

The Prospero software program: an alternative tool for the study of social representations

Nikos Kalampalikis, Fabrice Buschini

▶ **To cite this version:**

Nikos Kalampalikis, Fabrice Buschini. The Prospero software program: an alternative tool for the study of social representations. European Review of Applied Psychology / Revue Européenne de Psychologie Appliquée, 2002, 52 (3-4), pp.241-251. halshs-00494628

HAL Id: halshs-00494628
https://shs.hal.science/halshs-00494628

Submitted on 24 Jun 2010

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

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

European Review of Applied Psychology, 52(3-4), 2002, 241-251.

The Prospero software program: An alternative tool for the study of social representations

Nikos KALAMPALIKIS & Fabrice BUSCHINI*

Laboratoire de Psychologie Sociale
École des Hautes Études en Sciences Sociales

The authors would like to thank anonymous reviewers for their helpful comments on an earlier version of this article.

Abstract:

This article is in keeping with a reflection on the language and the numerous computer tools available for its analysis in social psychology. The article stresses the necessity to avoid the trap that consists in blindly using the results of computer textual analyses. The researcher should actually wonder about the epistemological presuppositions of the software programs he uses and must check if those programs do fit with his own theoretical frame. Thus, for the study of social representations, software programs based on the standard method of content analysis are presented as the most relevant. Results from empirical research studies, carried out with a software of this kind (Prospero), provide illustrations on the possibility and the benefits of creating dynamic and flexible networks of categorizations under three different qualitative forms: Synonymy, analogy and taxonomy. In accordance with the theory, these categorizations allow to grasp the processes of objectification and anchorage, which are constituent of social representations, within their context of linguistic communication.

Key words: Synonymy, analogy, taxonomy, social representations, textual analysis.

Résumé :

Cet article s'inscrit dans une réflexion autour du langage et des nombreux outils informatisés dont nous disposons pour son analyse en psychologie sociale. Il insiste sur la nécessité d'éviter le piège consistant à utiliser aveuglement les résultats d'analyses textuelles informatiques. Le chercheur doit en effet s'interroger sur les présupposés épistémologiques des logiciels qu'il utilise et vérifier qu'ils s'accordent avec son propre cadre théorique. Ainsi, pour l'étude des représentations sociales, les logiciels basés sur la méthode classique de l'analyse de contenu sont présentés comme les plus pertinents. Les résultats de recherches empiriques, réalisées à l'aide d'un logiciel de ce type (Prospéro), viennent illustrer la possibilité et l'intérêt de créer des réseaux de catégorisations dynamiques et flexibles sous trois formes qualitatives différentes : la synonymie, l'analogie et la taxinomie. En phase avec la théorie, ces catégorisations permettent d'appréhender les processus d'objectivation et d'ancrage constitutifs des représentations sociales dans leur contexte de communication langagière.

Mots-clés : Synonymie, analogie, taxinomie, représentations sociales, analyse textuelle.

* The second author is now at the university of Geneva.

European Review of Applied Psychology, 52(3-4), 2002, 241-251.

The Prospero software program: An alternative tool for the study of social representations

The analysis of the textual data is a major issue for social psychology. Only those data enable us to study the language with a sufficient distance. As a photography, they provide us, better than a sound recording, a vision that is at the same time permanent and isolated from a phenomenon in which we are constantly immersed and involved. This textual crystallization of the language, like retranscriptions of conversations or words associations, of press articles, web pages, texts or various books etc., this visible trace of the speech thus, often constitutes the raw material by which we can understand the social phenomena, and in particular that of social representations (Moscovici, 1984; 2001).

Recently still, the only possible way to analyze these textual data was a long and exhausting manual classification work in order to achieve a content analysis, forwarded by statistical analyses of obtained classes or categories. Nowadays, the volume of the textual data presented in various fields has been increased drastically due to the development of the new means of communication as well as numerous published materials. Faced with this, the traditional content analysis (Bardin, 1977) being already an enormous task, is now condemned to become a titanic and impossible work if the researcher does not impose to himself some limits. Nevertheless, the high-tech progress, which is responsible of this growth, gives us at the same time the opportunity to control it.

Thanks to this technological revolution, we can now benefit from the storage capacities and speed of data treatment, which enable us to deal with a quantity of information we could not imagine before. Moreover, many software programs dedicated to the textual analysis are already using those tremendous resources (e.g. Alceste, Tropes, Nudist, IN VIVO, Ethnograph, etc.). Although, it is clear that at this point we need to show caution. If it was possible to be indulgent with the traditional content analyses, due to the extent of required work, one ought to remain critical with computer-assisted textual analyses. A researcher who is interested in the study of social representations must take seriously in consideration both data processing tool and language.

1. Computer assisted textual analysis

In our opinion, thoughtfully use of the data-processing tool means to leave a kind of numerical behaviorism where one is satisfied in entering any form of data, as an input, and impatiently waiting for some results ready to use, as an output [1]. Without being specialists in data processing, we should stop considering the computer as a new *black box* and should become aware of the involved process. It means we should face a variety of the existing software programs and what they imply, particularly in the textual analysis. In the light of the growing approachability of the computer's universe with its deluding nature, like a magical illusion, one easily forgets that behind any data-processing software program stands more or less specialized philosophy. The textual analysis software programs do follow this rule. Moreover, as we can find different types of them, it means that there are also in this field different ways of understanding the text, therefore the language. A recent article of Jenny (1997) underlines extremely well this fact. The author has listed nearly thirty software programs of textual

[1] Very often, the only shown reflection of the virtual behaviorist is in preparing his data.

analysis and specified their own specifications, as well as their assumptions and theoretical orientations. Moreover, he reminds us that, behind these software programs there is often a statistical approach of the text, which is neither neutral. Jenny brings together the various software programs into four large groups according to their specific approaches of the text: lexicometric, socio-semantic, approach in networks of associated words, and that of a propositional (or clausal) and predicative analysis of the speech. In the present article where we want to consider the computer assisted textual analysis in relation to the study of social representations, the two last approaches will not be accounted for, due to their lesser relevance for that topic.

1.1. Lexicometric analysis: to give meaning to the calculations

Among these families, the most largely used approach in the field of the social representations is now indubitably the lexicometric one (De Rosa, 1995; Lahlou, 1998). According to proper terms' of Jenny, it « consists in comparing lexical profiles (relative distributions of the lexical occurrences, without need for preliminary reading) inside a corpus or between textual corpora, with several alternatives according to the manner of cutting or not the corpora, of lemmatizing or not the "lexemes" (or words), and of including or not the "blank words", etc. » (1977, p. 66).
Our position is that, if one intends to analyze texts or speech in the aim of studying social representations, the qualitative approach, which does not exclude a quantification of the data, must be privileged. Accordingly, the traditional analysis of contents is particularly suitable.

1.2. Analysis of contents: the meaning first and foremost

The principle of the contents analysis is, above all, based on the conception of more or less global categories or classes which have each their own meaning, and then on the distribution of the various elements of the texts within these classes with an aim of description and comparison. The last stage of the analysis is then a calculation, which may appeal to the descriptive or inferential statistics. It corresponds to the approach Jenny called socio-semantic. According to him, software programs of this kind proceed « by segmenting the corpus in relevant units of significance and by achieving a multidimensional categorization in accordance with the grids of conceptual analysis specific to each research (from a traditional point of view of *a posteriori* coding, where a researcher reads the text, "marks" and codes himself the units of meaning of the corpus), and by a possible use to statistical methods » (1997, p. 66).
For the software programs based on the lexicometric approach, it is exactly the inverse. All starts with computation. And it is the computation that will generate classes or groups of elements of the texts. The last stage then consists in giving a meaning to the classes obtained. A researcher is often faced with a huge amount of pages full of graphics and a shortened grammar of mathematics, which give him the ambivalent feeling of the possession of a « serious and successful » (i.e. scientific) analysis. But he also needs to find the keys for those enigmas. The two approaches are thus radically different: in the first case one makes calculations starting from the meaning, in the other, one finds meaning from the calculation.

Insert schema 1 here.

1.3. To choose a method suitable to the field investigated

In the field of the social representations, where symbols, images, and meaning itself are of a great importance, the lexicometric approach seems less appropriate than the content analysis.

The meaning lies in words[2], their using, their sociocultural context of production and re-production: It is not created by unspecified statistics. Moreover, two significant critics can be addressed to the lexicometric approach.
The first one refers to the practice of obtaining classes through cutting the text in arbitrary units. A number and the contents of these classes can vary, according to the size of the units, allocated by default or by the user.
The second critic is that the meaning *a posteriori* given to the classes according to their contents is sometimes forced or artificial, but almost always difficult, like when you have to find a meaning for a dimension given by a factor analysis (*cf.* Lorenzi-Cioldi, 1997). Moreover, it is not always simple to know if these classes explain more the meaning than the structure of language.

Neither the content analysis is safe from any critics. Generally, it is criticized for the subjectivity in making choices of the classes or categories, as well as in distributing the elements of the texts within those categories. Without claiming to bring a final solution to this problem, we will show it is possible to minimize it. But, let us first finish this quick overview of the computerized textual analysis in specifying that, if the lexicometric method does not seem well adapted for the study of the social representations, it can nevertheless be extremely valuable in other fields where one is more focused on the structure of the language For the field of social representations, we think that only software programs based on the traditional content analysis can be productive.

2. The language

In the field of the social representations, the language occupies a role of mediator between the subject and the represented object, which is real or ideal (Jodelet, 1989a). The representation of the social is inseparable from a setting in speech relating to it. The social construction of reality thus largely depends on both the various structures specific to each language as autonomous systems, and all the potential linguistic games within those structures (Whorf, 1956; Wittgenstein, 1965; Orwell, 1950; Bruner, 1991). For the researcher, the language gives, at the same time, a privileged access to the study of the common sense that is elaborated, shared and spread within, and to the different forms of the knowledge, thinking and acting in relation to the social object. Already in 1961, Moscovici affirmed that « to know a thing socially is to speak it » (1961, p. 148) and, since, several studies in social psychology confirmed the existing close relationship between social representations and communication systems in the organization symbolic system of the knowledge and the language (Jodelet, 1989b; Markova, Foppa, 1990; Lahlou, 1998).

Although it is essential for the social psychologist to carry out a reflection on the language, here we would not know nor do not want completely to accomplish this complex task. We simply will propose some tracks of reflection allowing to reconsider the traditional content analysis within the framework of social representations studies, in the light of the new possibilities given to us by the data-processing tool. More precisely, we will focus on the aspects of categorization or thematization, which are, in our opinion, crucial in the theory that occupies us.

The traditional content analysis was like the labor of Sisyphus, with its classifications to be continuously repeated, its transitory thematic categories disappearing with each research. On

[2] To say that the word possess a meaning does not at all exclude the analysis of the systems of representations inside which this « possession » finds its legitimacy (cf. Althusser, 1976 ; Van Dijk, 1996).

the contrary, due to data-processing storage the new content analysis gain a memory. Once created thematic categories do not disappear any more. Moreover, they can be re-used and expanded by each new research. This way, it is possible to carry out a cumulative set of different categories resulting from various research and various researchers. However, this does not mean that this tool one day might become exhaustive neither that the more we use it, the less work will be required. The problem of the ecological validity still remains. Is it possible to generalize and apply already existing categories to other materials or are they peculiar to one particular research? Do they reflect a consensual vision of reality or are they the product of the researchers' idiosyncrasy? We will try to answer this kind of questions by distinguishing various types of categorization that are, according to us, relevant for the theory of social representations.
First, we will quickly present the textual analysis software program, which provoked our reflection on the problem of categorization.

3. The Prospero software program

Prospero is a textual analysis software program, which was created and developed by a team of sociologists. They wanted, above all, to solve methodological problems similar to those we have already mentioned for social psychology (technological progress, size of the corpora, subjectivity of the content analysis, co-construction of the meaning, etc).

Accordingly, this software program gives the possibility « to make the specialized expertise or the solitary scholarship more relative, and to allow the production of collective proofs on the forms of memorizing, the categories of analysis and the modes of proof developed by the researchers. It is thus a question of making possible an accumulation of knowledge without dissociating it from the acts of reflexivity. In this way, the principal function of Prospero is to allow the researchers to conduct experiments on large corpora of texts based on the interactive creation of an "other scene" which can double, without any reduction, the relation of intimacy linking researchers to the texts they are investing » (Chateauraynaud & Torny 1997, pp. 430-431, our translation).

The meaning of a corpus of texts analyzed with Prospero appears right away through qualitative categorizations that are generated by a set of two levels of representation, which constantly remain under the control of the researcher. The first level refers to the particular status allotted to each element of the text (at the grammatical, syntactic and semantic level). This is a first-degree identification of each element within the « puzzle » of texts. The second level, aims to cover specific semantic and pragmatic dimensions and to study their dynamic evolution (Chateauraynaud, Torny, 1999). Researcher's task is to specify the meaning in using the flexibility of the software program and particularly by refining and enriching the two proposed levels according to the characteristics of the material to be analyzed and to the specific aims of the research.

At the first stage, while leaving the texts intact, the program reads the whole corpus to be analyzed and assigns its words to the following categories, constituting the first semantic structure, the first level of representation (which reflects the traditional grammatical structure of the language):

a) The « *entities* » (people, objects, concepts, impressions, emotions) that correspond at the grammatical level to the names or nominal groups.
b) The verbal forms (called « proofs ») that can be compared at the grammatical level with the verbs or verbal groups.

c) The « *qualities* » (qualifications - attributions of qualities or states) that concern the qualifying adjectives.
d) The « *markers* » (adverbs, elements who take an active part in creating the meaning of the statements)
e) The « *tool words* » (various articles, pronouns, conjunctions, operators etc.) and « *numbers* » (numerical values in figures and letters).

The following table gives an outline in the way in which the words of the corpus are distributed within the categories:

Insert Table 1 here

The indefinite forms correspond to the elements that do not appear yet in the basic semantic structure. The researcher can then choose to allot them to one of the above categories of which they belong, if he finds relevant to include them in the analysis.

Once the identification of the textual elements is established, the more important part of the analysis is focused on the elaboration of three conceptual instruments that lead to the second level of representation of the analyzed speech: « *fictitious beings* », « *categories* », and « *collections* ».

3.1 Fictitious beings

The concept to be fictitious corresponds to the various manners of naming an object. It thus enables to consider the anaphoric processes and the links of co-references, to take the measurement of the variations (attributive, metonymic, metaphorical) produced around the same object (human, concept, concept etc.) and to follow the appearance of naming or strange designating and also to characterize their context. With this structure it is possible to follow the evolution of the set of designations referring to certain objects defined by the researcher through the corpus and, therefore, to see in each text, designations or properties usually ignored, forgotten or left in an implicit state.

3.2 Categories

The categories are built by regrouping the different families of elements (« entities », « proofs », « qualities », « markers ») under the form of long lists, which point to the same relevant universe. Thus, the categories contribute to the identification of *kinds of speech*, which characterize the texts. In this manner, an access to the relation of representation between words and the used conceptual universe is possible. The categories always have a status and a dynamic position, and are ready for the necessary modifications. Of course, the same element can refer to different categories according to the way it is used in different discursive universes, thus respecting its polysemy.

3.3 Collections

The collections contribute on a conceptual and more holistic level to the appearance of the systems of representations by regrouping the elements of the corpus referring to those systems. The elements of the corpus are, this time, gathered according to their membership of a generic category following the method of the natural science.

The aim of this brief presentation was not to explain how to use the software program. Our intention was only to emphasize the two levels of representation already described together with the functionality of this basic framework for discourse analysis[3]
Further, we will present how the analytical possibilities offered by Prospero can become relevant for the analysis of the social representations.

4. Three types of categorization

The framework of the Prospero software program lead us to consider the various manners of categorizing which can be relevant to particular aspects of the social representations. Those are both general and particular. General, in the sense they generate a consensual vision of the reality surrounding us, and particular, insofar as this vision will be inflected according to the various social groups which privilege their own aspects or facets of the same representation of reality (Doise, 2001). On the other hand, the representations are for the human groups a means of controlling and understanding the reality, which surrounds them, and also a means of communicating between them on this reality. Starting from these aspects of the social representations, we will show one can distinguish three great forms of categorization of the meaning - the synonymy, the analogy and the taxonomy - which differently refer to the objectification and the anchoring, the two constituent mechanisms of the social representations (Buschini & Kalampalikis, 2002). In order to avoid talking at cross purposes, illustrations will be presented for each form of categorization.

4.1. Synonymy

This manner of categorizing corresponds under Prospero to the creation of the fictitious beings. One should not understand here the synonymy in a strict sense of an identity of meaning between terms, but rather, following the distinction of Frege (1971) between meaning and reference, in a sense of an identity of the referents coming from terms having a slightly different meaning. Thus, the object of categorization here is the referent. Synonymy refers to the aspect of the social representations that concerns the appropriation or the control of reality achieved by the process of denomination (Kalampalikis, 2002; Mauss, 1978; Moscovici, 1999, 2000; Philogène, 1999). This aspect is more often partial than global. All the work will thus consist here in gathering within the same category various denominations of the same object[4], various possible meanings of the same referent. It is not a strict synonymy, but a « social synonymy » within the meaning of Quine (2000)[5].

4.2. Illustration of synonymy

In a research study on the values and representations engaged in medicalized male contraception (Apostolidis, Buschini, Kalampalikis, 1998)[6], the reference to the condom was over-represented compared to the other means of contraception. This was because the

---

[3] Within this elementary framework, the software also allows to compare under-corpus according to external variables (sex, matrimonial status, age, dates, etc.) or to internal variables (thematic such as groups which use most of the words in –ism or not, texts referring or not to a particular concept, etc.). It also allows to identify semantic networks, to follow the evolution of the discourse on a temporal dimension, or to proceed to specific interrogations on the corpus using a set of formula easily modulated by the researcher.
[4] Object should be here understood in a broad sense: it can be concrete (individual or thing) or abstract (concept, discipline, etc.).
[5] One can wonder besides whether strict a synonymy really exists (e.g. odor *versus* scent).
[6] The full report of this study, carried out from 48 nondirective interviews, can be downloaded at the following address: ftp://www.ehess.fr/pub/lps/cmm.pdf

contraceptive dimension of the condom was associated with a prophylactic dimension in the relevant context of the AIDS. The creation of the fictitious being « *condom* », i.e., the various denominations used by our subjects to speak about the condom (« *condom* », « *hoods* », « *rubber* », « *latex* », « *condom* », « *French letter* », *etc.*), allowed us to catch a synthetic vision of an object which otherwise would have remained fragmented, and to gather the various universes activated. Thus, the mention of this object is located throughout the verbal productions, its weight is measured in terms of frequency (global and local), and its relations with the other elements of the text appear inside a context (global or local network).

On this subject, it is important to stress that the association between condom and prophylaxis, to the detriment of the contraceptive aspect, is more determining in the discourses of men, comparing to those of the women. Indeed, for the women, and mainly the older ones, the use of the condom more spontaneously encounters the dimension of contraception. Moreover, differences in terms of denomination of this object appear according to the sex of the people interviewed. Without the creation of the fictitious being *condom* these conclusions would have been more difficult to drawn.

Another study undertaken on the actual experiences of change within a big company and on the representations of its future (Lage et al., 1999) provides us an additional illustration of synonymy. It shows how the different denominations of the common referent « the employers » are distributing according to the different functions of the company's staff (salesmen, technicians and retrained).

One can thus notice that the commercial function mentions very little the employers and when it does it, it is rarely as « bosses » but rather under a more technocratic wording (leaders, managers, etc.). To the contrary, technicians talk essentially of the « boss » and use only more rarely other words.

In the category of retrained employees, corresponding to the people from the technical function being reassigned to the commercial one, appears a use of synonymy that perfectly reflects their particular situation. While keeping the terminology of its origin, it appropriates new names close to the commercial environment or imagined as such. The « boss » is always very present, but new declensions of the referent that are closer to the commercial universe appear too. Thus, in front of a same object (the employers), each sub-group of the population uses a specific vocabulary, which reflects the particular inflections with which one can appropriate a common representation.

Insert Table 2 here

Synonymy mainly shows how a specific object (the employers) or a part of a new representation could be objectified through its different denominations. It also provides information on the way different groups could anchor their shared representation into specific semantic universes.

4.3. Taxonomy

With taxonomy, we are finally in a more general area of the social representations where reality is already well defined. The object of categorization is then the class, the list or, in the Prospero's words, *the collection*. The question is to gather within a same unit all the elements having a same characteristic in common, as the natural science had laid down nomenclatures and drawn up catalogues of the life that surrounds us. The categories resulting from taxonomy will be consequently different from the others for two reasons. First, they will be made of lists longer than those resulting from synonymy or analogy are, and they could also be divided into subcategories. Second, as they reflect a kind of common or encyclopedic knowledge, these categories can be, at one point, exhaustive. For example, today, one can easily draw up the list

of all the countries of the world, within which some countries could be gathered (e.g. according to their continent or their economic richness). The categories thus established are unquestionable, only the historical process can change them radically. The list of all the countries of the world was modified after the fall of the Berlin Wall and, more recently, after the war in ex-Yugoslavia. And it could be the same with the current events in Iraq.

In the now familiar field of medicalized male contraception, it was thus possible to establish the list of known contraceptive means divided according to their generic specificity:

Insert Table 3 here

4.4. Illustration of taxonomy.

In a recent study, Pétard, Kalampalikis and Delouvée (2001) examined the manner in which the history of the social psychology was constructed between 1946 and 2001 in the French speaking handbooks of this discipline.
For this purpose, they proceeded among other things to a taxonomic classification of all the mentioned authors in the history chapters they analyzed (n=25) through different time periods. By this mean, they were able, on the one hand, to group quantitatively all authors' names mentioned in their corpus (n=702), and on the other hand, to highlight, using a chronological classification containing six time periods, the genealogy of discipline, as the history chapters of the social psychology handbooks present it.
They also showed that, out of all the authors mentioned, more than the half (55.8 %) was quoted during the period following the war until nowadays (1950-2000). However, if our interest is focused on the authors who create a consensus, meaning here those who are at least present in one fourth of the full corpus, it is easy to realize that, on the other hand, only 6,55% of the mentioned authors (n=46) fulfill this condition. On the other hand, only 20 % of these consensual authors fall in this same period (1950-2000). On the contrary, these most typical authors come essentially from the 1900-1950 period (53.3 %), whereas out of all the mentioned authors (n=702), this period only represents 28.3 %.

This example of taxonomy shows the benefits of constituting an objective taxonomic universe of reference with regards to the corpus of analyzed texts and, at the same time, it emphasizes the need to highlight the specific activations of this taxonomy according to the different corpora or sub-corpora analyzed. The study briefly mentioned here demonstrates the possibility of highlighting the anchoring of an entire discipline in a specific period of time, and the relative consensus which characterizes the original context in which they appear, that is, the history chapters of social psychology. Beyond, it reflects the process of anchoring a new representation into the pre-existing socially shared reality.

4.5. Analogy

The regrouping by analogy is achieved under Prospero by the creation of various categories (*cf* § 3.2). Comparing to synonymy, the analogy is more typical of the usual categorization used in the traditional content analysis, where the criteria of the regrouping is the idea or the concept. The analogy is primarily related to the comprehension and the control of reality and refers to partial or general aspects of the social representations. To categorize consists here in putting together all the terms referring to the same order of idea within a same set.
A significant part of the above mentioned research study on male contraception was focused on fears and side effects. For a large majority of the interviewed people, medicalized contraception is never a natural process, since it acts against a phenomenon inherent to the

nature of human being, namely the reproduction, so as to prevent its finality. Thus, here arose the need for categories relating to these dimensions. Given is an example of the analogical category built around the concept of anxiety:

Insert Table 4 here.

In order to precise the distinction between the synonymy and the analogy, it will be noticed that the analogy does not point to a common referent, but that it can include several distinct referents belonging to the same concept. Thus, in our example, the gathered terms are not different denomination of the same object « anxiety », but various states referring to the idea of anxiety. No need to say that this category can be expanded with further researches where it could be re-used.

4.6. Illustration of analogy.

The applied study on male contraception (Apostolidis, Buschini et Kalampalikis, 1998) provides us another interesting example of categorization based on analogy. This concerns an analogical category created within the grammatical field of verbs. In this context, the authors were interested in the emotional feeling conveyed by the discourse and particularly relevant for the object of this specific study. Table 4 provides a comparison of analogies obtained for the positive and negative feelings. One can notice there that, out of all the interviews (n=46), people express much more positive than negative emotions, by the verbs they use[7]. But, in return, the variety of negative expressed emotions is clearly more important than for the positive emotions.

Insert Table 5 here

With the small quantity of information provided in this table, one already gains information of quality, which makes obvious the benefits of analogy. It allows the access to both the processes of anchoring and objectification in order to understand them better. Whereas taxonomy allowed principally to shed a light on the phenomenon of anchoring and synonymy rendered objectification more obvious, analogy gives us access to the dynamic and sometimes complex interaction between both processes, as it will be summarized in the following section.

4.7. Synthesis

The examples that we have just presented only provide a concise and general idea of the research studies from which they are draw out. The aim was solely to illustrate our argument on the three forms of categorizations presented and not to account for the empirical research studies that the reader can refer to in other respects.
In terms of Logic, synonymy and taxonomy come both under a definition by *extension*. The common referent for synonymy and the class or the collection for the taxonomy will be defined by the addition of their elements like in the linguistic process of *denotation*. Conversely, the analogy depends on a definition by *comprehension*, since the categories based on ideas or concepts are made with the various *connotations* of these ideas or concepts. One will find in the following table a summary of the various characteristics mentioned for each of the three forms of categorization.

[7] Of course, the authors have been careful to check those verbs were not uttered within the interviews in the negative form.

Insert Table 6 here

5. Interest for the study of social representations

Among the various definitions of the social representations, the one recently given by Moscovici (1998) seems to us particularly relevant for our matter. In his definition, Moscovici considers the representations in two ways. In a static view, they can be seen as theories around a topic, allowing to describe and classify people and things, to make them understandable. From a more dynamic point of view, they can be considered as flexible networks in the form of encyclopedia of ideas, metaphors or images. This second manner of defining the representations clearly emphasizes the need to achieve various forms of categorizations, in order to understand the complex relations of these networks.

As we already mentioned, synonymy is related to the mastery and the control of reality. Synonymy is relevant to grasp the way by which the groups or the individuals appropriate the various aspects of a representation by naming them. It is thus a means to have an access to the inflections and sociocultural variations of a same « theory » about a representational object.

Within its different grammatical levels and according to the degree of activation of the different categories, the analogy provides information on quite a few aspects of the protean process of anchoring. First, with the categories of substantives and adjectives, it informs on the semantic shell of the representation. But, thanks to the categories of verbs, it can also indicate to what extent the representation acquire instrumental aspects. The typological function of the representations could partly be seized by the categories of tool words, such as personal or possessive pronouns.
The categories of the analogy can also be built to investigate social representations under a specific theoretical approach. For example, one could study a representation within the linguistic theory of the speech acts, by differentiating categories of verbs according to their illocutionary force. But the risk is then important to fall into the trap of categories sometimes too specific or local which are only relevant to a specific research. And, once again, one can be accused of a lack of objectivity.
One can however expect for the analogy the existence of consensual categories crossing the different grammatical levels. We think here of the analogical categorizations that are close to the thêmata (Moscovici and Vignaux, 1994; Markova, 2001). They are generally related to dimensions such as time, space, morals, etc., and have their negative counterpart. Whether an analogical category is objective and consensual can be proven by the increased number of research studies. A category one cannot use in different studies will certainly be too subjective or specific. But, on the contrary, if it passes the test of successive research studies, it will then be close to the idea of thêmata.

As for taxonomy, it points to a consensual representation of the world. It accounts for a level of the knowledge, an encyclopedic learning, of the world. That is a « déjà-là » situation, both arranged and ordered, on which the new representations are grafted. The activation of one or another taxonomic categories will thus provide information on the way new representations are anchored, at a cursory look, inside a body of knowledge objectified by the consensus. The categories implied in taxonomy are amongst the least subjective and local because they describe an established consensus of reality.

These three forms of categorization have a strong relationship with the objectification and the anchoring, the two major processes involved in the development of the social representations. Let us briefly recall that *to objectify* is to make words correspond to things, in other words, to

constitute socially the reality. *To anchor* means to name and classify, in order to assimilate and/or to differentiate. If the objectification explains how the represented elements are integrated as terms of social reality, anchoring allows to understand the way they contribute to express, interpret and constitute social relations (Moscovici, 1961).

The relation of synonymy, analogy and taxonomy with these two processes is neither unilateral, nor exclusive, but multiple and dynamic. Given the fact we focused this article quite exclusively on the textual discourse, this relation is even partial.
Synonymy refers to the wording of a material or conceptual object, which, due to its propagation within various communities, will find in each of them various meanings pointing to the same referent. In this sense, synonymy takes a full part in the process of objectification. Here, the cultural or linguistic environment, the context of enunciation, the speech acts, the linguistic games, and the communication, will be significant factors in the choice of the words.
Taxonomy clearly refers to the process of anchoring. It constitutes a framework of pre-existent and objective meanings, and it achieves a requirement of order, that is, it distributes within various classes the elements that can then become relevant variables for a new representation or a new research study.
Whereas synonymy and taxonomy highlight in a privileged way one of the two processes, without excluding the other, analogy brings as much information on the objectification than on the anchoring, as the schematization suggested in table six indicates it, and as one can see it in the following table.

Insert Table 7 here

6. Conclusion

In the framework of this article, we have presented, in a rather concise manner, but nevertheless with many illustrations, a software program for discourse analysis which answers, according to us, to the requirements and principles of research in social psychology, and more particularly to those from the field of social representations. The reader who is anxious to learn more on the functioning of the software program may usefully refer to the references previously mentioned. But, beyond the software program, we wanted to present here a method allowing the exploration of the various aspects of the constituting language of social representations.
This starting point must be pursued in a twofold direction, because questioning the status of discourse analysis in this theoretical framework cannot be achieved without a parallel reflection on the status of language itself.

References

Althusser, L. (1976). Idéologie et appareils idéologiques d'Etat, *in* L. Althusser, *Positions*, Paris, Éditions sociales, pp. 67-125.

Apostolidis, T., Buschini, F. & Kalampalikis, N. (1998). *Représentations et valeurs engagées dans la contraception masculine médicalisée.* Rapport de fin de recherche, sous la direction scientifique de Denise Jodelet, Contrat INSERM n° 4M015B, Paris, Laboratoire de Psychologie Sociale de l'E.H.E.S.S., 183 pages (ftp://www.ehess.fr/pub/lps/cmm.pdf).

Bardin, L. (1977). *L'analyse de contenu*. Paris, PUF.

Bruner, J. (1991). *…car la culture donne forme à l'esprit : de la révolution cognitive à la psychologie culturelle.* Paris, Éditions Eshel.

Buschini, F. & Kalampalikis, N. (2002). La synonymie, l'analogie et la taxinomie : trois formes de catégorisation pour l'étude des représentations sociales, *in* C. Garnier & W. Doise (eds), *Les représentations sociales. Balisage du domaine d'études*. Monréal, Éditions Nouvelles, pp. 187-205.

Chateauraynaud, F. (2002). Prospéro, une méthode d'analyse des controverses publiques, *in* P. Blanchard & T. Ribémont (eds), *Méthodes d'outils des sciences sociales. Innovation et renouvellement*. Paris, L'Harmattan, pp. 61-84.

Chateauraynaud, F. & Torny, D. (1999). *Les sombres précurseurs. Une sociologie pragmatique de l'alerte et du risque.* Paris, Éditions de l'EHESS.

De Rosa, A. S. (1995). Le "réseau d'associations" comme méthode d'étude dans la recherche sur les représentations sociales : structure, contenus et polarité du champ sémantique. *Cahiers Internationaux de Psychologie Sociale*, 28, 96-122.

Doise, W. (2001). *Droits de l'homme et force des idées*. Paris, P.U.F.

Frege, G. (1971). *Ecrits logiques et philosophiques*. Paris, Seuil.

Jenny, J. (1997). Méthodes et pratiques formalisées d'analyse de contenu et de discours dans la recherche sociologique française contemporaine. Etat des lieux et essai de classification. *Bulletin de Méthodologie Sociologique (B.M.S.)*, 54, 64-112.

Jodelet, D. (1989a). *Folies et représentations sociales*. Paris, PUF.

Jodelet, D. (dir.) (1989b). *Les représentations sociales*. Paris, PUF.

Kalampalikis, N. (2002). Des noms et des représentations. *Cahiers Internationaux de Psychologie Sociale*, 53, 20-31.

Lage, E., Personnaz, M., Madiot, B., Buschini, F., Dargentas, M., Givry, C., Kalampalikis, N. & Lima, L. (1999). *Les vécus du changement et la représentation de l'avenir à EDF-GDF.* Rapport de fin de recherche pour le G.R.E.T.S. de EDF, Paris, Laboratoire de Psychologie Sociale de l'E.H.E.S.S., 252 pages.

Lahlou, S. (1998). *Penser manger*. Paris, PUF.

Lorenzi-Cioldi, F. (1997). *Questions de méthodologie en sciences sociales*. Lausanne, Delachaux et Niestlé.

Markova, I. (2001). Social Representations and Communicative Genres. In F. Buschini et N. Kalampalikis (dir.), *Penser la vie, le social, la nature. Mélanges en l'honneur de Serge Moscovici. Papers in honour of Serge Moscovici.* Paris, Éditions de la Maison des sciences de l'homme, pp. 219-235.

Markova, I., Foppa, K. (1990) (Eds). *The dynamics of dialogue.* London: Harvester Wheatsheaf.

Mauss, M. (1978). *Sociologie et anthropologie*. Paris, PUF, Collection Quadrige (première édition : 1950).

Moscovici, S. (1961). *La psychanalyse, son image et son public*. Paris, PUF.

Moscovici, S. (1984). Introduction : le domaine de la psychologie sociale. In S. Moscovici (dir.), *Psychologie sociale*. Paris, PUF, pp. 5-22.

Moscovici, S. (1998). History and actuality of social representations. In U. Flick (Ed.), *The Psychology of the Social*. Cambridge: Cambridge University Press, pp. 209-247.

Moscovici, S. (1999). Noms propres, noms communs et représentations sociales, *Psychologie et Société*, 1, 81-104.

Moscovici, S. (2000). What is in a name ? In M. Chaib & B. Orfali (eds), *Social representations and communicative processes*. Jönköping, Jönköping University Press, pp. 12-28.

Moscovici, S., Vignaux, G. (1994). Le concept de Thêmata. In C. Guimelli (dir.), *Structures et transformations des représentations sociales*. Neuchâtel, Delachaux et Niestlé, pp. 25-72.

Orwell, G. (1950). *1984*. Paris, Gallimard.

Pêcheux, M. (1969). *Analyse automatique du discours.* Paris, Dunod.

Pétard, J.-P., Kalampalikis, N. & Delouvée, S. (2001). L'histoire de la psychologie sociale dans ses manuels, *Les Cahiers Internationaux de Psychologie Sociale*, 51-52, 59-80.

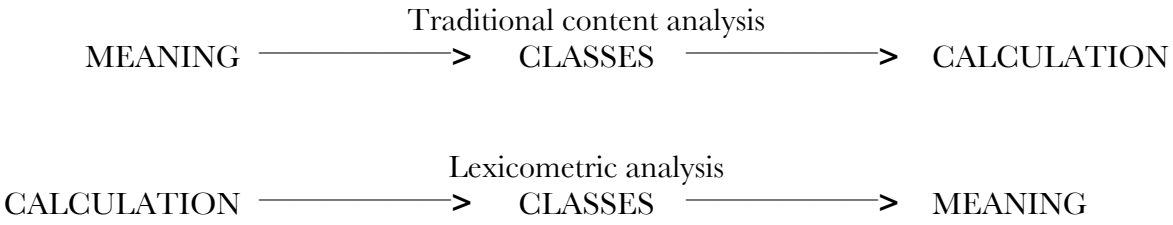
Philogène, G. (1999). *From Black to African American. A new social representation.* Westport and London, Praeger.

Quine, W.V.O. (2000). *Le mot et la chose*. Paris, Flammarion.

Van Dijk, T.A. (1996). De la grammaire de texte à l'analyse socio-politique du discours. In S. Moirand (éd.), *Le Français dans le monde*, num. spécial, Le discours: enjeux et perspectives. Paris, Hachette, pp. 16-29.

Whorf, B. L. (1956). *Language, Thought, and Reality. Selected writings of Benjamin Lee Whorf.* New York: The Technology Press of Massachusetts Institute of Technology and John Wiley and Sons, Inc.

Wittgenstein, L. (1965). *Investigations philosophiques*. Paris, Gallimard.



Schema 1: Opposed processes of the two analyses

Basic categories	**Examples**
Entities	*life, confidence, couple, problems, body, love, nature, town...*
Qualities	*naturalness, reliable, physical, practical, sure, badly, artificial...*
Proofs	*(to) take, (you) know, (it) irritates, (to) cheat, (to) doubt, (I) love, (to) act...*
Markers	*really, a little, always, too, not, without, finally...*
Tool words	*I, you, who, you, what, it is, it, that, but, for, something...*
Numbers	*twenty, 3, 100, 100%, fifty, three, 4/5, 1980...*
Indefinites	typing errors, onomatopoeic terms, neologisms, specific terminologies, etc

Table 1: illustration of the first level of representation

	Total corpus		**Salesmen**		**Technicians**		**Retrained**	
Denominations	**Frequency**	**%**	**Frequency**	**%**	**Frequency**	**%**	**Frequency**	**%**
Boss	39	52	1	14,3	17	68	21	48,8
Leaders	12	16	3	42,8	4	16	5	11,6
Bosses	8	10,7	0	0	3	12	5	11,6
Managers	5	6,7	0	0	1	4	5	11,6
Rulers	4	5,3	0	0	0	0	4	9,3
Manager	3	4	1	14,3	0	0	1	2,3
In high reaches	2	2,7	1	14,3	0	0	1	2,3
Graduates of *École Polytechnique*	1	1,3	1	14,3	0	0	0	0
Top managers	1	1,3	0	0	0	0	1	2,3

Table 2: synonymy for the referent « the employers », according to the professional assignment.

COLLECTION: CONTRACEPTIVE MEANS
CLASS: generic abstinence, coitus interruptus, diaphragm, hormones, masturbation, mutual masturbation, onanism, ogino's method, Ogino, pills, pill, contraceptive pill, injection, injections, condom, oral-genital relations, withdrawal, sodomy, spermicide, intra-uterine device, sterilization, hormonal treatment. *CLASS: masculine* coitus interruptus, condom, cunnilingus, injection, Injection, male pill, men pill, men pills, retention of sperm, vasectomy, Vasectomy. *CLASS: feminine* vaginal shower, sponges, fellatio, fellation, termination of pregnancy, tying of the Fallopian tubes, tubal ligation, female pill, female condom, coil.

Tableau 3: Collection of contraceptive means with its sub-classes.

CATEGORY: anxiety
panic, pangs, agitation, distress, anguishes, anxiety, apprehension, apprehensions, bile, excess stomach acid, nightmare, fear, fears, distrust, doubt, doubts, dramatization, effervescence, fright, embarrassment, sword of Damocles, terror, scare, frenzy, fret, obsession, disadvantages, worried, concern, anxious, concerns, insecurity, unquiteness, the spectrum of, lack of insurance, bad blood, mistrust, nervousness, panic, distressing situation, stress, suspense, terrors, stage fright, worry, scared stiff

Table 4: category of entities on the theme « anxiety »

Positive feeling (148)		Negative feeling (45)	
Instances	Frequency	Instances	Frequency
I love	*85*	*to shock*	*9*
to love	*27*	*to irritate*	*9*
I like	*16*	*shocking*	*8*
They love	*6*	*makes a problem*	*5*
You love	*4*	*to hate*	*5*
To have an orgasm	*4*	*to appall*	*3*
appreciate	*2*	*you fear*	*2*
enjoy	*2*	*to die*	*2*
(s)he loves	*1*	*to shake*	*1*
(s)he has an orgasm	*1*	*feel guilty*	*1*

Table 5: Analogical categorization for the positive and negative feelings.

	SYNONYMY	ANALOGY	TAXONOMY
Object of the thematization	*Referent*	*Concept, Idea*	*Class, List, Collection*
Logic of definition	*Extension*	*Comprehension*	*Extension*
Form of the signification	*Denotation*	*Connotation*	*Denotation*
Relation to the reality	*Control, Take over*	*Understanding, mastery*	*Consensus*
Type of construction	*Denominative*	*Semantic*	*Encyclopedic*

Table 6: Three forms of categorization

Forms of categorization		**Main processes of social representations**	
		Anchoring	**Objectification**
	Synonymy	*x*	X
	Analogy	X	X
	Taxonomy	X	*x*

Table 7: Respective relevance of the different forms of categorization for the processes of anchoring and objectification (small italics indicate a lesser -not an absence of- *relevance compared to capital letters).*