A method to extract social representations from linguistic corpora

2nd International Conference on Social representations. Rio de janeiro, 1994.

(1996) - LAHLOU, Saadi. A method to extract social representations from linguistic corpora. *Japanese Journal of Experimental Social Psychology*. 1996, 35, 3, pp. 278-291.

A method to extract social representations from linguistic corpus

Saadi Lahlou.

EDF-Direction des Etudes et Recherches ; GRETS

and Laboratoire de Psychologie Sociale, EHESS. Paris

Tel : (33-1) 47 65 52 21. E.mail :Saadi LAHLOU@grets.der.edf.fr

Summary : A new method is presented, which enables extracting the pattern of social representations of an object from corpora in natural language "about" this object.

Interrogation of a source of common knowledge (a representative sample of a population, a dictionary, a set of articles or books), yields a corpus of linguistic statements concerning the object. In the case of individuals, an open question on free association (What comes comes to your mind about...?). In the case of dictionaries, the set of all definitions of synonyms and analogues of the word in question is used.

The corpus is then processed with a software that breaks up the corpus into statements (e. g. : sentences), and then makes a classification of those statements, on the basis of co-occurrence of lexical traits. Each class is considered as a basic nucleus of the representation, characterised by typical lexical traits. Multivariate analysis enables to represent the relationship of those nuclei and traits in a semantic space of connotations.

Demonstration of the method is presented on two corpuses about "eating", (1) coming from a survey using the free association technique on a 2000 sample representative of the French population, (2) 544 definitions of synonyms and analogues of "to eat" from a large dictionary. Results are quite similar on both corpuses ; they yield a very clear model of the social representation of "eating", which is coherent with the findings by qualitative methods in the literature.

An epidemiological approach of representations

Communication and life in common imply that members of a society share a common, and pragmatic, view of the objects in the world they live in. Such common mental constructs are called "social representations" (Moscovici, 1961). They are "a kind of knowledge, socially constructed and shared, having pragmatic purpose and contributing to build a common reality for the community"(Jodelet, 1989).

Social representations can be figured as a combination of basic nuclei -that is some basic concepts or traits, that are mentally associated in the subject's minds (Abric 1984, 1993, Flament 1993). This model has been supported with some empirical evidence (Vergès, 1994 ; Guimelli et Rouquette, 1992). We adopt here an epidemiological view (Sperber, 1989), where a social representation can be considered as a cultural specie, a population of individual representatives (mental representations) widespread in the minds of human subject ; just like the specie of, say, "butterflies" is a population of individual butterflies scattered about the ecosystem. In this ecological perspective, characteristics of the social representation can be inferred from observation of a sample of individual representatives.

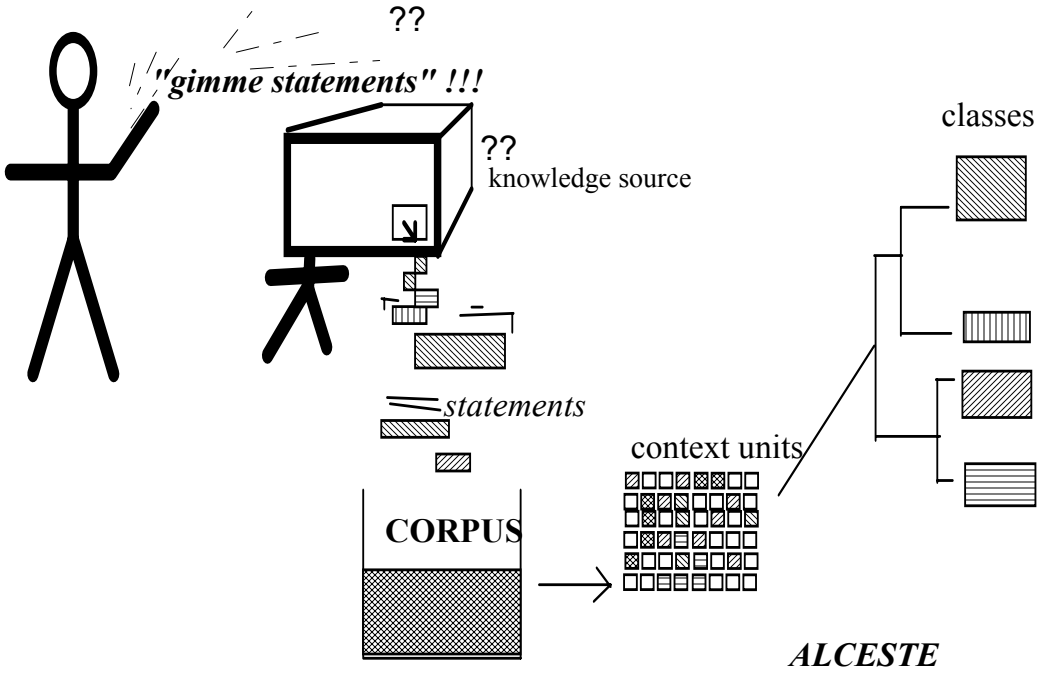
Getting a clear view of social representations is difficult, since they are mental constructs and can only be inferred from observation and declarations of the subjects. Classical methods soon reach limits in the investigation of social representation when the basic nuclei are not previously known, as investigation is biased by the researcher's preconceptions. E.g., to go on with our metaphor : should one include the "chrysalides" in the description of "butterflies" ? In the study of representations, qualitative approaches yield detailed information, but no quantified data, which makes it difficult to distinguish the basic nuclei from less central associations, that would be common only to the subjects interviewed. On the other hand, questionnaire surveys produce biases, and they can never assure that the questions map *all* relevant dimensions of the representation. Behavioural observations are a must, but they are usually so technically heavy that they are not used in the first place (Wagner, 1994).

We propose a new method, based on natural language analysis, using the "free association" technique, and allowing investigations at a population's level. Although it does not solve all the problems, it seems to be an easy first approach to get quickly the main dimensions of social representations, approach that can then be completed then with more classical methods. An interesting feature of the method is that it can also be applied quite cheaply on non-human cultural sources (e.g. dictionaries), sparing hereby the burden of interviewing thousands of subjects with results nearly equivalent.

This method will be demonstrated through empirical investigation about "eating", on two corpuses obtained from sources of "common knowledge" : a sample of 2000 French adults, and a dictionary. Both yield the same model of "eating" in the French population.

The method : statistical analysis of lexical data

The method aims at spotting the basic nuclei of social representation, in the discourse produced by an informed source. First, a corpus of statements (sentences in natural language), about the same object (here : "eating") is obtained through the free association technique (Freud, 1923 ; Jung, 1933). The sentences obtained are aggregated into a single corpus. The corpus is then processed with a statistical analysis of lexical data software, ALCESTE (Reinert, 1983, 1990), yielding classes of statements that have similar lexical content.



Those semantic classes are considered to be the basic nuclei of the social representation, and the global pattern linking them will be said to be the "paradigm" underlying the social representation of "eating" (Beaudouin & Lahlou 1993). This method is a kind of quantified, detailed, content analysis. As it is computerised, it can be applied to very large sets of statements, far beyond the capacity of human analysis ; it is also free of preconceptions, since there is no interpretation by the software, but mere computing of lexical co-occurrence.

Classes are built on the principle of putting together statements that are close to each other within a cluster (analogy), and different from statements in other clusters (contrast). This is a principle similar to human natural categorisation and pattern recognition, as described by Rosch (1975) and the Gestatlttheorie. Analogy and contrast are made on the bases of significant traits. In our case, the traits will be lexical (words, lexical roots), and the mathematical technique used is descending classification ("segmentation") on the basis of a table crossing sentences and traits. Algorithms used are derived from multivariate analysis techniques (principal components analysis, dynamic clouds).

For instance,

des crudités de la **viande** et des **légumes** de la salade et un **fruit**

will be close to :

la **viande** les **légumes** dessert les **fruits** le pain le fromage

because they have in common the lexical roots **viand+**, **legum+**, **fruit+**,

and far from

aller au restaurant avec des copains

because they have no lexical trait in common. The technique enables taking account of indirect links between sentences, yielding that :

(1) des crudités de la viande et des légumes de la salade et un fruit

is classed with :

(4) repas copieux avec entrée plat_résistance et dessert

although they have no lexical trait in common, but their association is indirect through sentences like :

(2) a un bon *repas entrée plat_résistance* fromage **salade** *dessert* café

(3) la **viande** les **légumes** *dessert* les **fruits** le pain le fromage

those two sentences being close of the former ones, made close to each other by the co-occurrence of ***fromage*** and ***dessert***, and the fact that, as a cluster, the result has a maximum contrast with other clusters.

Although the software is deaf to meaning, the results make sense, as one can easily judge since the translation in English is :

(1) fresh vegetable meat and greens salad and a fruit

(2)a good meal first course main course green salad dessert coffee

(3) meat vegetables dessert fruits bread cheese

(4) abundant meal with first course main course and dessert

The material : texts obtained through the free association method

Common sense can be expressed by different "minds". The most often interviewed are the native ("man of the street"), and the "expert". But non human sources can also be consulted with profit, for instance the dictionary, precisely designed to express common sense into public knowledge. We shall apply our technique to representative sample (2000 subjects) of the French population (Lahlou, 1992, 1993), and to a famous French dictionary, Le Grand

Robert (9 volumes, 100 000 lexical entries) (Lahlou, 1994a & b). The principle of the method is identical for both sources : the source is stimulated with the name of the object, and asked to produce statements relating to that object.

Human subjects have been asked[1] the open-ended question : “Si je vous dis *manger*, quels sont les cinq premiers mots qui vous viennent à l'esprit?” (if I tell you “eating”, what are the five first words that come to your mind ?). Answers yield the corpus "Natives". Here are some answers.

```
* faim, nourriture, joie (hunger, food, happiness)
* cuisiner (cooking)
* fourchette casserole (fork, pan)
* crabe homard poisson (crab, lobster, fish)
* j'ai faim (I'm hungry)
*  viande,  légumes,  régime,  diététique/etc.  (meat,  vegetables,
dietetics, etc.)
* à grossir, manger=grossir pour moi, j'ai pris 10 kilos (to fatten,
to eat = to fatten for me, I've taken 10 kilos)
* nourriture restaurant sortir. (food, restaurant, going out)
```

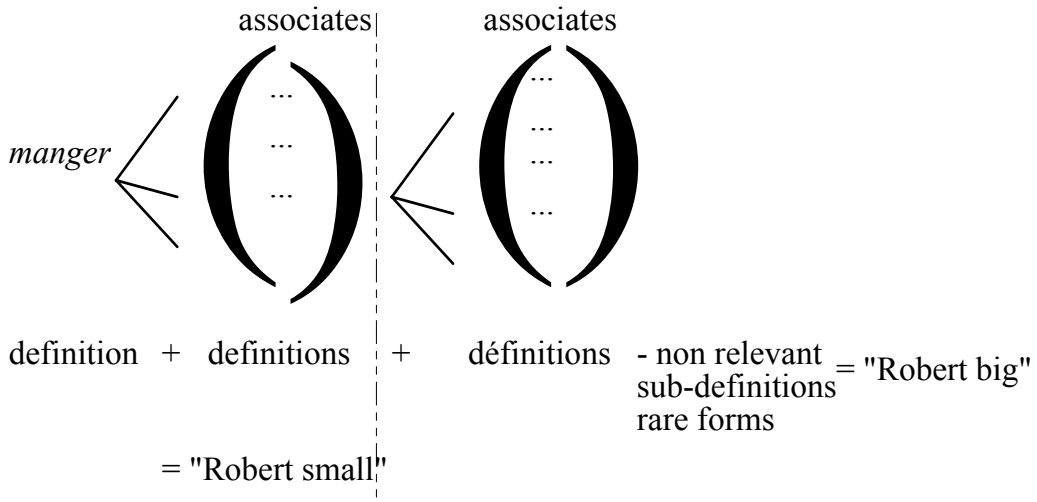
This corpus contains 12,000 occurrences, with a vocabulary of 3,309, reduced to 308 distinct roots after eliminating roots with few (<4) occurrences, and "tool words" (articles, prepositions...).

The dictionary "Le Grand Robert" (1991, electronic version) also yields analogous indications. E. g. for the entry "manger" (to eat), it features a list of more than 140 synonyms, analogues, contraries and derived terms. A corpus ("Robert small") containing the 140 definitions of first level associates was constituted.

In turn, each of those 140 associates has associates itself. We made a list of those associates of first and second level, that includes 544 words ; and collected all the corresponding definitions in a single corpus[2], called "Robert big". This corpus contains 137,576 occurrences, with a vocabulary of 16,896.

[1] In Spring 1990, Credoc’s periodic survey "Aspirations", 2000 subjects representative of the adult French population (face to face interviews, at home, by professional interviewers).

[2]After elimination of some rare words and sub-definitions that wore not relevant (e. g. integration in the mathematical sense).



Here is, for example, a (short) definition of a verb, that has 3 sub-meanings (to make swallow, to gulp down, to learn massively):

> ***INGURGITER*** *v. tr.*
> *(Sujet nom de personne).*
> *_ 1. Rare. [a] Introduire dans la gorge, faire avaler (qqch.) à (qqn). > Enfourner, entonner. Ingurgiter qqch. à qqn. La potion qu'on lui a ingurgitée. S'ingurgiter qqch., l'absorber. «En s'ingurgitant un sévère apéritif» (Huysmans, à rebours).*
> *[b] Ingurgiter ses leçons à un élève. - S'ingurgiter un énorme traité.*
> *_ 2. (1840). Avaler avidement et en quantité (qqch. : aliment, boisson). > 1. Boire, déglutir, engouffrer. Faire ingurgiter qqch. à qqn.*
> *_ 3. (1856). Absorber massivement un savoir sans pouvoir l'assimiler.*
> *> Apprendre.*

EATING as evoked by the dictionary

The corpus "Robert small" was broken down into 1220 statements. Tool words (articles, prepositions etc.) were not retained, and the words have been lemmatised (reduction to lexical roots : verb to the infinite, adjectives to singular etc.). Finally 877 lexical traits (the most frequent) were retained for analysis.

Besides *eat* (*manger* 443 occurrences) the most frequent traits are *take* (*prendre,* 662), *do* (*faire,* 258)*, someone(-body) (quelqu'un, 243), table (243), foodstuff (aliment, 190), mouth (bouche, 184), meal (repas, 166), feed (nourrir), touch (toucher), attack (attaquer), food (nourriture), something (quelque chose), mouth (gueule), swallow (avaler)* and *taste* (*goût).*

Analysis yields four classes :

TAKING (*PRENDRE)*

This class (the largest) is mainly characterised by verbal traits, which is remarkable since the software is deaf to grammatical classes when it classifies :

take, touch, attack, smbdy, smthg, cut, hand, subject, burn, gnaw, contact, weapon, consider, strong, pounce/melt, crunch, attain, opponent, enemy, catch, fire, hit, destroy, pull, pinch, act, fight[3].

This class of intake and appropriation is heavy with violent, agonistic, aggressive connotations.

FOOD *(NOURRITURE) :*

Typical traits of this class are foodstuffs, or food categories.

food, feed, bread, nourishment, diet, live, fasting, sustenance, plant, necessary, supply, deprivation, milk, dietetic, (starvation) diet, element, meat, nutri (-tion, -ent), breast, sustain, sour, sweet, child, cake, sap, starve, vegetable, shortage, product, water, beverage, fresh, animal, fruit, lean, flesh, breed, idea, god, egg[4].

One notes the presence of *breast* and *milk*, among other foods.

MEAL (REPAS)

Typical traits are nouns of meal objects other than foods, relating to the people present, or the table arts, timing, occasion. This class is centred on ritual, instrumental and social aspects of the intake. It gives precisions the circumstances and means of the intake.

meal, table, dinner, service, crockery, soup, buffet, invite, serve, lunch, dish, restaurant, dessert, host, trait, guest, canteen, snack, fork and spoon, hotel, together, utensil, plate, wash, breakfast, treat, region, evening, hour, order, room[5].

---

[3] prendre, touch+, attaqu+, qqn, qqch, entam+, main+, sujet, consum+, ronge+, contact+, arme+, consider+, fort+, fondre., croqu+, atteindre, adversaire, ennemi, attrape, feu, coup, detruire., tirer., pince+, agir., combat

[4] aliment+, nourr+, pain+, nourrit+, regime, vivre., jeun+, subsistance, vegetal+, necessaire+, fourn+, privation+, lait, dietetique+, diete, element+, viand+, nutri+, sein, sustente+, ame+, sucre+, enfant+, gateau+, seve+, affame+, legum+, liquid+, manque+, produit+, eau, boisson+, frais, animal<, fruit+, maigr+, chair+, eleve+, idee+, dieu, oeuf

[5] repas+, table+, dine+, service+, vaisselle+, soup+, buffet+, invite+, servir., dejeune+, plat+, restaur+, dessert+, hote, trait+, convive+, cantine, collation+, couvert<, hotel+, ensemble, ustensiles, assiette+, laver., petit-dejeuner, regal+, region<, soir+, heure<, ordonn+, salle

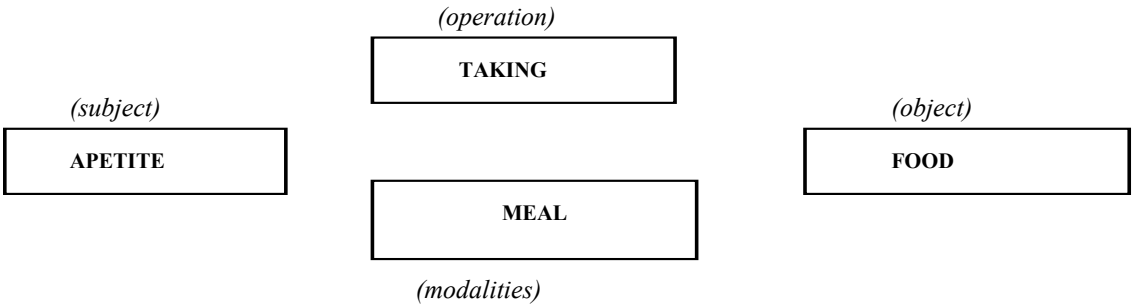
APPETITE (APPETIT)

The traits obviously characterise the desire for food, and more generally libido.

glutton, appetite, eat, greedy, guzzle, gluttonous, gorge, stuff, voracious, excess, ogre, avid, fulfil, cram, hunger, appetite, suck, carnivorous, eye, avidity, satisfaction, wolf, desire[6].

Desire is expressed with a connotation of greediness, voraciousness that is typical of intense desire. This violence, slightly surprising in such a corpus, is not an artefact. Further analyses on the larger corpus confirm and enhance this aspect.

This classification can be considered as the *central paradigm* of the social representation of eating, consisting of four basic nuclei.

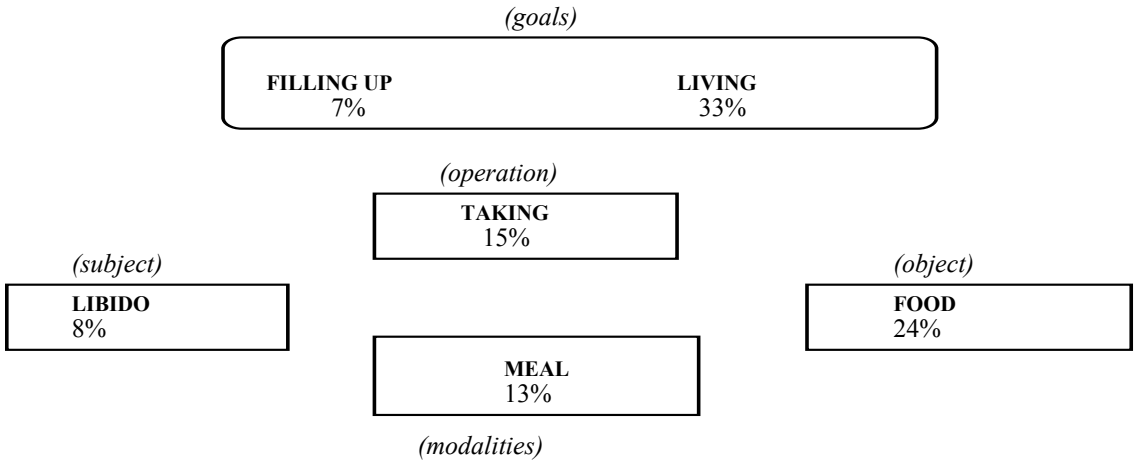


Further analysis on the larger corpus "Robert big" yields two extra classes of supra ordinate goals (FILLING UP, LIVING) that focus on the social, moral aspects (keeping balance, sharing...). These classes show how the representation of eating is connected with a more general world of the subject, his goals and his conception of life.

This analysis also yields deeper and purer concepts for the previous nuclei : the "appetite" class looses its feeding connotations (which rally the "FOOD" class) and becomes a more general hunger, with sexual and cognitive connotations (e. g. "sexual", "curiosity", ...) which makes us identify it with a basic "LIBIDO" drive.

This representation of eating can be pictured as the following combination of basic nuclei :

[6]glouton+, appet+, mange+, gourmand+, goinfr+, goulu+, rassasi+, empiffr+, vorac+, exces, ogre+, avide+, assouvi+, bafre+, faim+, appetits, aspir+, carnassier+, yeux, avidite, contente+, loup+, desir+

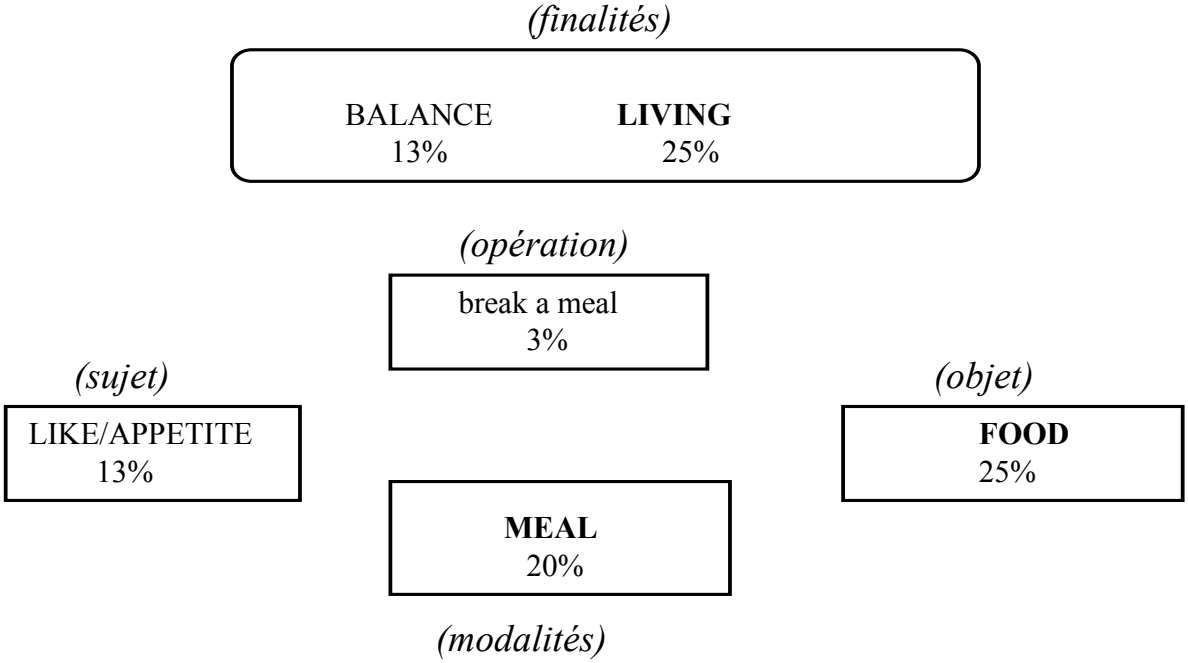


Therefore, dictionary analysis with our method painlessly yields *knowledge modelling* concerning our object (eating). The model is explicit and simple, as could be expected from a source that has didactic intentions, like the dictionary.

This blind method gives results that are very stable to parameter modifications in the analysis, analysis of sub-corpuses, pre-treatment by reduction through syntactic analysis, and/or semantic networks (Constant, 1991, Pigamo, 1990).

The human subjects

The analysis of free associations of "eating" (2000 subjects, Spring 1991) yields results quite similar to the dictionary, although a bit poorer.



The most significant difference is that the central class "TAKE" has almost totally disappeared, leaving a faint notion of *breaking the fast*, or *taking a meal*. This can be understood since, as the survey is a face to face dialogue, this trait can have been omitted by interviewees because it is too obvious (it would indeed sound a bit stupid to say that eating evokes *taking* food) This may explain why in other protocols, researchers do not get all the nuclei from the subjects, and are therefore obliged to design method where the "operators" are explicitly asked to subjects (as in the SCB model, by Guimelli & Rouquette, 1992)

This implicitation can also be interpreted, psychoanalytically, at the light of the fact that it is precisely this nucleus that has the most animal/primitive/aggressive connotation, and would therefore be rejected into the unconscious.

Whatever, this shows that : (1) the method is reliable enough to give similar results in two very different sources of the supposed same representation, (2) that the dictionary is not only less expensive to interview but also more precise because more structured and didactic, and has less inhibitions to talk than the native.

This structure is interesting in itself, since it features clearly the basic nuclei of the representation of eating for French people.

The analysis of the dictionary also allows some interesting archaeological findings. Language is a kind of collective memory ; in the networks of explicit definitions seem to lie some ancient meanings, that evoke Jung's "collective unconscious", and might enlighten some of the magical beliefs that can be found in our cultures.

We believe the investigation technique we propose is valuable for research purposes, and for more pragmatic purposes (e.g. : marketing).

The ontogenesis of social representations

Comparison of our results with the sociological and psychological literature on food shows many findings coherent with previous theories (Lahlou, 1994b), and enlightens some aspects of the now classical "incorporation paradigm" (Frazer, 1890 ; Rozin, 1988, 1990 ; Fischler 1990 ; Rozin, Millman & Nemeroff, 1986). This principle, a cultural universal, says, roughly, that people believe *they become what they eat*.). We shall discuss here only some results showing how the method can give clues on the genesis of social representation.

As we saw, the central nucleus TAKING has some aggressive connotations. This feature is coherent with the findings of Freud (1921), tracing the identification principle back to the oral phase of the organisation of libido, in which one incorporates, by the act of eating, the

objects desired and liked, and in doing so, one destroyed it as such. As described in the classical Vocabulary of Psychoanalysis, by Laplanche and Pontalis :.

> "In fact, three meanings are present in incorporation : to take pleasure in having an object penetrating oneself, destroying this object, assimilating the properties of this object by keeping it inside oneself." (Laplanche & Pontalis, 1967, 1990 p. 200) (italics mine)

Our material brings evidence to this interpretation since the action aggressive connotations of the TAKING, logically unexpected, are typical of the subject/object ambivalence in the oral and pre-objectal phases of development. Also coherent with Freud's views is the fact that the FOOD nucleus contains traits typical of the breast feeding relationship, as well as some religious connotations.

If we now look at the general structure of the representation, is interesting to note that those elements constitute the bases of a typical causal script, with a *agent*, an *action*, and a *patient*. This classical causal script is here completed with goals, and means. In fact, the scheme obtained is nothing else than the schematization of the behavioural process of eating. This is coherent with the old idea that representations are learned from experience. In this perspective, the structure found is roughly a translation of behaviour (as subjectively perceived, including emotions) into "ideas". The structure emerging through analysis can be considered as a record of previous behaviours, with sedimentation starting from infancy, and probably, through cultural and genetic transmission, way back into our hunter-gatherer's historical past. This brings empirical support to the view that representations have their origin in biological functions, as expressed by Moscovici (1974), or Spitz : .

> "the transition from somatic to psychological is a continuous one, and therefore the prototypes of psychic nuclei must be searched for in physiological functions and somatic behaviour" (Spitz, 1965, 1968, pp. 78-79)

In this perspective natural language could therefore be considered as a cultural sediment, keeping vivid traces of former cultural and self development, a kind of collective memory, where one could find traces of mankind's past and childhood, just as we can see in Le Grand Robert, in watermark, Man-the-hunter with his predator instinct, and Man-the-infant with his oral drive. This is coherent with the old idea that language is at the same time a vehicle and an image of the culture (Durkheim, 1895, Whorf, 1956).

Pragmatic aspects of the representation

Our method allows further investigations of representations, that gives a clearer view of how the subjects use them as behavioural guidelines. We shall now analyse a corpus of answers to the question "If I tell you eating well, what comes to your mind ?" (Spring 1990, 2000 ,

representative of French adults). Because of the adjective "well" this question induces answers on what *is to be done* by the subject. Free associations are then oriented towards behaviour, enlightening the pragmatic aspect of social representations. This aspect is a most prominent function of social representations, which are a picture of the world intended to serve the pragmatic goals of individuals in their everyday life (Moscovici, 1961).

After cleaning up (suppression of *don't know*, *that's all*, etc.), reduction of locutions, composed nouns, and excluding articles, the corpus contains 13,504 useful words, with a vocabulary of 1,408 different graphic forms, of which 999 appear 3 times or less. The mean per sentence is 6.75 useful words.

Most frequent traits are :

eat (manger, 871), good (bon 716), meal (repas 453), balance (équilibre 407), not (pas 400), I (je 290), it is (c'est 219), not (ne 218), well (bien 212) , in by (en 181), one, we (on 176), healthy (sain 166), do (faire 165); that (que 138), meat (viande 134), restaurant, with (avec), vegetable (légume), thing (chose), for (pour), too much (trop), self (se), hunger (faim).

The most frequent word chains are *good meal* (bon repas) and *eat balanced* (manger équilibré), quite frequent, but also many negative chains : *not too much, no more, not always no excess*, that most often limit : *fat, heavy, rich foods, sugar,* and *meat*.

Superficial analysis confirms that we are on the same grounds of the central paradigm, but that sentences have a prescriptive connotation.

Global analysis yields 7 classes :

To eat one's fill (10-14%)

The first is based on the repletion feeling : to eat normally, but not too much, just enough to satisfy the hunger. Some answers :

To eat correctly, to eat one's fill (manger correctement manger à sa faim)

To fill up one's stomach (à se remplir le ventre)

to eat normally, to feed myself just remaining with a small hunger (me nourrir normalement, me nourrir en restant avec une petite faim)

etc.

This is the representation underlying :

-Have some *more* of this pie ?

-*No* thank you, it is delicious, but it wouldn't be *reasonable*.

Eat what I like (14-17%)

This class is centred on pleasure, and desire is the criterion of choice.

To eat what I fancy, what I take pleasure in (manger ce qui me plaît ce-que j'ai le plaisir à manger)

It's great, there's the pleasure of savouring good things (c'est formidable il-y-a le plaisir de déguster des bonnes choses)

Etc.

Balanced and healthy (16-20%)

This class mixes the strive for natural products, and a balance of nutrients, it is typically a dietician's discourse.

To eat balanced, without chemicals or preservers (manger équilibré, sans produit chimique ni conservateur)

To eat healthy, varied food s and fresh products (manger sainement, une nourriture variée et produits frais)

Food balance, protein,; lipids, glucids, vitamins, minerals (l'équilibre alimentaire protéines lipides glucides, vitamines sels minéraux)

Etc.

This is more frequent among women, and upper middle class. *Glucids, proteins, lipids, vitamins, energy, fresh, natural* appear frequently, and also the refusal of : *chemicals, canned products, frozen foods, artificial colouring.*

Not too much fats or sugar (11-13%)

This class is of forbidding and restriction, against fats and sugar , and excess in general.

Not too fat not too much sugar and not too many calories (pas trop gras pas trop de sucre et pas trop de calories)

But not eat too heavy (mais pas manger trop lourd)

Diet cholesterol too much sugar (régime cholesterol trop de sucre)

etc.

This is more frequent among elder people

Foodstuffs and meal course (15-19%)

This class, very stereotyped, describes the traditional sequence of a French meal course.

A good meal : first course, main course, cheese salad dessert coffee (un bon repas : entrée plat de résistance fromage salade dessert café)

Meat vegetables, bread and a glass of wine ! and cheese (viandes fruits légumes, pain et un coup de rouge! et du fromage)

Etc.

Traditional home-made dishes (4-6%)

This class insists on family aspects of food ; tradition and home making are the marks of a "roots" cuisine with strong affective tone.

A good grub French style home made cooked dishes (une bonne bouffe à la française plats cuisinés maison)

Fine-cooked dishes pastry (des bons petits plats pâtisserie)

To have a good meal with my mom, the good old family cooking of my childhood (faire un bon repas avec maman, la bonne cuisine familiale de mon enfance)

A family table, happiness harmony and good cooking (une table familiale gaie présence d'harmonie et de bonne cuisine)

etc.

Conviviality (12-14%)

This class evokes feast, banquets, and friends : the social aspect is central.

To have a super good meal with friends (faire un super bon repas avec des amis)

A good grub with pals, like a country barbecue (a une bonne bouffe avec des copains style barbecue a la campagne)

Feast, good wine (bon gueuleton, bon vin)

etc.

This is typical of young men, inhabitants of mid-size towns, factory workers, employees.

Restaurant (6-8%)

This is a gastronomic class in the French sense of the term, with decorum, table arts, of which the restaurant is typical as the place where the art of eating takes place as a ritual.

Gastronomic restaurant (restaurant gastronomique)

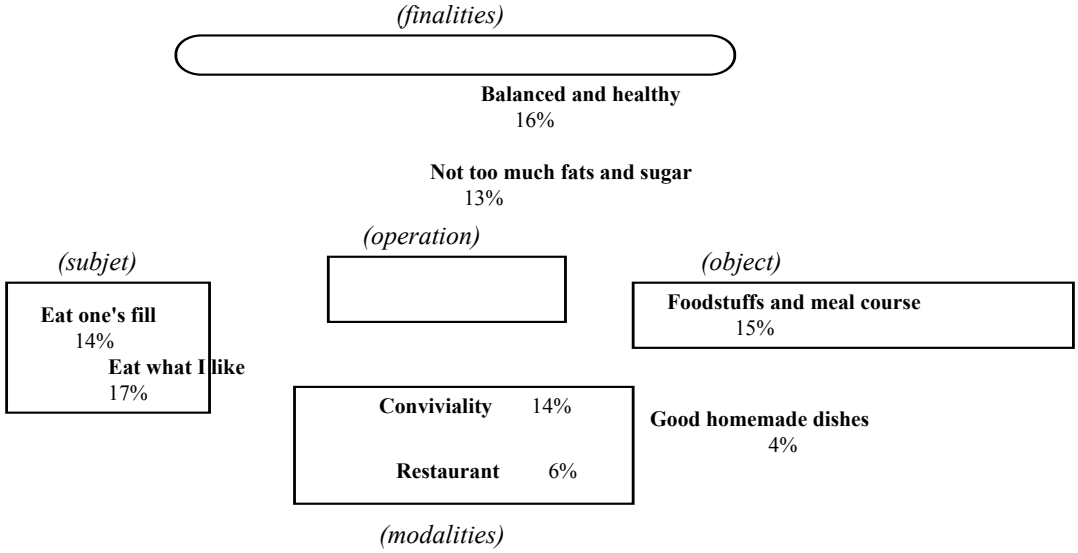
Go to a beautiful restaurant with refined menus, pâté (aller dans un beau restaurant avec des menus raffinés du foie_gras)

Go out at the restaurant with my wife (faire une sortie au restaurant avec ma femme)

Etc.

This is typical of young people, inhabitants of big towns, men, unmarried couples, higher education, upper classes

Factor analysis of the corpus yields a connotation space where two words are all the more close than they co-occur in the sentences. For instance, *restaurant* is very close of *gastronomy*, and *cheese* of *dessert*. In this space the classes are organised around 4 poles. The first relates to the body's voice (***Eat one's fill***, ***Eat what I like***), the second to the respect of some ethics -in fact, dietetics- (***Balanced and healthy, Not too much fats and sugar***) ; the third to social aspects (***Conviviality***, ***Home-made***, ***Restaurant***), and the last, which is also the most stable and structured, pragmatic (***Foodstuffs and meal course***). A mapping is given in annex, and detailed results in (Lahlou, 1992). It is interesting to replace those classes in the frame of the central paradigm.



As one can see, the classes don't fit well in the frames. Indeed, ***Eat one's fill*** and ***Eat what I like*** are centred on the subject. But the first class is deported to the pragmatic aspect, as it contains as well motivation and sensation as a stop condition. The second is clearly deported towards the object, it mentions not only the desire, but also the foods, either as a general

token (*something I like, things that*...) but sometimes explicitly (*chocolate, wine, sweets*...). And both classes are rather verbal, the answers often contain the trait *eat*.

In the modalities, if the classes ***Conviviality*** and ***Restaurant*** are well in the frame, the class ***Home-made***, even if the connotations and many traits clearly place it near the two former, contains explicit references to foodstuffs (the *traditional dishes*).

The classes ***Balanced and healthy***, and ***Not too much...*** have ethical and even moral connotations. but the balanced class is expressed in terms of object characteristics as well as in terms of principles, and the class not too much contains a reference to action (avoidance) and to objects named by their characteristics (*fat, sweet*) or even by names (*fats, alcohol, canned foods*).

Finally, the ***Meal course*** class, although it is purely objectal in trait contents, presents a systematic order that can only be qualified as actional, and prescribes the sequence in which foods must be eaten.

The central frame, TAKE, remains empty. The operator has been totally made implicit, it has been absorbed in the other classes.

Most of those effects can be explained by some classical aspects of the representational scripts theory ; they basically derive from the fact that scripts are simple, pragmatic, goal oriented schemata that people use to guide their behaviour.

First, representational scripts are concrete and synthetic : they articulate the basic nuclei of representation into simple action rules. All scripts are the combinations of a couple of basic nuclei, usually including the operator TAKE, so that the script becomes an autonomous pragmatic prescription with it own "runtime" motor. In this respect, representations are indeed are at once structure and process : structures that articulate in the form of process. What we see her is the same representation as supra, but structured into behavioural prescriptions by articulation of the basic nuclei.

Second, they are simple and unambiguous.

Third, they are systematically focused on the conclusion (finality of the action), compared to the basic representation. This effect is easily readable on the figure through the systematic deportation of the classes to the right side of the frame.

The central nucleus of the paradigm, he operator TAKE, because it is obvious in an everyday life context, is implicit in the subjects' answers (an effect that was predictable). It has been incorporated in each script so that each script is a self sufficient "runtime" rule with its own action motor.

Finally, let us mention that the subjects did not all say the same things. Statistical analysis shows some traits are typical of some categories. For instance, men rather talk about : *bon(good), foie gras (pâté), qualité (quality), banquet (feast), vin (wine), pomme. de terre (potato), français (French); choucroute (sauerkraut), sauce, frites (French fries), charcuterie (delicatessen) copains (pals), steak, canard (duck), restaurant, and boeuf (beef)* ; and women about : *équilibré (balanced), légumes (vegetables), sain (healthy), laitages(dairy products), vitamines (vitamins) kilos, lait (milk), varié (varied).*

Young people rather talk about in *bouffe (grub), copains (pals), restaurant, couscous, dessert, gâteau (cake)* ; and the elderly about : *peu (few), sans (without), excès (excess), modérément (moderately), raisonnable (reasonable), and cholesterol*.

Crossing representations with actual behaviour shows, as could be expected, significant interaction (Lahlou, 1994b) : representations are like mental organs, some of their nuclei develop more when frequently used by the subject, while others, although still present, atrophy. For instance, elder people, who are more restricted in their intake behaviour because of health concerns, will display representations where the "balanced", and "not too much" scripts are more prominent ; whereas youngsters, focused on the social, fulfilling, and enjoyable aspects of eating, display more developed "eat what I like", "eat my fill", and "conviviality" nuclei.

Conclusion

Lexical analysis allows an anatomy of social representations at a scale that was hardly accessible with qualitative techniques. Results obtained on food representations are coherent with the classical findings. Beyond investigation on human subjects, the technique enables analysis of material coming from documentary collective sources, like dictionaries (but many such sources could be considered). Those conceal implicit models of knowledge which apparently account for many "live" representations that can be observed in the subject. Also, the study of linguistic sources, as language acts as a collective memory, can yield patterns of a former state of representational development, in sociogenesis or epigenesis.

The nuclei of the basic paradigm can be articulated into pragmatic scripts, providing economical guidelines for everyday behaviour. Detailed analysis can therefore yield some elements for behaviour prediction. Hereby, although the method still remains to be improved for better accuracy, it opens interesting perspectives in the study of human thought and behaviour in general.

This method opens a new field of linguistic material to psychosocial investigation on a large scale. The main advantages are : the possibility of investigating quickly and cheaply a quantity of data that is beyond capacity of a single analyst with classical content analysis,

robustness of results, limitation of analyst's biases, possibility of historical and international comparison.

References :

ABRIC, Jean-Claude (1984). A theoretical and experimental approach to the study of social representations in a situation of interaction. in. R. Farr and S. Moscovici, eds. Social Representation, Cambridge, Cambridge U. P. . pp. 169-184.

ABRIC, Jean-Claude (1993). L'étude expérimentale des représentations sociales. in Jodelet, D. éd. Les représentations sociales. Paris, PUF, 3ème éd. , pp.187-203 .

BEAUDOUIN, Valérie, LAHLOU, Saadi (1993). L'analyse lexicale : outil d'analyse des représentations. Cahier de recherche n°48. CREDOC, Paris.

CONSTANT Patrick(1991). *Analyse syntaxique par couche*. - Paris, Thèse de doctorat ENST.

DURKHEIM, Emile (1895). Les formes élémentaires de la vie religieuse. , 1991. Le livre de poche. Editions Générales Françaises. Paris.

FISCHLER, Claude (1990). L'homnivore. Odile Jacob, Paris.

FLAMENT, Claude (1993). Structure et dynamique des représentations sociales. in Jodelet, D. éd. Les représentations sociales. Paris, PUF, 3ème éd. , pp.204-219.

FRAZER, J. G. (1890). Le rameau d'or. Paris, Robert Laffont. 1981.

FREUD, Sigmund (1921). Psychologie des masses et analyse du moi. . Oeuvres complètes. Volume XVI. Paris, P.U.F. 1991. pp. 1-84.

FREUD, Sigmund. "Psychanalyse" et "Théorie de la libido". 1923. Oeuvres Complètes. Volume XVII. Paris, P.U.F. 1992. pp. 181-208.

GUIMELLI, Christian, ROUQUETTE, Michel-Louis (1992). Contribution du modèle associatif des schèmes cognitifs de base à l'analyse structurale des représentations sociales. Bulletin de Psychologie. Tome XLV, n°405, janvier-février . pp. 196-202.

JODELET, Denise éd. (1989). Les représentations sociales. Paris, Presses Universitaires de France, 1993.

JUNG, Carl Gustav (1933). Dialectique du moi et de l'inconscient. Idées, Galllimard, Paris. 1964.

LAHLOU, Saadi (1992). *SI/ALORS : BIEN MANGER***Erreur ! Aucun nom n'a été donné au signet.** *? - Application d'une nouvelle méthode d'analyse des représentations sociales à un corpus constitué des associations libres de 2000 individus*.- Paris, CREDOC, Cahier de recherche du CREDOC, n°34.

LAHLOU, Saadi (1993). Lexical analysis : an approach to social representations of food. European Interdisciplinary Meeting : Current Research into Eating Practices. Potsdam, Germany, October 14-16.

LAHLOU, Saadi (1994a). Modélisation des représentations sociales par l'analyse lexicale des énoncés de dictionnaires : une nouvelle approche pour la psychologie sociale. Colloque international Consensus ex Machina. Paris, Sorbonne, 20-23 avril. ALLC-ACH-Inalf/CNRS-ENS Saint Cloud.

LAHLOU, Saadi. (1994b). Penser Manger : les représentations sociales de l'alimentation. Thesis (to be submitted 12/1994). EHESS.

MOSCOVICI, Serge (1961).- *La psychanalyse son image et son public*. -Paris, Presses Universitaires de France, 1976 .

MOSCOVICI, Serge (1974). Quelle unité : avec la nature ou contre ? In : L'unité de l'homme. Pour une anthropologie fondamentale. Tome III. Morin, E. , Piattelli-Palmarini, M. Eds. Editions du Seuil, Points. Paris, 1974. pp. 286-319.

PIGAMO, Frédéric (1990).- *Outils de traitement sémantique du langage naturel* - Paris, Thèse de doctorat ENST.

REINERT, Max (1983).- "Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte". - *Les cahiers de l'analyse des données*, Vol VIII, n° 2.

REINERT, Max (1990).- "ALCESTE, une méthode d'analyse des données textuelles. Application au texte «Aurélia de Gérard de Nerval" *Bulletin de Méthodologie Sociologique,* 26.- pp. 25-54.

ROSCH, Eleanor R. (1975).- “Universals and Specifics in Human Categorization”, R. BRISLIN, S. BOSNER & W. LONNER eds. *Cross Cultural Perspectives on Learning*, Halsted, New-York.

ROZIN, Paul (1988). Cultural Approaches to Human Food Preferences. in : Nutritional Modulation of Neural Functions. New-York, Academic Press, . pp. 137-153.

ROZIN, Paul (1990). Social and Moral Aspects of Food and Eating. In : The legacy of Solomon Asch : Essays in cognition and social psychology. Irvin Rock (ed.). Lawrence Erlbaum Associate Publishers, Hillsdale, New Jersey. 1990. pp. 97-110.

ROZIN, Paul, MILLMAN, Linda, NEMEROFF, Carol (1986). Operation of the laws of sympathetic magic in disgust and other domains. Journal of Personality and Social Psychology, 1986, Vol. 50, n°4, pp. 703-712.

SPERBER, Dan (1989). L'étude anthropologique des représentations : problèmes et perspectives, in Jodelet, D. éd. Les représentations sociales. Paris, Presses Universitaires de France, 1989. pp. 115-130.

SPITZ, René A. (1965). De la naissance à la parole. La première année de la vie. Bibliothèque de Psychanalyse. Paris, P.U.F. 1968.

VERGES, Pierre (1994). Approche du noyau central : propriétés qualitatives et structurales. In : Guimelli, C. (éd.) Structures et transformations des représentations sociales. Delachaux et Niestlé, Neuchâtel, pp. 233-253.

WAGNER, Wolfgang (1994). The Fallacy of Misplaced Inentionality in Social Representation research. Manuscrit non publié, Université de Linz, Autriche, 26 p. 1994.

WHORF, Benjamen Lee (1956).- “On the Connection of Ideas”. - in *Language, Thought and Reality*. John B. Caroll ed. The MIT Press, Cambridge, Massachussetts.- pp. 35-39.

Connotation space of lexical traits (corpus "Eating well")