique grâce à des développements issus des sciences du TAL sur des corpus thématiques d'apprentissage (Suignard, et al., 2014) mais ces processus sont encore trop coûteux. Une autre option consisterait à composer avec le texte brut, sans corrections, et utiliser *in fine* une unité statistique différente. En effet, (Bécue-Bertaut, 2003) a montré sur un corpus de réponses ouvertes que « *la lemmatisation modifie davantage la structure que la normalisation* (orthographique) ». Elle explique, au moyen de l'analyse factorielle multiple sur tableaux de contingence, que l'information véhiculée dans un corpus brut, corrigé ou lemmatisé est différente et que la structure est moins modifiée en passant du texte brut au texte normalisé que du texte normalisé au texte lemmatisé. Au lieu de choisir une unité statistique plutôt qu'une autre, elle propose d'ailleurs de les utiliser simultanément. Cette piste n'a pas été explorée ici, car si l'idée est séduisante, on peut d'ores et déjà émettre la réserve que cette solution est très dépendante du corpus et des opérateurs de saisie. L'avantage d'une lecture intégrale, même contraignante, est d'acquérir une bonne connaissance du corpus et de soulever des questionnements qu'on souhaitera explorer une fois le corpus prêt à être analysé.

3.1.4 Analyses statistiques descriptives

C'est le package `tm()` de R qui est utilisé pour réaliser les statistiques descriptives du corpus. Tout d'abord, le nombre de formes graphiques du corpus, le nombre de formes distinctes ainsi que le nombre d'hapax (mots qui ne sont cités qu'une seule fois) sont calculés et utilisés pour caractériser le corpus. La courbe de Zipf (ou courbe des gammes de fréquences) est construite comme la distribution des occurrences en fonction du rang des formes graphiques sur une double échelle logarithmique avec l'objectif d'examiner de façon globale la distribution lexicale, sans considération de vocabulaire (Beaudouin & Hébel, 1994). La manière dont cette courbe s'écarte d'une droite renseigne sur la distribution des formes graphiques : plus les fréquences sont instables, plus la lemmatisation impactera le nombre de variables utilisées pour réaliser la classification (Beaudouin

& Hébel, 1994). Ensuite, le tableau lexical est agrégé selon des variables d'intérêt (âge, sexe, catégorie socio-professionnelle et IMC) et chaque forme graphique est testée pour mettre en évidence un sur ou sous emploi selon la variable d'agrégation. Si le test est significatif pour un mot, sa valeur test positive est conservée et un nuage de mot est généré avec le package wordcloud() de R pour l'ensemble des formes significatives. Les nuages de mots construits permettent d'avoir un aperçu rapide des différences de lexique qui peuvent exister au sein des catégories d'intérêt.

3.1.5 Classification et caractérisation des classes

Au vu des résultats exploratoires, que propose chaque logiciel, détaillés au paragraphe 1.3, on utilise la classification au sens de Reinert programmée dans le logiciel Alceste© pour traiter la question du bien manger en 2013 et révéler les classes de discours. L'outil intégré de lemmatisation est utilisé pour cette analyse car la lemmatisation a fait ses preuves sur la stabilité des classes et permet d'éviter la formation de classes de réponses artéfactuelles (Brousseau, 1999). La classification est double pour bénéficier d'une plus grande stabilité des classes. La longueur des unités de contexte paramétrée est celle par défaut (découpage entre 10 et 12 mots). Les clés d'analyse ont été modifiées pour tenir compte des mots outils - marqueurs d'une relation temporelle (ex :) ou d'une intensité (ex :) - ceux-ci étant susceptibles d'intervenir dans la discrimination des classes. Les classes obtenues sont caractérisées par les formes graphiques significativement plus représentées dans les classes au sens du χ^2 d'association. L'analyse conjointe des formes les plus caractéristiques et des fréquences de segments répétés permet de mieux interpréter la classification réalisée.

Tous les individus sont ensuite affectés à une classe de discours. Les classes sont décrites par des variables socio-démographiques (exogènes à la classification) au moyen de la fonction catdes() du package FactoMineR de R.

Pour chacune des classes, on calcule enfin un bilan lexical (nombre de mots, nombre de formes distinctes, nombre de hapax et longueur moyenne des réponses) afin de détecter des variations autour de la richesse du vocabulaire utilisé par chaque classe Alceste. Ceci est réalisé avec le package tm de R.

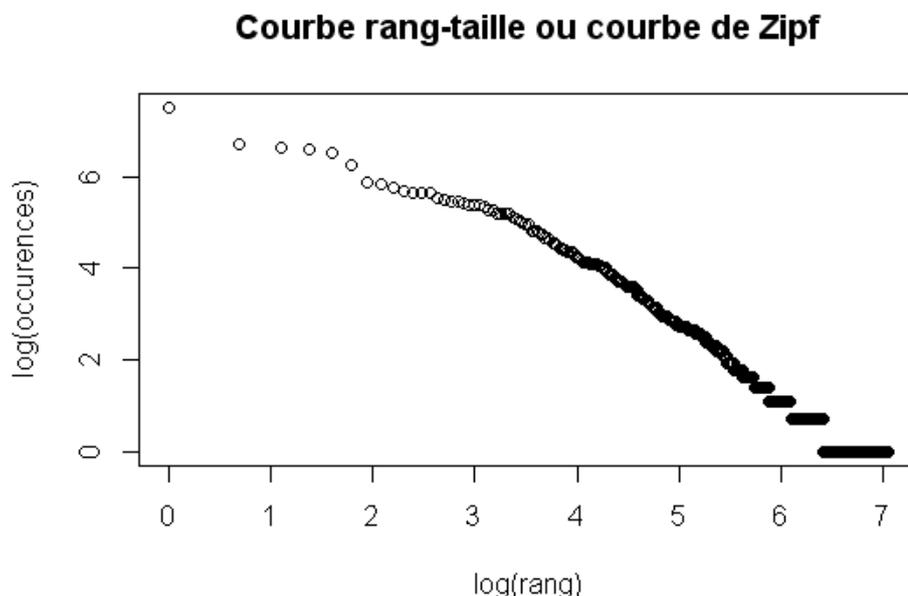
3.2 Résultats

3.2.1 Description du corpus

Le corpus contient un total de 19 571 formes graphiques parmi lesquelles on distingue 1 144 mots distincts (soit 5,8% du total). Le vocabulaire est donc particulièrement redondant. Parmi ces formes distinctes, 533 sont des hapax (soit 2,7% du total des formes et 46% des formes distinctes).

La courbe de Zipf (Figure 4) est particulièrement instable sur les hautes et basses fréquences. L'allure de la courbe est typique des corpus de réponses à des questions ouvertes. L'instabilité sur les hautes fréquences s'explique par le phénomène d'écholalie décrit au paragraphe **Erreur ! source du renvoi introuvable.** et l'instabilité sur les basses fréquences s'explique par la multiplicité des « auteurs » du corpus. En effet, chaque auteur apporte un vocabulaire particulier, la courbe s'écarte donc de la droite de régression de façon plus prononcée qu'en cas d'étude d'un corpus par un auteur unique.

Figure 4 - Courbe de Zipf du corpus "bien manger" 2013



Source : Enquête CCAF 2013, CREDOC

3.2.2 Catégorisation du vocabulaire selon le sexe, l'âge, la CSP et l'IMC

Les nuages de mots sont une grille de lecture intuitive et ludique pour représenter le vocabulaire typique d'une catégorie. Chaque forme graphique est typographiée dans une police d'autant plus

grande que sa valeur test est significative. Pour guider la lecture, 8 tranches de couleurs indiquent un même ordre de grandeur de valeur test.

a. Les formes graphiques caractérisant le sexe

- Les mots caractéristiques des femmes appartiennent plutôt au champ lexical de la nutrition et de la diététique (Figure 5). Elles citent de façon privilégiée des aliments ou nutriments (« légumes », « protéines », « laitiers », « sel ») ainsi que des formes graphiques relatives à la manière de les préparer (« sainement », « cuisiner ») ou de les ingérer (« moment », « plusieurs » du syntagme « plusieurs fois », « mangé »). Cette relation à l'alimentation, s'explique de façon anthropologique (Hubert, 2006). La mère, et ce dans toutes les cultures, sécurise son enfant en le nourrissant et notamment dès la naissance en l'allaitant. La mère garde ce rôle de prévention en santé au sein de la famille. Les femmes sont alors nettement plus sensibles aux liens entre alimentation et santé.

Quand elles parlent d'elles, les femmes utilisent le pronom personnel « moi ». On retrouve ici l'image du corps particulière à la femme qui craint de prendre du poids, concomitante à celle de suivre un régime. Les femmes s'expriment davantage en termes de besoins, d'attentes et d'apports susceptibles de répondre aux besoins du corps. Les représentations font apparaître les « bons » aliments en termes de bienfaits pour le corps. Cela explique une forte réceptivité des femmes vis-à-vis des recommandations nutritionnelles des pouvoirs publics et des industriels. On sait qu'un certain nombre de sites sur Internet consacrés à l'alimentation et à la nutrition sont expressément dédiés à un public féminin. Pour elles, une bonne alimentation répond à la prévention santé dans son rôle nourricier et correspond à un impératif social qui privilégie l'apparence, la présentation de soi. Chez les femmes, comme le montre Hubert (2006), les représentations du corps, les conceptions de la santé et les représentations ancestrales se télescopent et mettent en évidence un mélange confus sur l'alimentation.

- On retrouve dans les formes spécifiques des hommes le rôle nourricier de la mère et de la femme des anthropologues (Hubert, 2006). Dans les cultures patriarcales la femme ne peut être que mère et nourricière. Ces représentations anciennes témoignent d'une approche attachée à la satiété, au plaisir du ventre, au goût, et au plaisir du repas partagé (« amis »). Le champ lexical de la satiété est présent (« sa », « faim », « ventre »), le goût (« bon »), la convivialité (« amis »). L'adjectif pluriel « petits » provient du syntagme « petits plats ». La forme « femme » est présente ici car les hommes citent de façon privilégiée leur femme (et plus particulièrement leur cuisine) comme edificatrice de leur représentation du bien manger. Quand les hommes parlent d'eux, ils utilisent le pronom personnel éliidé « j' ». Le domaine culinaire reste au quotidien une attribution majoritairement féminine au sein des couples, ce qui explique la place du mot « femme » comme référence privilégiée des hommes lorsqu'il

s'agit de faire appel à leurs représentations. Le stéréotype de la reproduction de la cuisine maternelle comme définition de « *la seule vraie cuisine* » perdure. La femme est intimement associée dans l'imaginaire à la production culinaire, c'est à dire à la transformation des aliments en nourriture.

Le vin, qui n'apparaît pas dans le nuage « féminin », est le seul aliment cité et occupe une place notable chez les hommes.

Figure 5 - Nuage de mots des formes graphiques caractérisant le sexe



Source : Enquête CCAF 2013, CREDOC

b. Les formes graphiques caractérisant la catégorie d'âge

Le vocabulaire différencié par catégorie d'âge illustre des préoccupations différentes à travers les générations. Les 15-24 ans semblent plus imprégnés par le discours nutritionnel issu du PNNS (Programme National Nutrition Santé), les formes « *grignoter* », « *entre* », « *gras* », « *cinq* », « *jour* », « *fruits* » et « *équilibré* » émanent des messages sanitaires « *Pour votre santé, évitez de grignoter entre les repas* », « *Pour votre santé, mangez au moins cinq fruits et légumes par jour* », « *Pour votre santé, évitez de manger trop gras, trop sucré, trop salé* ». Leur aliment préféré semble être les « *frites* ». Les plus jeunes apparaissent comme davantage perméables aux recommandations parce qu'ils se situent dans une période du cycle de vie où ils tendent à s'émanciper des règles parentales sans que leur mode de vie ait eu le temps et l'occasion de se structurer. Cette incertitude les conduit donc à être plus réceptifs à ce type de message. On retrouve chez les plus jeunes, la place de la « *mère* » nourricière (Hubert, 2006). « *Pour les Juifs, les Chrétiens, les Musulmans, aimer sa mère veut dire qu'il faut manger ce qu'elle nous donne. La nourriture offerte est une preuve de son amour. Refuser la nourriture, c'est refuser l'amour. Il y a une charge affective énorme et une grande émotion exprimées dans l'offre à manger maternelle. On note, en anthropologie de l'alimentation, des différences entre la manière dont sont nourris*

filles et garçons, ou les préférés et les autres, au sein d'une même cellule familiale. Les garçons sont souvent favorisés, souvent plus gros, il est des cas où on assiste littéralement à des gavages maternels : *si tu m'aimes, mange !* ».

Les 25 à 34 ans utilisent de façon privilégiée la forme plurielle « *viandes* », généralement de façon contiguë aux formes « *légumes* », « *poissons* », « *féculents* » au pluriel également. Ils sont également caractérisés par la forme « *après* » du syntagme « *ne plus avoir faim après* », traduisant la recherche de l'effet de satiété. L'importance de la convivialité se manifeste par les formes « *famille* » et « *partager* », souvent cooccurrentes, et rappelle que, tandis que l'adolescence inaugure une période plus déstructurée dans l'équilibre et les manières alimentaires, la mise en couple et l'arrivée du premier enfant se caractérise généralement par un « *retour aux normes* » (Desjeux, 2006). Cette classe d'âge cite de façon privilégiée la forme « *jardin* » pour désigner la provenance des légumes dans la représentation qu'ils se font du « bien manger ». Ils évoquent la restriction à travers « *raisonnables* » du syntagme « *quantités raisonnables* ».

Les 35 à 44 ans s'approprient le vocabulaire des plats et des produits bien « *cuisinés* ». Ils se partagent avec les 25 à 34 ans la forme « *famille* ». La restriction n'est évoquée que par la forme « *priver* » mais le retour au texte montre qu'il s'agit plutôt de ne pas se priver. C'est la seule classe à évoquer significativement le prix dans ses représentations du « bien manger ». Ces tranches d'âge correspondent à la période du cycle de vie où l'on élève ses enfants. Il s'agit donc de les nourrir correctement, suffisamment, en respectant un certain nombre de règles, sans cesser de prêter attention à son budget en raison du poids des dépenses consacrées à l'entretien de la cellule familiale. Le terme « *famille* » indique l'importance donnée aux repas pris en commun, matérialisant le lien entre parents et enfants.

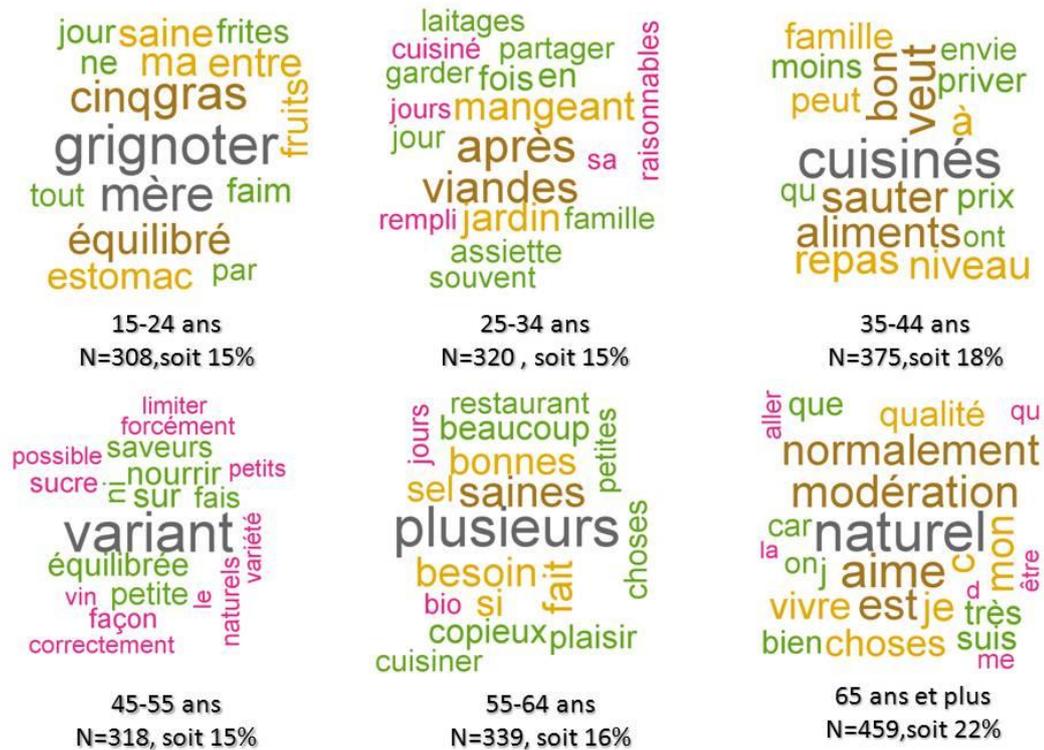
Les 45 à 55 ans ont un vocabulaire plus partagé et peu de mots sont très significatifs de cette classe. La forme « *variant* » du syntagme « *variant les repas* » vient s'ajouter au sur-emploi de la forme « *équilibrée* » du syntagme « *alimentation équilibrée* ». A cet âge, il s'agit de garder la forme en faisant attention à son alimentation, sans être pour autant dans une perspective directe de santé.

La catégorie des 55-64 ans ne se restreint pas. Au contraire, elle utilise de façon privilégiée des formes relatives au plaisir et à la quantité. L'importance de la convivialité se manifeste par le sur-emploi de « *plusieurs* » dans le syntagme « *à plusieurs* », que l'on peut mettre en lien avec celui de « *restaurant* ». Elle se préoccupe également de la présence de « *sel* » dans les aliments et semble particulièrement concernée par le « *bio* ». On y décèle les prémisses d'une préoccupation concernant sa santé directement impactée par l'âge. On peut y déceler également un effet

générationnel, les préoccupations relatives à l'environnement, auxquelles participent la filière « bio », étant davantage portées par les actifs les plus âgés.

Les 65 ans et plus semblent se tourner vers plus de simplicité : ils sur-emploient la forme « naturel » au singulier dans les syntagmes « manger naturel » ou « le plus naturel possible ». Les personnes les plus âgées ont sans doute mieux connu l'univers des campagnes, or les études alimentaires ont montré le lien entre familiarité au monde rural et attrait pour les produits naturels. La restriction est incarnée dans les formes « modération » et « normalement », à un âge où l'appétit tend à diminuer et où la digestion se fait parfois plus paresseuse. Cette classe est la seule à sur-employer des pronoms personnels pour parler de ses représentations (« je », « me », « mon »).

Figure 6 - Nuage de mots des formes graphiques caractérisant la catégorie d'âge



Source : Enquête CCAF 2013, CREDOC

c. Les formes graphiques caractérisant la catégorie socio-professionnelle

Les agriculteurs exploitants sont principalement caractérisés par deux formes du bien manger : « toujours », du syntagme « pas toujours la même chose » pour insister sur l'importance de la variété, et la forme « saveurs » pour souligner l'importance du goût. Il faut y voir sans doute l'effet d'un accès privilégié à des produits agricoles de meilleure qualité, notamment en termes de goût.

Les artisans, commerçants et chefs de PME sur-emploient les formes « *niveau* » et « *alimentaire* » dans les syntagmes « *au niveau du goût* » et « *équilibre alimentaire* ». Ils rejettent les « *plats* » tous « *prêts* ». La présence des formes graphiques « *femme* » et « *restaurant* » laisse à penser que leur représentation du bien manger est liée à une alimentation qui n'est pas préparée par leurs soins, correspondant à des modes de vie plus traditionnels avec des rôles plus définis entre les hommes et les femmes.

Les cadres et professions intellectuelles supérieures sur-emploient les formes ayant trait au champ lexical du goût (« *goût* », les syntagmes « *quelque chose de bon* » et « *bonnes choses* »). La convivialité est illustrée par la sur-représentation de la forme « *amis* ». Cette catégorie de CSP+ est la seule à évoquer le « *bio* » et le « *vin* ». On peut y déceler la place des déjeuners d'affaires dans ces catégories professionnelles, et plus largement celle des relations sociales et des réseaux de connaissances. La bonne nourriture et le bon vin y sont les garants de bons échanges et de moments de détente. Le recours au « *bio* » indique parallèlement un niveau de revenus permettant d'acquérir des produits plus onéreux.

Les cadres moyens et professions intermédiaires utilisent préférentiellement le vocabulaire d'une « *cuisine* » qualifiée de « *saine* ». Ils sur-emploient le mot « *gâteau* » comme la denrée-exemple à ne pas consommer. A noter également, l'emploi privilégié du pronom « *ses* » pour parler de « *ses envies* », « *ses repas* », « *ses menus* »...

Les contremaîtres, agents de maîtrise et techniciens comptent beaucoup de formes graphiques spécifiques mais avec une significativité moins élevée. C'est toutefois cette classe qui aborde avec son vocabulaire le plus de concepts différents. « *Moins* », « *limiter* », « *attention* », sont les formes les plus significatives et appartiennent plutôt au champ lexical de la restriction.

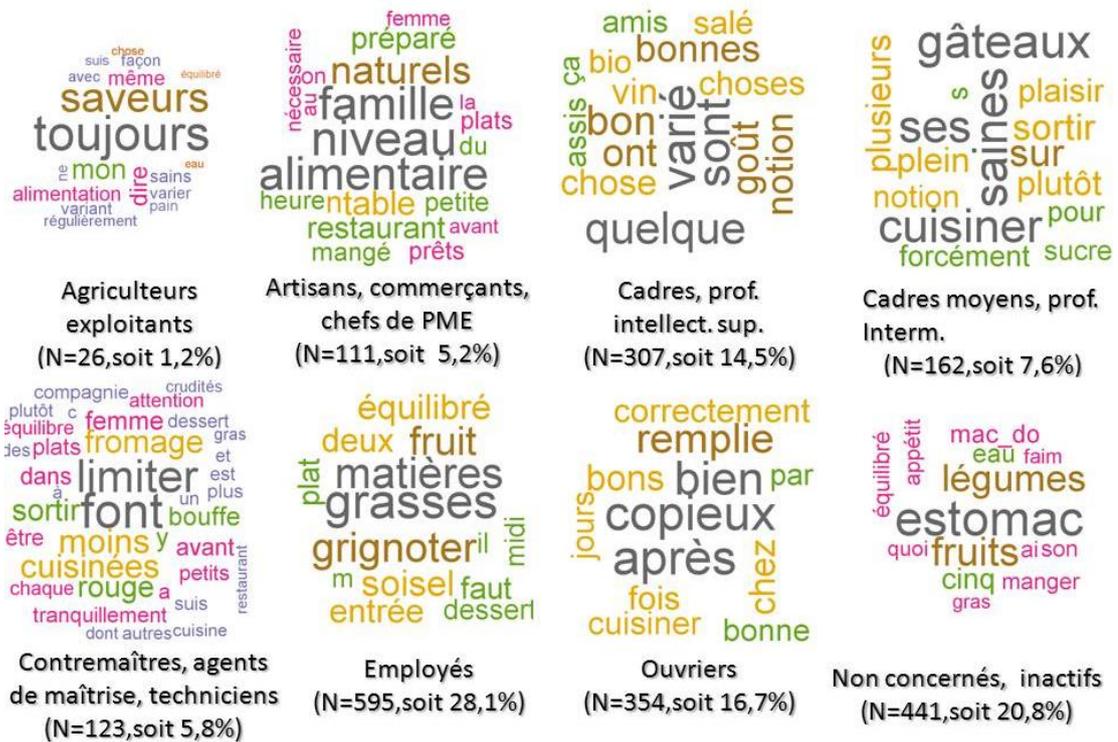
Les employés sur-emploient un vocabulaire lié à la nutrition et à la diététique : les « *matières grasses* » et « *grignoter* » sont à éviter et sur-représentés dans cette classe. De la même façon « *équilibré* » est significativement plus cité dans cette classe. Les formes « *entrée* », « *plat* » et « *dessert* » sont également plus fréquentes ensemble. Ces professions correspondent à un environnement urbain, bien exposé aux messages nutritionnels, faiblement novateurs et au contraire fidèles à des repas structurés.

Les ouvriers sur-emploient le vocabulaire de la satiété : « *copieux* », « *après* » du syntagme « *ne plus avoir faim après* », « *remplie* », « *correctement* ». A noter également, l'emploi de la préposition « *chez* » comme prescriptrice d'endroits où il faut, ou a contrario où il ne faut pas, se fournir en repas où en aliments (« *chez soi* », « *chez le producteur* », « *chez le boucher* »...). Le souci d'une alimentation suffisamment abondante renvoie aux besoins physiques de professions où

la dépense du corps reste importante et nécessite de l'énergie. Les différences selon les classes sociales observées dans les travaux de Bourdieu restent toujours vraies.

Les non concernés et les inactifs semblent, comme les 15-24 ans, plutôt imprégnés du discours nutritionnel et évoquent les modes de consommation en fast-food (« Mac Do »).

Figure 7 - Nuage de mot des formes graphiques caractérisant la catégorie socio-professionnelle



Source : Enquête CCAF 2013, CREDOC

d. Les formes graphiques caractérisant l'IMC

Les individus dont l'IMC indique un statut pondéral « maigre » utilisent davantage le vocabulaire lié à la quantité, via les formes « ai » de « j'ai », « ait » du syntagme « qu'il y ait », « assez », « assiette » et « rempli ».

Ensuite, les individus dont l'IMC indique un statut pondéral normal utilisent de façon privilégiée le vocabulaire d'une alimentation équilibrée (« fruits », « légumes », « variant », ne pas « grignoter »).

Les individus dont l'IMC indique un surpoids pondéral utilisent préférentiellement des formes graphiques qui touchent à plusieurs dimensions : la santé (« forme »), le « prix », la « convivialité », la notion de « saison » dans les « produits ». Dans tous les cas, cette catégorie a sur-employé le mot « important » pour décrire ses représentations du bien manger.

Enfin les individus dont l'IMC indique une obésité ont utilisé préférentiellement la forme « table », soit pour décrire la satiété en « *sortant de table* », soit pour souligner l'importance de manger « *assis à table* ». La satiété est citée via les mots « *satiété* », « *ventre* » et « *plein* ». Le mot « *plaisir* » est enfin caractéristique de cette catégorie.

Globalement, les individus dont l'IMC correspond au moins à un surpoids n'expriment pas un rapport problématique à la nourriture, ou un sentiment de culpabilité. On y observe au contraire, notamment, la notion de « *plaisir* », certes ainsi que d'« *équilibre* », et de « *convivialité* ».

Figure 8 - Nuage de mot des formes graphiques caractérisant l'IMC



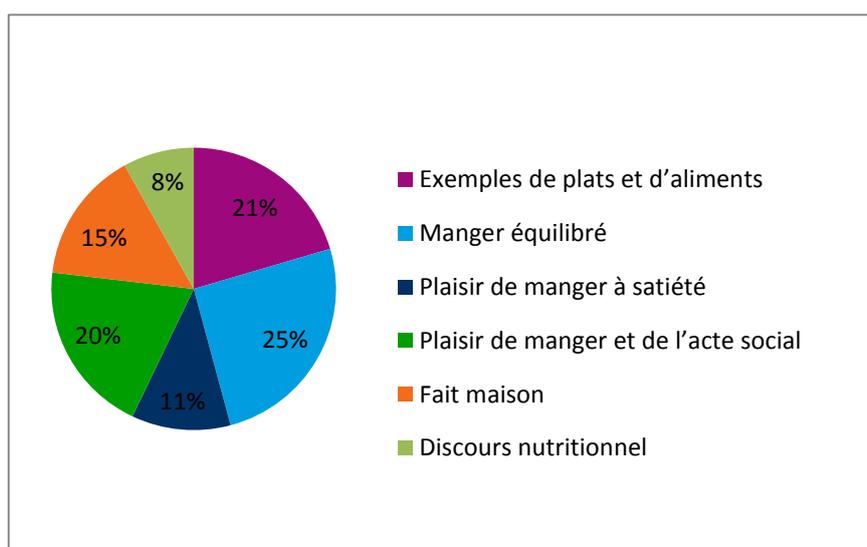
Source : Enquête CCAF 2013, CREDOC

3.2.3 Six classes de représentations du bien manger

La classification descendante hiérarchique de Reinert, décrite au paragraphe 2.1.4., produite par le logiciel Alceste a classé 2157 unités de contexte produites par 2120 individus en 6 classes de discours (18% d'unités rejetées non classées). 98% des réponses des individus sont donc analysées comme une unité de contexte, les réponses de 34 individus comptabilisent 2 unités de contexte et la réponse d'un individu en comptabilise 4 (Figure 9). Pour la moitié des réponses

comptabilisant plusieurs unités de contexte, leurs unités de contexte appartiennent à la même classe. Dans l'autre moitié des cas, il n'y a qu'une unité classée et l'autre est rejetée. On pourra donc faire l'approximation qu'à un individu correspond une seule représentation du bien manger. Les six représentations du « bien manger » mises en évidence par la classification sont les suivantes : manger équilibré (25%), exemples de plats et aliments (21%), acte social (20%), fait maison (15%), manger à satiété (11%), et discours nutritionnel (8%).

Figure 9 - Répartition des classes lexicales du corpus "bien manger" en 2013 :



Source : Enquête CCAF 2013, CREDOC

Parmi les 1 181 formes distinctes détectées, 238 ont été activement utilisées dans l'analyse. Le bilan lexical par classe (Tableau 2) montre que les classes lexicales ne sont pas équivalentes du point de vue de la richesse du vocabulaire intra-classe. La classe de la représentation du manger équilibré (26%) semble particulièrement pauvre (0,5 formes distinctes par individu et 0,3 hapax par individu) et la classe du discours nutritionnel (9%) particulièrement riche (2,1 formes distinctes par individu et 1 hapax par individu), alors que les autres classes sont équivalentes.

Tableau 2 - Bilan lexical par classe du corpus "bien manger" en 2013

Classes	individus	formes graphiques	formes graphiques/individu	formes distinctes	formes distinctes/individu	hapax	hapax/individu
Exemples de plats et d'aliments	352	4795	13,6 ±8	453	1,3	216	0,6
Plaisir de manger à satiété	194	2025	10,4 ±7	257	1,3	128	0,7
Plaisir de manger et de l'acte social	341	3836	11,2 ±6	450	1,3	224	0,7
Manger équilibré	438	2373	5,4 ±4	240	0,5	123	0,3
Discours nutritionnel	139	1844	13,2 ±7	296	2,1	141	1,0
Fait maison	260	3073	11,8 ±7	423	1,6	216	0,8

Source : Enquête CCAF 2013, CREDOC

3.2.3.1 Classe 1 : Exemples de plats et d'aliments, 20%

Cette classe contient des descriptions de la structuration d'un repas, de son ordonnance, éventuellement complétées d'un type d'aliment en particulier. Si l'on est loin de la complexité de l'ordonnance d'un grand repas français comme au XIX^{ème} siècle (mise en bouche, hors-d'œuvre, potage, relevé, entrée, rôti, entremets, fromage et desserts – incluant pâtisseries, entremets sucrés et mignardises), on en perçoit les reliefs. Ces descriptions adoptent l'un des piliers du modèle alimentaire français, celui d'un repas structuré, avec un nombre de plats successifs toujours pris dans le même ordre. Ce qui implique que les convives partagent le même menu. En dépit d'une tendance à la simplification des repas depuis quelques années, cette structuration demeure relativement prégnante, en particulier concernant les dîners. La mentalisation du bien manger, qui passe par ces descriptions du type « entrée, plat, dessert », passe également par la mention des familles d'aliments : d'abord les légumes et les fruits, ensuite la viande, puis les féculents. Les mots caractéristiques de la classe ci-dessous sont cités par ordre de significativité. Cette classe est particulière parce qu'elle donne des réponses en termes d'« objets concrets ». (Lahlou, 1992) dégage également cette classe dans l'enquête « Aspirations des Français » du CREDOC de 1991 quand il pose la question « Si je vous dit manger à quoi pensez-vous ? ». Il explique qu'« il n'y a pas de raisonnement, d'affects, ni d'abstraction raisonnée. Les réponses sont essentiellement descriptives et informatives ».

Mots caractéristiques (par ordre de significativité) : légume, fruit, viande, dessert, entrée, jour, féculent, fromage, poisson, trois, plat, repas, régulier, cinq, eau, boire, laitage, complet, heure, salade.

Quelques réponses typiques de la classe:

« avoir un plat de viande et poisson un peu de légumes et des féculents, du fromage et un fruit quelques fois ».

« avoir un repas matin, midi et soir. Une entrée, un plat, un dessert »

« avoir un repas de légumes, de féculents et de fruits trois fois par jour et des produits laitiers »

« c'est un bon repas composé d'une entrée, un plat copieux, une viande bien tendre, fromage, salade, et un dessert »

« un repas complet entrée plat fromage dessert »

Les caractéristiques des enquêtés ayant construit cette représentation du bien manger sont : sur-représentation significative des jeunes de 15 à 24 ans (20% dans la classe contre 14% dans la population), des femmes (59% dans la classe contre 52% dans la population), des habitants de

l'agglomération parisienne (26% dans la classe contre 17% dans la population), des employés (22% dans la classe contre 14% dans la population), , des individus avec un IMC de normo pondéraux (59% dans la classe contre 51% dans la population), des individus dont les revenus du foyer sont compris entre 18 294 € et 22 867 € (13% dans la classe contre 9% dans la population), ainsi que des individus dont le diplôme du chef de famille est un équivalent Bac +2 (BTS, DUT, DEST ou DEUG) (14% dans la classe contre 11% dans la population).

3.2.3.2 Classe 2 : Plaisir de manger à satiété, 11%

La satiété est la première dimension dans cette classe, il s'agit avant tout de manger « à sa faim », ni trop, ni trop peu. Avoir conservé les mots outils de la relation temporelle se justifie pleinement ici car ils permettent de dégager des mots caractéristiques comme « plus » et « après » fréquents dans les locutions « ne plus avoir faim après », plus particulièrement « en sortant de table ».

La notion de plaisir est assez importante dans la sensation de réplétion : il faut également que l'ingestion soit source de satisfaction. Cela passe par le contentement des envies (« ce que j'aime », « ce que j'ai envie ») et l'absence de privation (« sans se priver »). Les difficultés consécutives à la crise économique contribuent sans doute à rappeler l'importance qu'il y a à bien s'alimenter, à y trouver du plaisir. Cette dimension avait faibli en 2007 du fait de l'importance donnée cette année-là à la santé et à l'équilibre nutritionnel. Nous voyons en 2013 que les plaisirs de la table reprennent de l'ampleur.

Mots caractéristiques (par ordre de significativité) : Faim, plus, plaire, après, envie, table, rassasier, assez, ne, sortir, appétit, aimer, priver, satisfaire, souvent.

Quelques réponses typiques de la classe :

« sortir de table ne plus avoir faim »

« c'est plus avoir faim après et manger ce qu'on aime.

« satisfaire sa faim et ses envies »

« ne plus avoir faim, avoir mangé assez. Pour moi il y a une notion de quantité »

« ne plus avoir faim. Manger des plats qui me plaisent, que ce soit bon »

« manger à ma faim et ce qui me plaît »

« manger de tout, sans se priver, mais être raisonnable »

« manger des choses agréables au palais. Qu'on n'ait pas faim en sortant de table qu'on ait assez mangé sans manger trop »

Les caractéristiques des enquêtés ayant construit cette représentation du bien manger sont : sur-représentation significative des hommes (58% dans la classe contre 47% dans la population), des jeunes de 15 à 24 ans (22% dans la classe contre 14% dans la population) et des individus habitant des foyers de 5 personnes ou plus (9% dans la classe contre 6% dans la population).

3.2.3.3 Classe 3 : Plaisir de manger et de l'acte social, 19%

Dans cette classe, la dimension hédonique du bien manger prédomine. Il s'agit d'abord et avant tout de « se faire plaisir » (répété 10 fois), de « prendre du plaisir » (répété 48 fois). Ce plaisir à manger est illustré par l'intérêt porté à manger des bonnes choses manger des choses qui ont du goût, par une « bonne assiette » (6 répétitions) avec laquelle on va « se régaler ».

Le plaisir passe aussi par le temps consacré à manger, c'est-à-dire que bien manger est l'occasion de se réunir, notamment avec des amis ou en bonne compagnie, pour partager un moment plutôt long. Il s'agit de « prendre le temps » (15 répétitions) pour vivre l'instant présent et apprécier la convivialité autour du repas. C'est la seule classe où il est fait référence aux autres participants du repas et à l'importance du lien social.

Là encore, en période de crise, l'accent est mis sur la gourmandise, comme pour compenser les difficultés du quotidien. La gastronomie, le plaisir de cuisiner, la variété des aliments, la valorisation du goût et des « manières de table », la convivialité, témoignent tous d'une dimension culturelle revivifiée qui participe au réenchantement du quotidien. Ces éléments progressent au détriment des préoccupations relatives à la santé par l'alimentation qui avaient fleuri en 2007.

Cette dimension « santé » reste néanmoins présente : bien manger, c'est être, avoir ou maintenir une « bonne santé » (15 répétitions). C'est aussi être, rester et se maintenir « en forme » (10 répétitions).

Mots caractéristiques (par ordre de significativité) : plaisir, bonne, chose, faire, santé, prendre, assiette, goût, apprécier, convivial, vivre, agréable, forme, ami, temps, régaler, bien, compagnie, moment, restaurant.

Quelques réponses typiques de la classe :

« du plaisir, c'est bien se nourrir et être en bonne santé »

« se faire plaisir. Prendre le temps de manger et d'apprécier ce qu'on mange, que cela ait du goût »

« prendre du plaisir. Se régaler. La convivialité »

« me faire plaisir, avec la famille ou les amis »

« c'est prendre du plaisir. Que ce soit un moment de plaisir et de partage »

« être en bonne santé et se faire plaisir »

« manger des bonnes choses. Etre bien accompagné quelque chose de convivial quoi »

« c' est prendre son temps. Se faire un repas mijoté, préparé. Il faut une attente du repas.
La compagnie pendant les repas »

Les caractéristiques des enquêtés ayant construit cette représentation du bien manger sont : sur-représentation significative des individus de 55 à 64 ans (19% dans la classe contre 16% dans la population), des couples sans enfants (39% dans la classe contre 31% dans la population), des individus passant plus de 3h par jour devant un écran (67% dans la classe contre 60% dans la population), des foyers dont les revenus annuels sont compris entre 30 490 € et 38 112 € (13% dans la classe contre 9% dans la population), des inactifs (41% dans la classe contre 32% dans la population), des individus en surpoids (34% dans la classe contre 29% dans la population), et des individus qui déclarent ne pas savoir s'ils ont une bonne alimentation (1,3 % dans la classe contre 0,5 % dans la population).

3.2.3.4 Classe 4 : Manger équilibré, 26%

Cette classe, la plus volumineuse, est aussi la plus homogène. Les réponses sont particulièrement courtes (5 mots en moyenne) et le sous-corpus constitué par le vocabulaire de la classe particulièrement peu riche et varié. Les représentations du bien manger sont principalement axées sur l'équilibre (52% des formes « équilibré » sont comptées dans cette classe) et se déclinent en plusieurs idées autour de ce concept. Manger équilibré, c'est « varier les menus » (12 répétitions), c'est « manger sain » (34 répétitions). La forme « goûteux », même si sa fréquence est faible (12 occurrences) sur l'ensemble du corpus, est particulièrement représentée dans cette classe, elle est même systématiquement associée à la notion d'équilibre. Manger équilibré, c'est également manger « correctement » et dans des quantités « raisonnables ». La restriction n'est pas abordée aussi frontalement que dans la dernière classe évoquant le discours nutritionnel, mais le souhait de contrôler quantité et qualité est nettement perceptible.

Mots caractéristiques (par ordre de significativité) : Equilibré, manger, sain, varier, menu, goûteux, correct, raisonner.

Quelques réponses typiques :

« c'est manger sainement. C'est varier les menus ».

« manger équilibré et de façon saine. »

« manger équilibré et des menus variés. »

« manger correctement, équilibré »

Les caractéristiques des enquêtés ayant construit cette représentation sont : sur-représentation significative des femmes (58 % dans la classe contre 52 % dans la population), des 35-44 ans (21 % dans la classe contre 17 % dans la population), des individus dont le chef de famille est cadre, de profession intellectuelle supérieure ou de profession libérale (26 % dans la classe contre 19 % dans la population), dont le diplôme du chef de famille est de niveau 2ème ou 3ème cycle universitaire ou grande école (23% dans la classe contre 16 % dans la population), des individus qui comptent 2 actifs ou plus dans le foyer (76 % dans la classe contre 71 % dans la population), les individus les moins sédentaires qui passent moins de 3h par jour devant un écran (44 % dans la classe contre 39 % dans la population) et des individus qui déclarent avoir une bonne alimentation (76 % dans la classe contre 71 % dans la population). .

3.2.3.5 Classe 5 : discours nutritionnel, 9%

Dans cette classe, la plus petite de la typologie, le discours nutritionnel dispensé par les dispositifs gouvernementaux depuis que l'état s'est saisi de la question alimentaire est très pregnant.

En effet, depuis le 28 février 2007, la loi française impose aux marques de produits alimentaires (sauf à s'acquitter d'une taxe) d'introduire les messages sanitaires suivants dans leurs publicités et autres outils de communication :

- « Pour votre santé, mangez au moins cinq fruits et légumes par jour » ;
- « Pour votre santé, pratiquez une activité physique régulière » ;
- « Pour votre santé, évitez de manger trop gras, trop sucré, trop salé » ;
- « Pour votre santé, évitez de grignoter entre les repas ». (PNNS - Messages sanitaires sur les publicités, 2014)

Ces discours nutritionnels diffusés à la télévision et à la radio ont donc imprégné les esprits au point d'être repris presque mot à mot dans cette classe. 60% des formes « trop », « gras », « sucré » et « salé » sont rassemblées dans cette classe et les segments répétés « trop gras », « trop sucré », « trop salé » ont des fréquences élevées (respectivement 30, 15 et 8 répétitions). 50% des formes « grignoter » et « éviter » sont rassemblées dans cette classe. Rien d'étonnant à cela, le dispositif d'évaluation qui a mesuré les impacts de cette campagne d'information a mis en évidence que 87 % des Français accueillent favorablement les messages sanitaires insérés dans les publicités alimentaires, et surtout que 71 % d'entre eux les ont mémorisés (INPES, 2014).

Cette classe est donc aussi assez naturellement celle de la restriction : « ni », « pas trop » sont les mots outils qui consolident la classe.

Mots caractéristiques (par ordre de significativité) : Gras, trop, pas, sucre, ni, matière, grignoter, ne, éviter, attention, salé, toujours, graisse, grand, mac_do, abuser, importer, gros, truc, sauter.

Quelques réponses typiques :

« c' est manger équilibré il ne faut pas manger trop gras et trop sucré. Il ne faut pas manger de trop grosses quantités. »

« se régaler en mangeant. Faire attention. Manger pas trop gras ni trop sucré ni trop salé »

« c'est manger équilibré pas trop gras et pas trop salé et pas en trop grande quantité. »

« pas manger trop sucré, ni trop gras »

« manger simple, pas trop gras pas trop sucré pas trop salé. »

« manger pas trop gras et régulièrement en faisant attention aux quantités. »

« manger équilibré et pas de grosses quantités. Ne pas grignoter entre les repas »

Les caractéristiques des enquêtés ayant construit cette représentation sont : sur-représentation significative des femmes (65 % dans la classe contre 47 % dans la population), des jeunes de 15 à 24 ans (23 % dans la classe contre 14 % dans la population), des individus habitant des villes de plus de 100 000 habitants (42 % dans la classe contre 31 % dans la population), , des individus qui habitent dans le Sud-Est de la France (18 % dans la classe contre 12 % dans la population) et des individus les moins actifs (moins de 2h par jour d'activité physique) (69 % dans la classe contre 59 % dans la population).

3.2.3.6 Classe 6 : Fait maison, 15%

Cette représentation du bien manger met l'accent sur le produit, son origine et son histoire. Idéalement, il est frais, il a bon goût, il doit être de saison, « bio » ou local, voire même provenir du jardin pour garantir sa qualité présumée.

La façon dont il est accommodé, préparé et cuisiné est également importante dans les représentations. La cuisine doit être faite maison, les mets préparés simples ou traditionnels. Ces représentations se traduisent par un refus des produits surgelés ou des préparations toutes prêtes et une aspiration à plus de naturalité. Le « fait maison » inclut la notion de « produit maison » (c'est-à-dire produit de son jardin, ou produit localement, acheté directement au producteur).

Cette classe traduit le « fait maison » comme moyen de faire des économies en réponse à la crise (HEBEL, SIOUNANDAN, & LEHUEDE, 2009) ou encore comme solution pour se rassurer face à aux inquiétudes suscitées par la cacophonie nutritionnelle ou les scandales alimentaires (HEBEL, 2008).

L'enquête CCAF 2013 témoigne à cet égard d'une inquiétude liée au scandale à la viande de cheval commercialisée comme étant de la viande de bœuf, fortement médiatisée l'an dernier. Le climat de défiance qui en a résulté explique que les risques envisagés sur la consommation de plats préparés aient progressé. La méfiance vis-à-vis d'une alimentation de plus en plus industrialisée – notamment relative aux OGM, dont l'image s'est fortement détériorée depuis 2007 – entraîne une demande de contrôles sanitaires plus importants. Les industriels de l'alimentation sont les acteurs en qui on a le moins confiance et qui ont la plus mauvaise image.

Cependant, le « fait maison » ne repose pas seulement sur les peurs. Il est aussi une source de plaisir et de satisfaction. En témoigne la mouvance récente pour le « *fait soi-même* » comme réalisation de soi. Nous sommes en présence d'une cuisine choisie, synonyme de plaisir et d'épanouissement, se distinguant d'une cuisine contrainte, plus quotidienne.

La cuisine apparaît ainsi comme un moyen d'expression de soi, en même temps qu'une garantie en termes de provenance et de transparence. Le « fait maison » correspond le plus souvent à des ingrédients dont on connaît l'origine, à une production locale, finalement en harmonie avec la montée de la consommation « engagée » (en particulier dans la défense de l'environnement et des emplois).

Mots caractéristiques (par ordre de significativité) : produit, bon, qualité, frais, maison, cuisine, fait, préparer, naturel, simple, jardin, cuisiner, femme, local, bio, acheter, origine, traditionnel, prêt, saison.

Quelques réponses typiques :

« manger des produits frais cuisinés maison et ne pas acheter de produits tout prêts avec plein de cochonneries à l'intérieur cuisiner des produits du coin »

« manger des choses fraîches qui viennent du marché ou du magasin bio, fait maison »

« les repas que prépare ma femme qui est une bonne cuisinière avec des produits de qualité »

« des légumes du jardin et des produits dont on connaît l'origine »

« la qualité des produits et en frais, pas de surgelés »

« c'est manger dans la qualité des produits, des produits frais, pas forcément dans la quantité, de la bonne marchandise, sans produits chimiques, naturels »

« avoir des très bons produits principalement locaux et préparer généralement le plus simplement possible »

« faire la cuisine et manger des bons produits et de qualité »

« La bonne cuisine, les bons produits. Le plaisir de cuisiner »

Les caractéristiques des enquêtés ayant construit cette représentation du bien manger sont : sur-représentation significative des individus dont le diplôme du chef de famille est de niveau 2ème ou 3ème cycle universitaire ou grande école (23% dans la classe contre 16 % dans la population), dont les revenus du foyer dépassent 45 735 € (12% dans la classe contre 8 % dans la population) ou sont compris entre 22 867 € et 30 490 € (21% dans la classe contre 16 % dans la population), habitant le Sud Est de la France (17% dans la classe contre 12 % dans la population), vivant dans des villes de plus de 100 000 habitants (37% dans la classe contre 31 % dans la population) et ces individus déclarent avoir une bonne alimentation (76 % dans la classe contre 71 % dans la population).

4 ETUDE EN EVOLUTION DES REPRESENTATIONS DU BIEN MANGER DE 1988 A 2013

Après avoir mené cette comparaison entre différents logiciels d'analyse et identifié le plus adapté à notre problématique, nous souhaitons prolonger les travaux initiaux de Saadi Lahlou en mettant en évidence l'évolution des représentations sociales relatives à l'alimentation et au « bien manger » au cours des vingt-cinq dernières années.

4.1 Enjeux autour de l'évolution des représentations sociales du « bien manger »

Pour analyser l'évolution des représentations survenue au cours des 25 dernières années, nous nous appuyons sur la question ouverte qui a été posée lors des enquêtes CCAF de 1988, 1995, 2007 et 2013 : « Pour vous, qu'est-ce que bien manger ? ».

La question alimentaire dans le monde, en Europe et en France est l'objet de nombreuses préoccupations, d'études et de recommandations. L'autorité européenne de sécurité des aliments (EFSA) recommande à ses Etats-membres d'enquêter de façon quantitative et qualitative sur les habitudes de consommations alimentaires de ses ressortissants. En France, la mise en œuvre de programmes alimentaires (PNNS, PNA, plan obésité, etc.) a un impact quantitatif mais aussi qualitatif sur l'alimentation des Français. Des préconisations ont été émises, alors, passant notamment par l'apposition sur les produits de recommandations nutritionnelles (telles que « Mangez cinq fruits et légumes par jour »).

Ces programmes ne sont pas sans conséquences sur les représentations sociales du bien manger : on avait pu constater par exemple en 2007 l'impact des discours nutritionnels sur les représentations du « bien manger » (Mathé, Pilorin, & Hébel, 2008). Les nombreuses recommandations, s'ajoutant les unes aux autres, ont pu accentuer l'impression de « cacophonie » déjà pointée suivant le concept de C. Fischler (1990).

Parallèlement, les scandales alimentaires qui ont émaillé l'actualité depuis les années 1990 (1986-1996 : la crise de la "vache folle" ; 1999 : le scandale du poulet à la dioxine ; 2001 : épizootie de fièvre aphteuse ; 2002 et 2003 : vente de viande sous embargo chez Buffalo Grill ; 2003-2006 : épidémie de grippe aviaire ; 2011 : E.coli dans les concombres ; 2013 : fraude à la viande de cheval dans les lasagnes) ont contribué à fragiliser la notion de « bonne alimentation » en raison des risques sanitaires portés au débit de l'industrialisation des produits alimentaires.

Tous ces facteurs ont influencé les représentations sociales de la question du bien manger et qui doivent pouvoir se mesurer. L'analyse lexicale constitue l'une des méthodes aptes à cette mesure.

Les quatre séquences (années) que nous observons montrent des Français attachés à leur modèle gastronomique (au sens premier) en même temps que soucieux des conséquences de leur alimentation sur leur santé, tout en restant réticents à se plier à une norme (nutritionnelle).

On relève quelques évolutions dans le rapport que les Français entretiennent avec leur alimentation. Ce qui relevait auparavant d'un savoir « naturel », au sens d'intuitif, d'intériorisé, se trouve aujourd'hui davantage tributaire d'un savoir scientifique ou informatif qui donnent aux individus un sentiment d'ignorance et d'inquiétude. Cette prise de distance par rapport à leur alimentation peut être source d'effets positifs comme d'effets pervers. Elle peut laisser la place, outre à des messages contradictoires (la fameuse « cacophonie » décrite par C. Fischler (1990)), à des rumeurs de toutes sortes faisant peser le doute sur la provenance, la composition ou la qualité des aliments (Campion-Vincent, Renard, 2002). La distance concerne aussi le rapport à l'industrialisation de l'alimentation, et la dissolution de la confiance vis-à-vis de produits fabriqués loin et sans visibilité sur la manière dont ils sont effectivement produits.

Les approches nutritionnelles, qui ont mis en avant la responsabilisation et l'autonomie de l'individu, se sont révélées facteur d'anomie et source de vulnérabilité vis-à-vis des industriels de l'alimentation motivés à détruire tout ce qu'il y a de social dans l'individu (Ariès, 1997). Ces approches ont été favorisées, au détriment d'une approche plus collective permettant une rencontre véritable autour de la table, de cultiver le « sens gourmand » (Boutaud, 2005) en même temps que de résister à une standardisation générale en marche.

La diffusion massive de messages délivrant des objectifs nutritionnels, en particulier depuis le début des années 2000, a eu un effet anxiogène en faisant percevoir l'alimentation comme « un problème » et en tendant à réduire le lien de l'individu à son alimentation à la seule problématique sanitaire. L'année 2013 offre toutefois quelques indices d'un retour de balancier en direction d'un rééquilibrage entre les enjeux instrumentaux (la santé, l'équilibre) et les enjeux psychosociaux (le plaisir, le goût, la convivialité). Ce rééquilibrage a d'ailleurs été pris en compte par les pouvoirs publics, qui ont notablement infléchi leur discours en faisant évoluer les objectifs nutritionnels du PNNS vers une approche plus en harmonie avec les contextes culturel et social. Aussi, l'évolution des représentations sociales et du vocabulaire qui les donne à voir doit être appréhendé comme un univers de significations, résultat des influences réciproques des discours des individus et de ceux des institutions publiques ainsi que des intervenants divers concernés par la problématique alimentaire.

4.2 Matériel et Méthodes

Le volet comportement de l'enquête CCAF, dont la méthodologie est décrite au paragraphe 3.1, a été réalisé en 1988, 1995, 1997, 2000, 2003, 2007, 2010 et 2013. On dispose ainsi d'un matériau textuel similaire qui permet une analyse comparative et chronologique.

L'analyse en évolution a été réalisée sur 4 années de référence : on retient tout d'abord l'année 1988 car c'est la première année que la question « Pour vous qu'est-ce que bien manger ? » a été posée. Ensuite, les années 1995, 2007 et 2013, chacune espacée de 6 ou 7 ans, permettent d'observer des changements dans les représentations dans un laps de temps égal. Les années 1988, 1995 et 2007 est aussi la série qui a été étudiée par (Mathé, Pilorin, & Hébel, 2008), ce qui permettra de comparer les résultats obtenus.

Les recueils pour ces quatre années se sont fait dans des conditions similaires, à savoir, en face à face par des enquêteurs formés, qui ont pour consigne de relancer si nécessaire et de retranscrire mot à mot la réponse fournie.

Le matériel d'étude sur les 4 années a subi les mêmes traitements que ceux décrits au paragraphe 3.1 pour le corpus 2013 (correction orthographique, harmonisation du vocabulaire...).

Dans un premier temps, nous étudions en évolution la taille, l'étendue et la richesse du vocabulaire. Cette étape est réalisée avec le logiciel R. Ce choix s'explique par la facilité qu'offre ce logiciel d'extraction, de manipulation de données et de représentation graphique (dont représentations factorielles).

Ensuite, l'analyse de l'évolution du vocabulaire de 1988 à 2013 est réalisée par analyse factorielle des correspondances avec le logiciel R également, et ce pour les mêmes raisons. En effet, un moyen de visualiser graphiquement les changements de vocabulaire qui s'opèrent au cours du temps est de réaliser une analyse factorielle des correspondances sur le tableau lexical agrégé par année. Le tableau soumis à l'analyse contient quatre lignes qui correspondent aux quatre années et en colonne les mots. On cherche à extraire les plus grandes dimensions de variabilité de ce tableau, ce qui permet de visualiser le vocabulaire spécifique de chaque année. On retient l'intégralité du vocabulaire non lemmatisé y compris les mots outils pour réaliser cette AFC. En effet, comme l'explique (Brunet, 2000), si l'émission des formes diminue l'effet de chacune, comme l'analyse du tableau lexical fait la part belle aux fréquences élevées proportionnellement, cet inconvénient est gommé. La lemmatisation n'est pas donc indispensable lorsqu'il s'agit d'obtenir des visualisations. De même, sur un plan factoriel de ce genre, la présence de mot outils est pleinement justifiée car « *si ces mots outils caractérisent électivement certaines catégories, ils se positionnent dans leur voisinage, et peuvent être intéressants à interpréter ; si au contraire leur*

répartition est aléatoire, ils seront situés dans la partie centrale du graphique, sans en encombrer la lecture » (Lebart, 2001).

Enfin, l'évolution des représentations sociales est réalisée par classifications successives grâce au logiciel Alceste pour chacune des années à analyser. Le paramétrage du logiciel est repris comme suggéré dans (Mathé, Pilorin, & Hébel, 2008). Les classes obtenues pour chacune des années sont comparées sur la base de leur contenu et de leur proportion dans les représentations du bien manger mises à nu.

4.3 Résultats

4.3.1 Des corpora différents

Avant même de s'intéresser au sens véhiculé par les mots dans chacun des corpus, suffisamment d'éléments descriptifs permettent d'affirmer que les corpora sont différents. Les quatre corpora analysés totalisent 7665 réponses (Tableau 3).

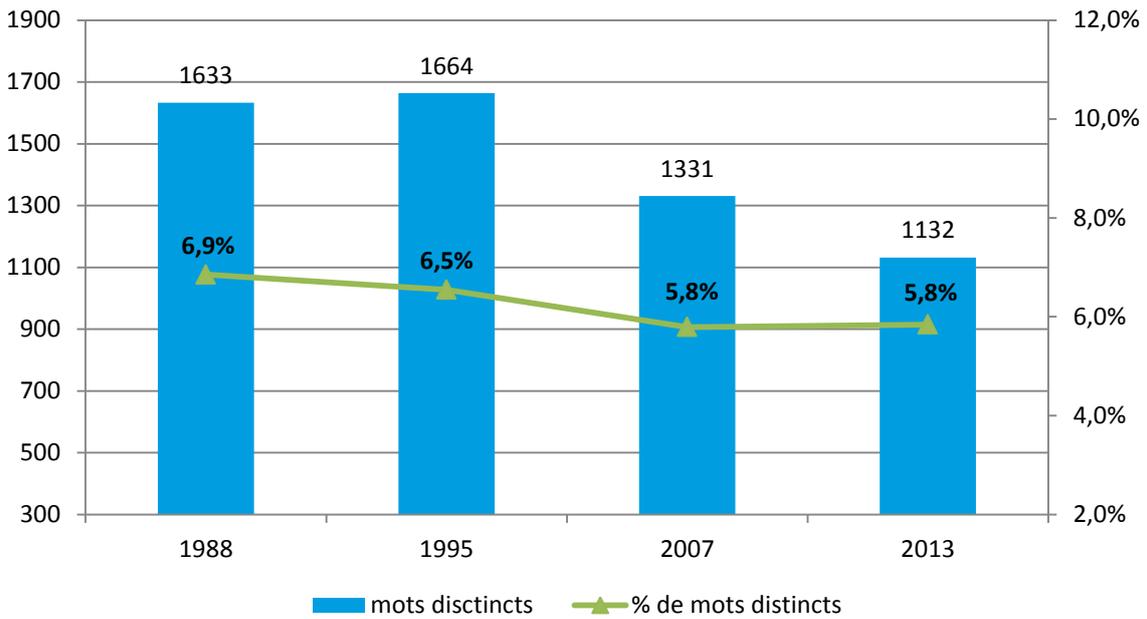
Tableau 3 – Bilan lexical des corpora des 4 années étudiées

Année	Nombre de réponses	Nombre de formes graphiques	Nombre de mots distincts	Nombre d'hapax
1988	1594	23827	1633	739
1995	1598	25415	1664	780
2007	2353	22988	1331	619
2013	2120	19380	1132	526
Total	7665	91610	5760	2664

Source : Enquête CCAF 2013, CREDOC

Malgré le fait que l'on dispose de moins de matériau textuel en 1988 et 1995 que pour les deux dernières années, l'étendue et la richesse du vocabulaire est cependant plus importante dans les deux premières enquêtes. En effet, même en comptant environ un quart de réponses en moins, le nombre de formes graphiques (Figure 10) ainsi que la proportion de mots distincts (Figure 10) est plus important dans les corpora 1988/1995 que 2007/2013. Les terrains d'enquêtes ont pourtant toujours été réalisés par le même institut, GFK-ISL, mais les modes de recueil évoluent passant du recueil papier en face à face à la forme électronique sous forme de saisie informatique (CAWI). La perte de richesse lexicale s'explique par les gains de productivité réalisés par l'Institut et le recueil informatique de l'information plus que par une perte de vocabulaire de la part des interviewés au cours du temps. En tout, les 4 corpora comptabilisent 2899 mots distincts. La proportion d'hapax est plutôt constante et représente 46%.

Figure 10 - Nombre et proportion de mots distincts par année analysée

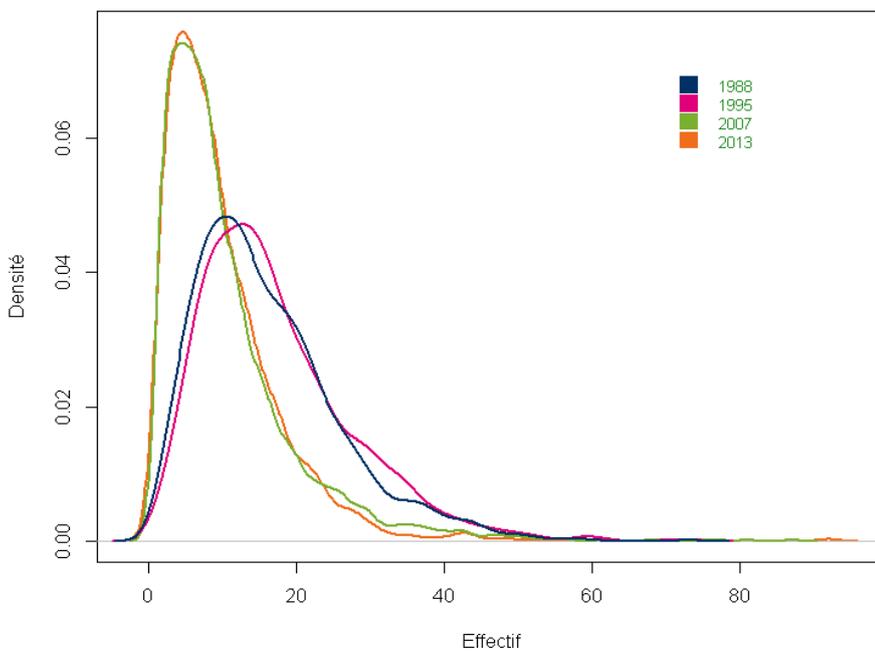


Source : Enquête CCAF 2013, CREDOC

En plus d'une « richesse lexicale » relative plus importante, les corpora 1988/1995 se distinguent par la longueur moyenne supérieure des réponses des enquêtés. En effet, la distribution (Figure 11) de la longueur des réponses montre un écart de 6 à 7 mots entre les corpora 1988/1995 et 2007/2013

Figure 11 - Distribution de la longueur des réponses

Distribution des longueurs de réponses

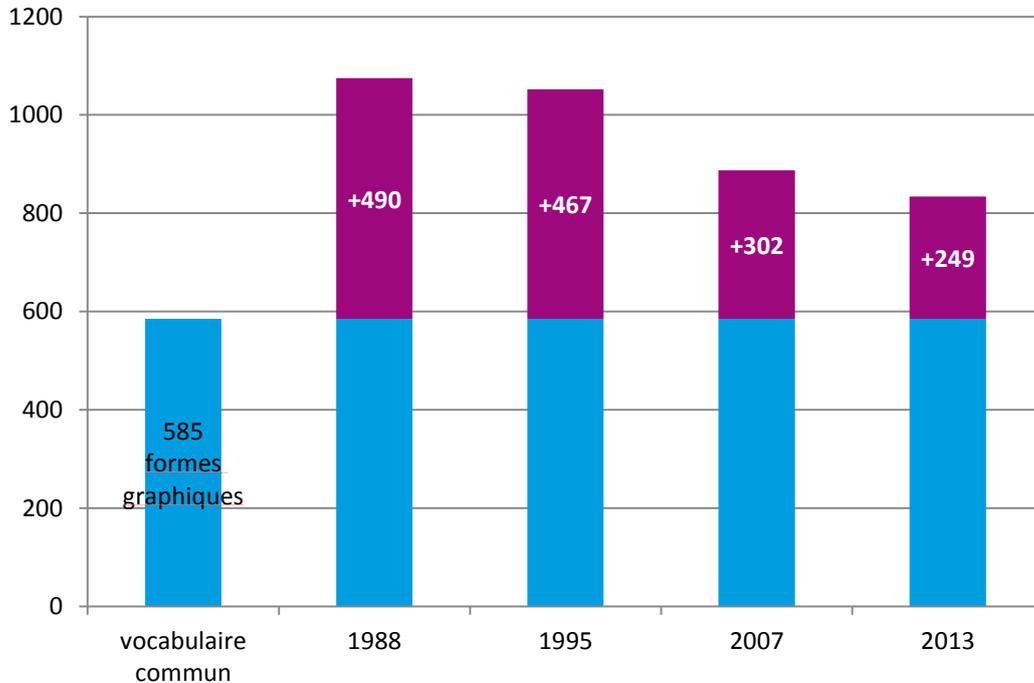


Année	Longueur moyenne des réponses
1988	16 ±9
1995	17 ±10
2007	10 ±10
2013	10 ±8

Source : Enquête CCAF 2013, CREDOC

Lorsqu'on analyse plus finement les mots distincts en présence dans chacun des corpus, on dégage un « tronc commun » de vocabulaire de 585 termes (Figure 12) auquel vient s'ajouter une proportion plus ou moins variable de vocabulaire spécifique d'une année :

Figure 12 - Vocabulaire commun aux 4 années analysées et vocabulaire spécifique à chaque année (en nombre de mots distincts)



Source : Enquête CCAF 2013, CREDOC

On constate encore une fois que les années 1988/1995 se distinguent de 2007/2013 par un vocabulaire spécifique plus conséquent.

Deux groupes de corpus semblent donc se distinguer. Le groupe 1988/1995 se caractérise par des réponses longues (plus de mots), avec un vocabulaire riche (plus de mots distincts) et original (plus de mots particuliers, d'hapax). Le groupe 2007/2013 quant à lui se caractérise par des réponses courtes (moins de mots), qui utilisent un vocabulaire consensuel (partagé par plus d'individus) et compte plus d'individus interrogés.

Deux raisons pourraient expliquer ces différences. Tout d'abord, on peut émettre l'hypothèse qu'entre 1995 et 2007 une « harmonisation » des représentations s'est opérée. Cela pourrait effectivement impacter la richesse du vocabulaire et la longueur des réponses, avec plus de réponses « toutes faites » et un inconscient collectif plus formaté. Ensuite, un changement méthodologique dans le recueil des réponses (relance, nombre de mots minimum...) pourrait expliquer ces variations. Il n'est pas à exclure non plus une conjonction de ces deux explications.

4.3.2 D'importantes variations de vocabulaire

Intéressons-nous maintenant à la nature du vocabulaire de façon qualitative. Une première approche pour appréhender le vocabulaire d'un ensemble de textes est de trier les formes graphiques ou les lemmes des textes et de comparer leurs fréquences et leur rang d'utilisation (Tableau 4).

Tableau 4 - Les 20 formes graphiques les plus fréquentes dans les corpora analysés

1988		1995		2007		2013	
forme	occurrences	forme	occurrences	forme	occurrences	forme	occurrences
manger	1146	manger	1065	manger	1977	manger	1831
bien	503	bien	540	équilibré	785	équilibré	677
repas	372	bon	428	repas	376	repas	291
bon	294	repas	349	bon	300	produits	252
choses	239	équilibré	255	tout	251	bon	244
viande	220	avoir	240	plaisir	244	légumes	240
aime	216	produits	240	faire	243	plaisir	236
légumes	215	aime	212	bien	235	bien	221
équilibré	214	faire	201	légumes	227	tout	221
dessert	185	qualité	201	avoir	225	faim	215
avoir	182	plaisir	194	aime	213	aime	210
tout	180	bonne	192	trop	210	faire	196
faire	176	tout	167	faim	190	avoir	180
plats	171	choses	164	produits	175	choses	158
entrée	153	plat	154	choses	146	trop	151
fromage	143	plats	135	fruits	141	fruits	144
plat	140	cuisine	134	varié	140	qualité	125
plus	139	chose	130	bonne	135	bonne	123
bonne	133	temps	130	aliments	126	quantité	117
temps	124	frais	129	quantité	126	bons	104

Source : Enquête CCAF 2013, CREDOC

Un certain nombre de mots conservent le même rang dans les classements de chaque année : « manger », « bon », « repas », « aime », « bonne » et « faire ». On constate par contre une descente au classement du mot « bien » qui passe du 7^{ème} (1988) ou 8^{ème} (1995) rang, au 19^{ème} et 21^{ème} rang en 2007 et 2013. Cela s'interprète, après analyse des segments, comme une diminution de l'écholalie : on ne répète plus à partir de l'enquête 2007 « bien manger c'est... » ou, dans le cas où ces mots sont prononcés, ils ne sont pas recueillis. Le syntagme « bien manger » est présent 184 fois dans le corpus 1988 alors qu'il n'est présent que 53 fois dans le corpus 2013.

La forme graphique « choses » au pluriel descend elle aussi dans le classement entre 1988 et 1995. Après analyse des segments, il semblerait qu'il y ait de moins en moins de description par l'exemple avec des locutions telles que « bonnes choses » (135 occurrences du segment répété en

1988 vs 81 en 2013) ou « manger des choses » (70 occurrences du segment répété en 1988 vs 58 en 2013). Il est possible que la forme graphique « produits » se substitue à « choses » via les segments « bons produits », « produit de qualité », « produits sains ».

De la même façon, la forme graphique « temps » du segment répété « prendre le temps » ou « prendre son temps » disparaît du classement des 20 formes graphiques les plus utilisées des corpora 2007 et 2013.

En revanche, les corpora 2007 et 2013 voient apparaître de façon plus fréquente les formes « plaisir » (à travers les segments répétés « se faire plaisir », « prendre plaisir à », « avec plaisir », « faisant plaisir »...) et « équilibré » (« manger équilibré », « manger équilibré et », « repas équilibré »...). Par ailleurs, la forme « faim » des segments répétés « manger à sa faim » ou « ne plus avoir faim » fait son apparition dans les classements 2007/2013. Ces segments peuvent évoquer la notion de satiété ou le poids de difficultés financières qui décrit le « bien manger » comme d'abord un besoin vital malaisé à satisfaire.

Plus largement et sans se cantonner aux 20 premières formes graphiques pleines, l'examen des formes graphiques de l'ensemble des corpora permet de dégager les mots qui apparaissent au cours du temps et ceux qui disparaissent des discours. Les tableaux 5 et 6 (ref) présentent les mots qui ont le plus perdu ou gagné de place dans le classement entre 1988 et 2013. Seules les formes graphiques avec une fréquence supérieure à 10 sont analysées : c'est le seuil, choisi arbitrairement pour ce corpus, qui permet de distinguer de réelles variations de rang et ne pas prendre en compte le « bruit » lié au vocabulaire d'individus particuliers. Les mots sont présentés par champ lexicaux, définis arbitrairement.

Tableau 5 - Gains de rang les plus importants sur le tableau des occurrences entre 1988 et 2013

"Les nouvelles tendances"		"le discours nutritionnel"	
bio	+434	salé	+673
assiette	+106	sucré	+533
maison	+54	eau	+347
plaisir	+48	gras	+95
		laitiers	+72
		fruits	+47
"la restriction et la surveillance"		"nutrition"	
raisonnables	+1452	protéines	+482
raisonnable	+1363	féculents	+163
grignoter	+1125	santé	+120
heures	+187	sainement	+75
éviter	+174	alimentation	+50
entre	+73	sain	+32
attention	+71		

Source : Enquête CCAF 2013, CREDOC

Le gain ou la perte de places est à interpréter avec précaution dans la mesure où le nombre et la nature des mots sont différents dans les corpora 2013 et 1988, c'est-à-dire que l'on s'expose à des gains ou des pertes exagérés. De plus, seules les variations entre les années 1988 et 2013 sont exploitées ; on ne tient pas compte des années intermédiaires. Cependant, ces observations donnent des résultats intéressants à discuter.

Les formes graphiques qui ont le vent en poupe appartiennent tout d'abord au champ lexical de la restriction et de la surveillance. Les formes « raisonnables » et « raisonnable » sont principalement assorties du mot « quantité » : il faut manger (de tout) en quantité raisonnable. « Grignoter » (+1125) est une forme qui apparaît en 2013, il faut éviter de grignoter « entre » (+73) les repas. De la même façon, « heures » prend 187 places car les dits repas doivent être pris à « heures régulières ». Somme toute, dans cette représentation restrictive de l'alimentation, il faut faire « attention » (+71). Cependant, ces formes n'entrent pas forcément en contradiction avec le respect du rythme des prises alimentaires reposant sur des moments fixés dans la journée, sans grignotage entre les repas. Il s'agit aussi de se discipliner collectivement pour manger en quantité raisonnable.

Ensuite, le champ lexical du discours nutritionnel se pérennise en 2013. « Salé » (+673), « sucré » (+533) et « gras » (+95) grimpent dans le classement, adjectifs dans lesquels on peut déceler l'écho du message véhiculé dans le cadre du PNNS « Pour votre santé, évitez de manger trop gras, trop sucré, trop salé ». « Laitiers » (+72) et « fruits » (+47) au pluriel connaissent une progression moins flagrante mais font également penser aux recommandations des « cinq fruits et légumes par jour » et des « 3 produits laitiers par jour ». Adossé au champ lexical du discours nutritionnel on trouve le champ lexical de la nutrition. « Protéines » (+482) au pluriel est le plus souvent cité dans « des protéines » et participe à la description de ce que doit contenir un repas. La forme « Féculents » (+163) s'insère dans le segment « des légumes et des féculents ». « Santé » (+120) se retrouve dans les expressions « bon pour la santé » « être en bonne santé ». La progression de ces formes traduit une préoccupation plus importante concernant l'alimentation.

A ce sujet, on doit souligner que cette progression ne fait pas apparaître les années intermédiaires. Or, la prise en compte de l'année 2007 montre une inflexion en 2013 relativisant la progression de l'approche nutritionnelle que dessine la courbe 1988-2013.

La dernière famille de mots à avoir la côte est intitulée « nouvelles tendances » car elle est le reflet de concepts émergents comme l'attrait pour le « bio » (+434), l'engouement pour le fait « maison » (+54) et la ferveur pour la gastronomie à la française des « belles assiettes » ou des « bonnes assiettes ».

En somme, le retour du plaisir n'efface pas les inquiétudes sur les risques alimentaires : En 2013, la santé n'est pas perçue comme mise en danger par une mauvaise alimentation qui serait le fait de son propre comportement (et dont on se sentirait coupable) mais du fait des produits eux-mêmes (tromperie, contaminations-dont OGM, défaut de traçabilité, intermédiaires trop nombreux). D'où le désir de revenir au local (comme garantie plus forte de produits sains). A ce titre, le « fait maison » apparaît aux yeux des consommateurs comme une garantie de manger plus sainement qu'avec des plats tout-prêts (cf. scandale à la viande de cheval vendue comme du bœuf en 2013). Cuisiner est aussi une source de plaisir en soi, et en perspective du point de vue du goût et de la convivialité, car on ne cuisine guère pour soi uniquement : la manière de s'alimenter des personnes seules en témoigne.

Trois autres registres décrivent les formes graphiques délaissées depuis 1988 (Tableau 6). D'abord le champ lexical du contexte de la prise alimentaire montre un décrochage de la forme « enfants » (-1015) en 2013. Ce mot disparaît presque du lexique du corpus. Les références à l'ambiance, qu'elle soit festive (« fête » perd 243 places) ou plus sereine (« calme » et « calmement » perdent respectivement 182 et 161 places) se font plus rares. Viennent ensuite les composants du repas qui semblent ne plus correspondre à la vision qu'ont les Français du bien manger en 2013 : exit le « café » (-998), les « sauces » (-920) et la « sauce » (-316), « la charcuterie » (-228) et dans une moindre mesure « le vin ». Le recul de « hors d'œuvre » (-936) et « mets » (-378) est peut-être autant dû à l'évolution du vocabulaire qu'à la déconstruction et la simplification des repas. Remarquons que « fruit » (-49) au singulier est un mot qui régresse alors même qu'au paragraphe précédent nous avons fait le constat de la progression de « fruits » au pluriel. Ceci s'explique par la diminution des descriptions de repas du type « une entrée, un plat, un fruit, un yaourt ».

Tableau 6 - Pertes de rang les plus importantes sur le tableau des occurrences entre 1988 et 2013 :

"le contexte"		"la composition des repas"		"la lourdeur calorique"	
enfants	-1015	café	-998	lourd	-957
fête	-243	hors-d'œuvre	-936	ordinaire	-882
calme	-182	sauces	-920	calorie	-226
calmement	-161	mets	-378	préparation	-193
ambiance	-144	sauce	-316	vite	-168
temps	-50	charcuterie	-228	léger	-146
		vin	-90	régime	-137
		fruit	-49	copieux	-41

Source : Enquête CCAF 2013, CREDOC

4.3.3 1988-1995 et 2007-2013 : deux univers contextualisés

En complément à cette première approche de l'analyse du vocabulaire, une autre façon d'explorer le contenu du corpus consiste à repérer les segments de texte les plus répétés. Il existe plusieurs façons de les identifier, notamment sur des critères de longueurs et d'occurrences dans les documents. Dans le Tableau 7, on extrait pour chaque corpus les motifs de texte de longueur de 4 termes au moins qui ont été repérés au moins 3 fois dans 3 réponses différentes.

Tableau 7 - Fréquence des segments répétés au moins 3 fois dans 3 réponses différentes dans les corpora analysés

1988		1995	
bien manger c est	74	bien manger c est	67
que l'on aime	59	ce qu'on aime	44
ce que l'on	57	que l'on aime	41
manger à sa faim	41	manger ce qu'on	39
manger ce que l	36	ce que l'on	38
manger ce que l'on	36	que ce soit bon	32
ce que l'on aime	34	manger à sa faim	30
pour moi c'est	32	manger ce qu'on aime	30
une entrée un plat	31	une entrée un plat	27
quelque chose de bon	27	faire un bon repas	26
2007		2013	
C'est manger équilibré	77	ce que j'aime	60
manger à sa faim	61	ce que l'on	58
ce que j'aime	56	manger ce que j	55
ce qu'on aime	47	que l'on aime	51
manger ce qu'on	44	manger à sa faim	47
un peu de tout	44	manger ce que j'aime	47
manger ce que j	42	C'est manger équilibré	40
ne plus avoir faim	41	ce que l'on aime	39
ce que l'on	40	C'est manger des	35
manger ce que j'aime	40	manger ce que	34

Source : Enquête CCAF 2013, CREDOC

En ayant bien à l'esprit qu'observer le texte à travers le prisme des segments répétés pose des hypothèses assez fortes (en particulier, on ne considère que des motifs de texte d'une longueur définie alors que le texte ne peut être réduit à cette seule dimension), on confirme tout de même que le phénomène d'écholalie est particulièrement présent, ou retranscrit comme tel, dans les corpora 1988/1995. En effet, le segment répété le plus fréquemment est « bien manger c'est » en écho à la question posée. Ce qui différencie également 1988/1995 des deux autres années considérées dans ce travail est la présence de l'expression « une entrée, un plat » qui suggère que le « bien manger » pour ces périodes s'illustre dans la description des séquences d'un repas.

Les dimensions « manger ce que l'on aime » et « manger à satiété » sont communes aux quatre corpora. Cependant la dimension de l'équilibre, comme l'avait déjà montré le tableau des fréquences et rangs des formes les plus citées, est plus abordée par les répondants des années 2007/2013.

Enfin, un dernier moyen d'explorer le contenu consiste à réaliser des comptages de réponses (Tableau 8), indicateur du caractère stéréotypé de celles-ci.

Tableau 8 – Les 5 réponses les plus citées dans les corpora des 4 années analysées

réponse	1988	1995	2007	2013
« manger équilibré »	7	9	155	140
« manger à sa faim »	5		11	26
« manger bien »				10
« manger ce que j'aime »	3		12	10
« manger de tout »			19	9
« bien se nourrir »			13	
« aller au restaurant »	6			
« ce qui plaît »	2			
Total	23	9	210	195

Source : Enquête CCAF 2013, CREDOC

Il existe plus de réponses-types dans les corpora 2013/2007 que 1988/1995. Même si le nombre de répondants est plus important, on trouve ici l'empreinte d'une homogénéisation du vocabulaire et des expressions concernant le bien manger. En 2013 et 2007, ce sont respectivement 140 et 155 enquêtés qui ont eu pour seule réponse « manger équilibré » à la question « Pour vous, qu'est-ce que bien manger ? » soit 6% des réponses. En revanche on ne retrouve cette réponse laconique que dans respectivement 0,5% et 0,4% des cas dans les corpora 1988/1995.

En forme de synthèse, on confirme les différences qui existent entre les corpora 1988/1995 et 2007/2013, soulignées plus haut. Soulignons en particulier les points suivants :

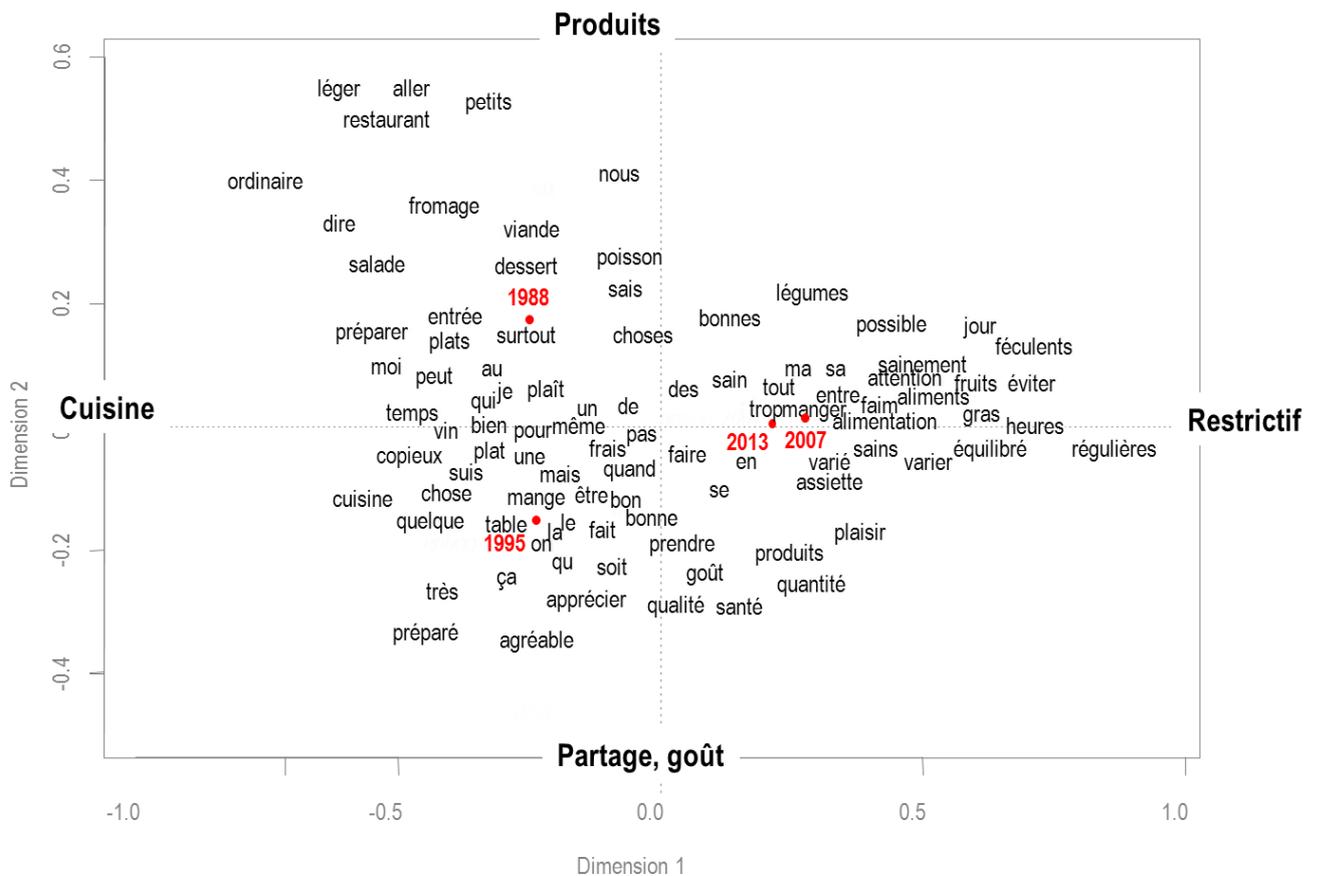
- il existe d'importantes variations autour d'un tronc commun de vocabulaire;
- les réponses sont de plus en plus stéréotypées, formatées ;
- il émerge des idées nouvelles (bio, fait maison) et des évolutions sont à l'œuvre dans le contexte de la prise alimentaire (enfants, restaurant...), en lien avec les mutations sociétales ;
- la « nutritionnalisation » des discours est palpable ;
- on note un recul des représentations descriptives (« une entrée, un plat, un dessert, un café »...).

4.3.4 Evolution des représentations du « bien manger » par analyse factorielle des correspondances : l'espace des représentations

La représentation dans l'espace des imaginaires alimentaires offre la possibilité d'observer sur un même plan les différences de vocabulaire selon les années, et les tendances dominantes pour chacune des quatre années. Cela permet aussi de visualiser les associations de mots apparues dans les réponses des personnes enquêtées, comme par exemple « aller » et « restaurant ».

Les individus (les années) et les variables (les formes lexicales) sont projetés sur le plan factoriel et séparés selon les plus grandes dimensions de variabilité du tableau, c'est-à-dire selon les plus grandes variations de vocabulaire. Deux années sont d'autant plus proches sur le plan factoriel qu'elles se ressemblent du point de vue du vocabulaire utilisé. Deux mots sont d'autant plus proches qu'ils sont utilisés la même année. Et enfin, une année est au barycentre des mots particulièrement cités à cette date. Seuls les 100 mots les plus contributifs aux deux axes sont représentés pour des questions de lisibilité (Figure 10).

Figure 10 - Plan factoriel de l'analyse des correspondances du tableau lexical agrégé par année étudiée



Source : Enquêtes CCAF 1988, 1995, 2007 et 2013, CREDOC

Le plan factoriel totalise $76,3+17,7 = 94\%$ de l'inertie totale. C'est une valeur élevée, témoin de la qualité de la représentation du plan. Cependant, la somme des valeurs propres vaut $0,05895 + 0,013646 + 0,004691 = 0,077287$ et égale l'inertie totale, c'est une valeur faible comparée à sa valeur maximum qui est égale à 3 (chaque valeur propre est comprise entre 0 et 1). Cette valeur extrême correspondrait à un vocabulaire spécifique et étanche pour chaque année étudiée, elle nous renseigne ici dans l'étude du nuage de point sur le fait qu'une bonne partie du vocabulaire est commun à toutes les années.

Il reste que les années 2007 et 2013 sont plutôt proches l'une de l'autre et plutôt éloignées de 1988 et 1995 sur l'axe horizontal. Ces deux dernières sont opposées également sur l'axe vertical.

Les pourcentages d'UC appartenant à des thèmes particuliers nous renseignent sur la prédominance d'une facette des représentations de l'alimentation propre à une période.

L'année 1988 : prépondérance du « Ce qu'on mange » (exemples de plats, d'aliments)

L'année 1988 est située à proximité du champ lexical du « bien manger par l'exemple ». Cette date est en effet au barycentre de mots qui décrivent l'ordonnance d'un repas (« entrée », « fromage », « dessert », « plats », « complet ») ou les aliments qui le composent, (« salade », « viande », « poisson »,). Ces exemples se retrouvent toutes les années mais l'année 1988 se caractérise effectivement par l'utilisation prépondérante de noms d'aliments ou de plats, en termes de fréquences, pour définir ce qu'est « bien manger ». La description factuelle prime sur un raisonnement plus abstrait. Les exemples se rapportent aussi aux séquences d'un repas structuré, comme « entrée » ou « dessert ». Bien manger, c'est manger un repas « complet ». Cela nous suggère qu'est fait référence, de manière implicite, à la structure traditionnelle du repas français.

Le pourcentage d'unité de contexte (UC) participant de cette classe varie en effet, de 30 % des UC en 1988, à 23 % en 1995, 16 % en 2007 et 21% en 2013.

Pour expliquer cette moindre présence de l'exemple, même si 2013 marque une reprise significative, on avance l'hypothèse d'une familiarisation au vocabulaire et aux concepts nutritionnels, ceux-ci étant employés au détriment de l'expérience vivante.

Les résultats sont proches de ceux relevés par (Mathé, Pilorin, & Hébel, 2008) : la question « Pour vous qu'est-ce que bien manger ? » est comprise comme une demande de conseils à laquelle les enquêtés ont plus tendance à répondre par des préconisations : « c'est manger un repas complet qui contient une entrée et un dessert » et « La qualité se trouve en sortant de l'ordinaire ». Cette année est aussi, en effet, proche des mots « ordinaire » et du couple « aller » / « restaurant ».

Le « ce qu'on mange » se rapporte également aux effets attendus des aliments consommés, en particulier du point de vue de la santé. De fait, l'utilisation des mots comme « calories » et « régime » montre la préoccupation d'un certain nombre d'individus dans leur manière de définir une bonne alimentation. La place prise en 1988 par les contenus du « bien manger » explique que les dimensions de plaisir et de convivialité soient moins mises en avant.

L'année 1995 : prépondérance du « Comment on mange » (convivialité, goût, plaisir, qualité)

L'année 1995 se situe plutôt au barycentre du champ lexical du « bien manger comme acte social ». Les mots « apprécier », « agréable », « préparé », « table », « cuisine » ont été particulièrement utilisés cette année-là. Encore une fois, les résultats sont proches de (Mathé, Pilorin, & Hébel, 2008) : l'alimentation est connotée positivement et est en lien avec le plaisir gustatif et le plaisir de l'acte social. Le suremploi du mot « déjà » est particulier à cette année d'étude. Le retour au texte nous informe que cela correspond à une hiérarchisation des concepts du bien manger, comme un retour à ce qui est important (*exemple : « se faire plaisir déjà et manger des bonnes choses, des aliments de qualité »*). Ce terme est présent moins de 15 fois sur les années 1988, 2007 et 2013 tandis qu'il est présent plus de 50 fois en 1995.

L'année 1995 est l'année faste du plaisir de manger, qu'il s'agisse de faire référence au goût des aliments ou à la convivialité. On peut y voir la quintessence du modèle alimentaire français, mariant le partage de la nourriture, le savoir-vivre et le bon goût des aliments.

Les mots tels que « savourer, déguster, raffiné, goût, ensemble » témoignent du plaisir associé au fait de « bien manger », tant en termes de contenu qu'en termes de forme.

La dimension sociale du « bien manger » apparaît clairement dans cette année 1995. Elle reflète cependant – avec une intensité particulière – une constante du modèle alimentaire français, qui se caractérise par une approche de l'acte de se nourrir moins utilitariste, valorisant le moment vécu plus que l'énergie absorbée. De ce fait, au rapport entre le mangeur et son aliment se substitue celui des mangeurs entre eux, notamment par les termes de « convivialité, présentation, compagnie ». Ce qui est mangé devient prétexte à être ensemble, à table, au cours d'un repas.

Les discours alimentaires propres à cette année-là mettent aussi en avant la dimension de plaisir du produit et des manières de préparer les plats (« cuisiner, préparation »).

La dimension de la santé n'est pas absente du rapport à l'alimentation cette année-là. Le terme de « besoins » et l'expression de « produits naturels » souligne cet aspect. Cependant, on doit

remarquer que le goût est devenu une composante indispensable des régimes, relativisant par-là l'attente purement instrumentale ordinairement attachée à ceux-ci.

L'année 2007 – Faire attention à ce qu'on mange (équilibre, santé, restriction)

L'année 2007 se caractérise par un vocabulaire orienté vers la santé. Le mot « équilibre » qui se situait en 1988 et en 1995 en 18^{ème} et 19^{ème} positions, en termes de fréquence de citation, arrive en 5^{ème} position en 2007. L'équilibre fait référence à une nourriture fraîche, naturelle, variée, consommée avec modération. On peut rappeler que dans la conception grecque (Galien) de la santé, la rupture de l'équilibre naturel, de l'équilibre des humeurs notamment, est perçue comme l'origine de la maladie.

Les individus correspondant à cette classe maîtrisent les termes scientifiques issus de la nutrition (lipides, glucides, protides, etc.) et se rapportent explicitement au discours nutritionnel ambiant en 2007.

Les manières de concevoir le rapport entre l'alimentation et la santé diffèrent selon les individus et au cours des années. En 1988, le champ lexical (« riche, lourd, calorie, excès ») pointe les excès alimentaires. En 1995, ce sont les mots « vitamines, glucides, protéines, lipides et éléments », directement issus des messages nutritionnels, qui marquent le début de la complexification et de la médicalisation des discours portés sur l'alimentation. En 2007, ce glissement sémantique va jusqu'à engendrer une nouvelle classe développant l'idée de la conscience d'un problème, à travers les mots « raisonnable, équilibré, quantité, normal, suffisant, éviter, grossir, santé ». Si en 1995, le naturel garantissait une « bonne alimentation » du point de vue de la santé, c'est la raison et la mesure qui l'assurent en 2007, comme l'indiquent les mots « raisonnable » et « éviter ».

De fait, l'année 2007 illustre un certain désenchantement de l'alimentation. Les individus expriment leur idée du « bien manger » par des mots dénotant la « conscience d'un problème », particulièrement relatif à la santé. L'augmentation de la diffusion des messages nutritionnels, particulièrement à partir de la mise en place du PNNS en 2001, se ressent dans les mots « fixe, régulier, peu et trop », qui connotent une restriction. Parallèlement, « le sucre, les féculents, l'eau, les laitages et le gras » apparaissent comme typiques de cette année 2007. Les produits dont on parle sont bruts (eau, sucre) ou regroupés en classes (les féculents), à l'instar des discours nutritionnels. En toute logique, à l'inverse de 1988, certains produits sont cités comme contraires à l'idée que l'on se fait du « bien manger » en 2007, tels que la viande, le fromage et le vin, qui sont significativement moins présents.

L'année 2013 – Prépondérance d'une alimentation saine faite soi-même (goût, qualité, santé)

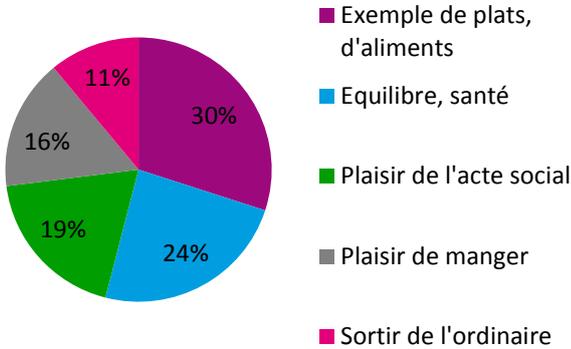
Les années 2007 et 2013 se situent à la croisée des champs lexicaux relatifs au goût et plaisir (« prendre », « plaisir », « goût », « qualité », « quantité ») d'un côté et celui de la diététique et la nutrition (« fruits », « légumes », « sainement », « jour », « manger », « varié », « nourriture », « éviter », « gras », « régulières », « équilibré », « santé ») de l'autre. Les résultats ici sont sensiblement différents des résultats présentés par (Mathé, Pilorin, & Hébel, 2008) dans la mesure où l'année 2013 apporte son lot de vocabulaire et modifie la forme du nuage de point étudié et projeté mais l'année 2007 reste cependant l'année de la restriction et d'un rapport nouveau à l'alimentation : plus raisonné et mesuré. Le terme « possible » mérite d'être commenté également pour l'année 2013 : dans le corpus, il est le reflet d'une conscience du compromis comme s'il existait un idéal du « bien manger » à atteindre, influencé par le discours nutritionnel (exemple : « *c'est manger équilibré et manger le moins de matières grasses possible* », « *manger à sa faim. Ce que nous avons envie de manger, si possible se faire plaisir* », « *c'est manger quelque chose de bon et si possible équilibré* »). En d'autres termes, le retour du plaisir ne signifie pas le retrait des préoccupations nutritionnelles et d'équilibre mais leur intégration dans un bouquet de représentations permettant d'envisager le plaisir sans abandonner ni non plus se concentrer exclusivement sur les aspects d'équilibre et de nutrition. Le « Bien manger » n'est plus un « problème » mais un objet d'attention.

Le placement des représentations sur un graphique permet de visualiser le mouvement de retrait par rapport au pôle « restrictif » en 2013. L'attraction amorcée en 2007 reste, certes, une tendance forte, mais elle ne s'impose pas de manière exclusive et définitive. Les représentations reprennent par exemple le chemin du goût, affirmant de cette façon que le manger « sainement » et « équilibré » ne suffit plus – en tout cas moins qu'en 2007 – pour déterminer la composition d'un menu.

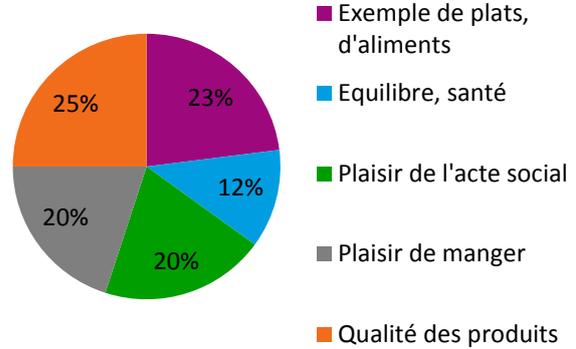
4.3.5 Evolution des représentations du « bien manger » par comparaison de classifications

L'analyse temporelle des représentations avec le logiciel Alceste produit 4 classifications successives qu'on compare du point de vue des proportions et de la nature des classes mises à jour.

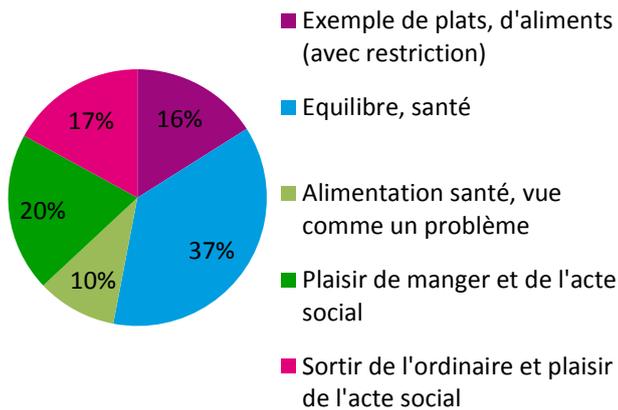
1988



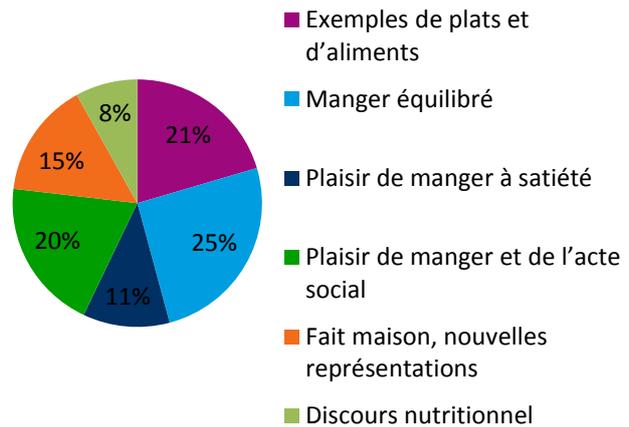
1995



2007



2013



On peut dès lors observer que **trois thèmes sont restés communs aux quatre années** : le thème de l'alimentation par l'exemple, le thème de l'alimentation plaisir (comprenant le plaisir de manger et le plaisir de l'acte social), et le thème de l'alimentation santé (regroupant deux classes en 2007 et en 2013).

Chaque année dispose ensuite d'au moins une représentation qui lui est propre :

- 1988 est marquée par la représentation "sortir de l'ordinaire" pour 11% ;
- En 1995, c'est la qualité des produits, représentant un quart des représentations, qui est mis en évidence dans les discours de cette année ;

- 2007 voit l'apparition de la représentation du bien manger comme un problème (10%) ;
- Enfin, 2013 s'illustre par deux nouvelles représentations : le fait maison (15%) et le discours nutritionnel (8%).

4.3.5.1 Les proximités entre 2013 et 1995

Les années 2013 et 1995 se ressemblent en bien des points, du point de vue des représentations :

- Les restrictions et les problématiques soulevées par l'alimentation en 2007 ont régressé et on constate un retour au plaisir (plaisir de manger lié à l'acte social (20%) et lié à la satiété (11%)) ;
- L'apparition de la classe dite du « discours nutritionnel » traduit, comme pressenti en partie 3, une homogénéisation des représentations par la standardisation des messages véhiculés par le PNNS. Le nutritionnel apparaît comme un prérequis ;
- Le thème de la **qualité des produits** caractérise fortement l'année 1995. Le vocabulaire allie la qualité des produits à la qualité de l'alimentation. Cette classe contient les mots « produits, qualité, frais, naturel, sain, goût, bon, saveur ». Absente en 1988 et en 2007, elle se retrouve en 2013 par le biais du « **fait maison** », sous la forme d'une réappropriation des produits alimentaires pour en maîtriser la production et la transformation, en parallèle de la montée de la production locale à favoriser. De ce fait, cuisiner à la maison apparaît comme un maximum en termes de proximité et de local ! A titre d'hypothèse, on peut avancer qu'en 2013, la mauvaise alimentation est peut-être moins le fait de son propre comportement que celui des produits eux-mêmes (tromperie, contamination, inquiétudes sur les risques alimentaires). La défiance encourage à cuisiner par soi-même, à manger « local » (origine, circuit court, traçabilité jugée plus fiable) ;
- Une autre facette du « bien manger » est représentée par une classe formée par le champ lexical de la sortie, justement absente en 1995 et en 2013 mais qui constitue 11 % des UC en 1988 et 17 % en 2007. Il s'agit pour les deux années, des mots « restaurants, aller, ordinaire, sortir, amis, inviter, extérieure ».

4.3.5.2 *L'année 2013 en recentrage par rapport aux préoccupations nutritionnelles de 2007*

Dans les représentations du « bien manger » observées dans les enquêtes, la dimension du plaisir partagé est prédominante en France jusqu'en 2007. En 2007, un changement de tendance est observé : le mot le plus cité à la question « Si je vous dis 'bien manger', à quoi pensez-vous ? » est : « équilibre ». La dimension restrictive de l'alimentation fonctionnelle est nettement plus présente qu'en 1995 et 1988, notamment chez les plus jeunes, ce qui traduit au premier abord un changement dans les connaissances mais qui pourrait indiquer un renversement générationnel dans les représentations sociales.

Entre 1988 et 1995, les représentations qui s'orientaient sur la santé perdaient de l'ampleur. Celles-ci prennent une place très importante en 2007 mais tendent à régresser de nouveau en 2013 (sans cependant perdre leur importance). En effet, 26 % des réponses se classent dans la catégorie équilibre en 2013, contre 37% en 2007 (et seulement 12 % en 1995), signe d'un retrait notable de cette dimension. On doit souligner qu'il s'est écoulé douze ans entre 1995 et 2007, mais seulement six entre 2007 et 2014 : la décade n'est peut-être que relative, mais elle peut aussi n'être qu'à mi-chemin.

Quoi qu'il en soit, la dimension culturelle (par exemple le plaisir lié à la « table ») progresse significativement entre 2007 et 2013, marquant un certain retrait des préoccupations relatives à la santé. Certains retournent vers le « fait maison » pour maîtriser le contenu des aliments ingérés et la recherche de produits plus naturels remplacent la notion de produits sains. D'autres choisissent de développer leurs talents culinaires, qui leur permettent d'allier le plaisir (de préparer, de déguster) et la confiance (fait par soi-même à partir de produits bruts) dans son alimentation.

Les classes associées au **plaisir de manger** (« faim, aimer, manger, agréable, content, bouffer ») et au **plaisir de l'acte social** (« temps, prendre, famille, ami, retrouver, ensemble ») restent stables sur les années 1988 et 1995 (environ 20 % d'UC pour chacune d'elle dans ces deux années), mais subissent une fusion en 2007. On retrouve les aspects du plaisir de l'acte social et de l'acte de manger dans une même classe totalisant 20 % des UC.

On observe une prédominance de l'alimentation « plaisir » en 1995 (35 % en 1988, 40 % en 1995 puis 20 % en 2007 et enfin 31% en 2013) et une augmentation forte du thème de l'alimentation santé en 2007 tandis que 2013 marque une décade de cette dimension (24 %, 12 % puis 47 % en 2007 et enfin 33% en 2013).

Le plaisir de l'acte social demeure une constante : il est l'une des seules composantes du modèle alimentaire français qui soit durable et qui se maintienne. La diffusion soutenue de messages

nutritionnels, à partir de la mise en place du premier PNNS en 2001, semble cependant avoir amplifié chez les Français un sentiment de culpabilité déjà suscité par le discours médical et la vogue des régimes aminçissants à partir des années 1980-1990 (Fischler, 1993B, « Le bon et le sain » in Autrement). Elle a en tout cas conforté la sensibilité du consommateur quant au lien entre une bonne alimentation et un bon état de santé, malgré une légère tendance à la baisse en 2013 : 87%, contre 89 % en 2007 mais 75 % à la fin des années 1980. Dans la même idée, si en 2007, seuls 69 % des Français considéraient avoir une bonne alimentation, ils sont 72% en 2013. Cette tendance est peut-être le signe d'un retrait de l'angoisse générée par certains discours, du fait d'une inflexion en la matière, notamment depuis la mise en place du PNA. Il est à cet égard significatif que le programme « nutrition-santé » ait été remplacé par un programme « alimentation ». En février 2011 en effet, le Plan National de l'Alimentation (PNA) est mis en place, prenant la succession des PNNS. Ce plan, qui met très nettement en avant le modèle alimentaire, marque un revirement notable : *« Les Français ont toujours eu une relation privilégiée avec leur alimentation. Cette relation a profondément marqué l'histoire, la culture et le mode de vie de notre pays. Manger n'est pas seulement un acte fonctionnel, c'est un plaisir quotidien et partagé qui obéit à certaines règles : la convivialité, la diversité alimentaire, le respect d'horaires fixes et la structuration des repas autour de trois plats principaux. Ces règles définissent un modèle alimentaire français. L'équilibre de ce modèle expliquerait pour partie qu'en France, seulement 14% des adultes ont des problèmes d'obésité contre plus de 30% aux Etats-Unis »*

De même qu'il était périlleux de déduire la montée des représentations caractérisées par les questions de santé et de nutrition aux messages nutritionnels du PNNS dans la période précédant 2007, cela l'est tout autant concernant le revirement de tendance relevé en 2013 dans le sillage du PNA. Encore une fois, sans parler d'un effet de causalité, on peut faire l'hypothèse que cette inflexion n'est pas sans lien avec les nouveaux messages des pouvoirs publics ou ceux relatifs au classement du repas gastronomique français sur la liste du Patrimoine culturel immatériel de l'UNESCO.

CONCLUSION

Cette analyse sur une question posée depuis 25 ans selon la même méthodologie (en face à face) a permis d'une part de mettre en évidence que les outils d'analyse statistiques mis au point au CREDOC au début des années 90 avec le développement du Logiciel ALCESTE de Max Reinert sont les plus adaptés au traitement des questions ouvertes. L'avènement du Big Data ne s'est pas traduit par le développement de nouvelles méthodes d'analyse de données. L'analyse lexicale mise au point par Max Reinert reste la plus stable (comparativement à Cohéris Spad). Cette stabilité est obtenue grâce à l'enchaînement de plusieurs outils d'analyse de données classiques. La méthode ALCESTE commence par une classification descendante, ce qui permet de ne pas inclure les petits noyaux très typés qui provoquent dans une analyse factorielle la création d'axes liés à des artefacts n'ayant pas de sens lorsque l'on travaille sur du sens et des mots. Ensuite le logiciel reclasse les points peu typiques et crée un groupe de non classés. Une analyse factorielle est ensuite réalisée sur l'appartenance aux classes. De nouveaux logiciels sont apparus, très complets comme IraMuteQ basés sur le même algorithme qu'Alceste et plus complets avec l'analyse des similitudes, mais leur manque de convivialité nous a conduits à conserver Alceste pour la classification. L'analyse en évolution des réponses à la question « Pour vous, qu'est-ce que bien manger ? », a combiné plusieurs approches, les comparaisons des mots et des classes entre les différentes années, mais aussi des analyses factorielles classiques. Comme l'ont montré tous les travaux des 25 dernières années sur l'analyse lexicale au CREDOC, ces outils peuvent être appliqués à de très nombreux corpus (presse, textes littéraires, objets complexes, ...) et fondent de nombreux espoirs pour l'avenir du Big Data.

D'autre part, cette analyse est riche d'enseignements sur l'évolution des représentations mentales de l'alimentation. Le lien fait entre l'alimentation et la santé en 2007, dont la progression est considérable, a pu être analysé au moins en deux courants : la recherche de l'équilibre et l'alimentation vue comme un problème. Cette dernière modalité constitue une nouveauté qui vient contredire l'approche traditionnelle de l'alimentation en France (plus familière, en particulier, aux aspects gustatifs et conviviaux). Les Français, qui ont été l'objet d'une diffusion massive de messages nutritionnels induisant une approche plus utilitariste de l'alimentation, à partir du premier PNNS en 2001, se trouvent tirillés entre la tradition culinaire nationale et l'augmentation des pathologies du comportement alimentaire.

Cependant, l'apparition d'une classe du « fait maison » en 2013 est significative d'un retour du plaisir dans l'alimentation, tant du point de vue de la préparation que du point de vue du goût attendu. Elle illustre aussi un déficit de confiance vis-à-vis des produits industriels tels que les produits transformés ou les plats préparés, faisant percevoir le « *fait maison* » comme garantie

d'un « *bien manger* ». La crise de la fraude à la viande de cheval est passée par là et à contribuer à créer un climat de défiance marqué vis-à-vis des produits industriels. Les préoccupations des Français semblent s'éloigner des injonctions nutritionnelles (classe du discours nutritionnel : 9% en 2013) sans cependant renier le souci de manger « *équilibré* » (26%).

Il apparaît aujourd'hui que les injonctions nutritionnelles ne sont pas nécessairement productrices d'améliorations en termes de santé publique. Les nouvelles orientations prises par le PNA, dont nous avons parlé plus haut, témoignent de cette prise de conscience du terreau culturel sur lesquels sont bâtis les modèles alimentaires. L'appel aux caractères essentiels du modèle alimentaire français laisse penser que c'est en s'appuyant sur ceux-ci qu'il devient possible d'introduire des principes de nutrition et d'équilibre, sans quoi tout enracinement est voué à l'échec.

INDEX DES TABLEAUX

Tableau 1 - Fonctionnalités des principaux logiciels d'analyse textuelle français.....	20
Tableau 2 - Bilan lexical par classe du corpus "bien manger" en 2013	37
Tableau 3 - Bilan lexical des corpora des 4 années étudiées	49
Tableau 4 - Les 20 formes graphiques les plus fréquentes dans les corpora analysés	52
Tableau 5 - Gains de rang les plus importants sur le tableau des occurrences entre 1988 et 2013	53
Tableau 6 - Pertes de rang les plus importantes sur le tableau des occurrences entre 1988 et 2013 :	55
Tableau 7 - Fréquence des segments répétés au moins 3 fois dans 3 réponses différentes dans les corpora analysés	56
Tableau 8 - Les 5 réponses les plus citées dans les corporas des 4 années analysées	57

INDEX DES FIGURES

Figure 1 - Courbe de Zipf du corpus "bien manger" 2013.....	29
Figure 2 - Nuage de mots des formes graphiques caractérisant le sexe	31
Figure 3 - Nuage de mots des formes graphiques caractérisant la catégorie d'âge	33
Figure 4 - Nuage de mot des formes graphiques caractérisant la catégorie socio-professionnelle ..	35
Figure 5 - Nuage de mot des formes graphiques caractérisant l'IMC	36
Figure 6 - Répartition des classes lexicales du corpus "bien manger" en 2013 :.....	37
Figure 7 - Nombre et proportion de mots distincts par année analysée.....	50
Figure 8 - Distribution de la longueur des réponses	50
Figure 9 - Vocabulaire commun aux 4 années analysées et vocabulaire spécifique à chaque année (en nombre de mots distincts)	51
Figure 10 - Plan factoriel de l'analyse des correspondances du tableau lexical agrégé par année étudiée	58

BIBLIOGRAPHIE

- BASTIN, G. (2002, Novembre). *Note sur la méthode ALCESTE*. Disponible sur MELISSA (Mettre En Ligne les Sciences Sociales Aujourd'hui) - ENS CACHAN: <http://www.melissa.ens-cachan.fr/spip.php?article200>
- BEAUDOIN, V. (2000). Statistique textuelle : une approche empirique du sens à base d'analyse distributionnelle. *Texte!*
- BEAUDOIN, V., BOISBOUVIER, N., HEBEL, P., LITMAN, S., RACAUD, T., & EVANS, C. (1993). *L'analyse lexicale : outil d'exploration des représentations*. PARIS: Cahiers de recherche du CREDOC.
- BEAUDOIN, V., & HEBEL, P. (1994). *Avancées en analyse lexicale*. Paris: Cahiers de recherche du CREDOC - n° 61.
- BECUE-BERTAUT, M. (2003). Comparaison des structures induites sur un ensemble de réponses ouvertes par le choix de l'unité statistique. *Corpus*.
- BECUE-BERTAUT, M. (2003). Comparaison des structures induites sur un ensemble de réponses ouvertes par le choix de l'unité statistique. *Corpus [en ligne]*, <http://corpus.revues.org/28>.
- BECUE-BERTAUT, M., & PAGES, J. (2000). Analyse factorielle multiple intra-tableaux. Application à l'analyse simultanée de plusieurs questions ouvertes. *JADT 2000 : 5èmes journées internationales d'analyse statistique des données textuelles*. Lausanne.
- BENZECRI, J.-P. (1973). *L'analyse des données 2 : L'analyse des correspondances*. Paris: Dunod.
- BENZECRI, J.-P. (1981). *Pratique de l'analyse des données : linguistique et lexicologie*. Paris: Dunod.
- BOUCHET-VALAT, M., & BASTIN, G. (2013). RcmdrPlugin.temis, a Graphical Integrated Text Mining Solution in R. *The R Journal*, pp.188-196.
- BRIAN E. (1983). Le système d'enquêtes sur les Conditions de Vie et Aspirations des Français 1978-1984, phase IV. Pratiques et perceptions des transports. Rapport CREDOC, n°4865, Vol III.
- BRIAN E. (1984). Analyse des données lexicométriques. Elaboration des programmes. Rapport CREDOC, n°4909.
- BRIAN E. (1986). Techniques d'estimation et méthodes factorielles, exposé formel et application aux traitements de données lexicométriques. Thèse de Docteur Ingénieur, Orsay.

- BROUSSEAU, A.-D. (1999). *Impact de la lemmatisation sur la robustesse des typologies lexicales - Rôle des seuils de sélection de mots analysés*. Paris: Cahiers de recherche du CREDOC n°131.
- BRUGIDOU, M., ESCOFFIER, C., FOLCH, H., LAHLOU, S., LE ROUX, D., MORIN-ANDREANI, P., & PIAT, G. (2000). Les facteurs de choix et d'utilisation de logiciels d'Analyse de Données Textuelles. *5èmes Journées internationales d'Analyse Statistique des Données Textuelles*. Lausanne.
- BRUNET, E. (2000). Qui lemmatise dilemme attise. *11èmes rencontres linguistiques en pays rhénan* (pp. pp. 7-32). Strasbourg: Scolia n°13.
- COLLERIE DE BORELY, A. 1998. Etude des réseaux de mots – Confrontation des résultats issus de deux méthodes d'analyse textuelle. Alceste et Leximappe. *Cahier de Recherche CREDOC*, n°115.
- DESJEUX D., 2006. La consommation. Que sais-je ? Presses Universitaires de France
- DEGENNE, A., & VERGES, P. (1973). Introduction à l'analyse des similitudes. *Revue française de sociologie*, pp. 471-512.
- FALLERY, B., & RODHAIN, F. (2007). Quatre approches pour l'analyse des données textuelles : lexicale, linguistique, cognitive, thématique. *XVIème Conférence Internationale de Management Stratégique*. Montréal.
- FEINERER, I., HORNIK, K., & MEYER, D. (March 2008). Text Mining infrastructure in R. *Journal of statistical software*, Volume 25, Issue 5.
- FISCHLER, C. (1990). *L'Homnivore*. Paris: Odile Jacob.
- FISCHLER, C. (1993), « Le Bon et le Sain(t) », Paris, Autrement, série « Mutations », n° 138, pp. 111-124.
- GUERIN-PACE, F., & GARNIER, B. (2010). *Appliquer les méthodes de la statistique textuelle*. CEPED.
- HEBEL, P. (2008). Le retour du plaisir de cuisiner. *Newsletter du CREDOC - Consommation et modes de vie n° 217*.
- HEBEL, P., SIOUNANDAN, N., & LEHUEDE, F. (2009). *Le consommateur va-t'il changer durablement de comportement avec la crise ?* Paris: Cahier de recherche du CREDO n°268.
- HEBEL P. ET PILORIN T. 2012. Comment les discours nutritionnels influencent-ils les représentations de l'alimentation ? *Cahiers de nutrition et diététique* (2012) 47, 42–49.

HUBERT, A. (2006). « Nourritures du corps, nourritures de l'âme. Emotions, représentations, exploitations ». Le mangeur-ocha.com. Disponible sur :

<http://www.mangerbouger.fr/IMG/pdf/EtudeMessagesSanitaires-2.pdf>

JENNY, J. (1999). Pour engager un débat avec Max Reinert à propos des fondements théoriques et des présupposés des logiciels d'analyse textuelle. *Langage et société - Volume 90*, pp.73-85.

JODELET, D. (1997). Les représentations sociales, Paris : PUF.

LAHLOU, S. (1992). *Si/Alors : "Bien manger" ? - Application d'une nouvelle méthode d'analyse des représentations sociales à un corpus constitué des associations libres de 2000 individus*. Paris: Cahier de Recherche - CREDOC.

LAHLOU, S. (1995). *Penser manger - les représentations sociales de l'alimentation*. Thèse de doctorat en psychologie sociale.

LEBART, L. (2001). Traitement statistique des questions ouvertes : quelques pistes de recherche. *Journal de la société française de statistique, tome 142, n°4*, pp. 7-20.

LEBART, L. (2004). Validité des visualisations de données textuelles. *JADT 2004 : 7èmes journées internationales d'Analyse statistique des données textuelles*. Louvain: Presses Universitaires de Louvain (PUL).

LEBART, L., & SALEM, A. (1994). *Statistique textuelle*. Paris : Dunod.

LUNG, F., & VENDASSI, P. (2012). *Diagnostic et évaluation : la boîte à outil du sociologue*. Paris: L'Harmattan.

LUONG, X. (1988). *Méthodes d'analyse arborée. Algorithmes. applications*. Thèse de doctorat - Paris V.

MATHE, T., PILORIN, T., & HEBEL, P. (2008). *Du discours nutritionnel aux représentations de l'alimentation*. Paris: Cahier de recherche du CREDOC.

MORINEAU, A. (2012, Janvier 4). *Histoire du logiciel SPAD*. Disponible sur deenov: <http://www.deenov.com/blog-deenov/histoire-du-logiciel-spad.aspx>

PNNS - Messages sanitaires sur les publicités. (2014, juillet). Disponible sur www.mangerbouger.fr: <http://www.mangerbouger.fr/bien-manger/infos-nutritionnelles/messages-sanitaires-sur-les-publicites.html>

REINERT, M. (1983). Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte. *Cahiers de l'analyse des données*, pp. 187-198.

- REINERT, M. (1999). Quelques interrogations à propos de l' "objet" d'une analyse de discours de type statistique et de la éponse "Alceste". *Langage et société - Volume 90*, pp.57-70.
- REINERT, M. (2000). La tresse du sens et la méthode Alceste : Applications aux rêveries du promeneur solitaire. *JADT 2000 : 5èmes journées internationale de l'analyse des données textuelles*. Lausanne.
- REINERT, M. (2001). Approche statistique et problème du sens dans une enquête ouverte. *Journal de la société de statistique de Paris, tome 142, n°4*, pp. 59-71.
- ROUANET, H., & LEPINE, D. (1976). A propos de "l'analyse des données" selon Benzecri : Présentation et commentaires. *L'année psychologique*, vol. 76, n°1. pp.133-144.
- SUIGNARD, P., KERROUA, S., PERADOTTO, A., LAGARDE, D., GUENET, A.-L., GUILLOT, L., . . .
DESPLAS, J. (2014). Correction orthographique de réclamations clients. *12èmes journées internationales de l'analyse des données textuelles*. Paris.