Viva la nanorevolucion! A Semantic Analysis of Spanish National Press

Giuseppe A Veltri *

School of Political, Social and International Studies, University of East Anglia

31/01/2012

1 Abstract

This study analyses nanotechnologies anchoring and codification in the Spanish national press to determine the thematic contexts in which this technology has been discussed. Latent semantic analysis was applied to identify themes based on semantic clusters and their longitudinal evolution. This analysis was carried out on a corpus of more than 600 articles from the most prominent Spanish national newspapers and includes articles from 1997 to 2009. Findings indicate an overall positive coverage and dominant thematic clusters related to national policies, economic development and business opportunities. Surprisingly, controversies surrounding nanotechnology are present in the early years of coverage but have become marginal over time, in contradiction to a general trend that emerged from previous studies on media representations of new technologies.

2 Introduction

Nanotechnology has frequently been described in the media, by policymakers, and even by scientists, as the technology that will provide the catalyst for the next industrial revolution. Yet, such utopian scenarios are also often accompanied by dystopian predictions within the current discursive arena. In July 2004, for instance, ETC Group (which stands for Erosion, Technology and Concentration), a Canadian environmental group, successfully lobbied for the attention of Britain's Prince Charles: soon after reading an ETC report on nanotechnology, His Majesty made a public statement on the risks of this new technology, and the mass media coverage of his view and the debate that followed brought nanotechnology to the

*Electronic address: G.Veltri@uea.ac.uk

attention of a much wider public. News of the royal opinion travelled quickly around Europe and to the United States, and phrases such as nano bots, nanomachines and grey goo began to appear more frequently in many media outlets. Such images were in sharp contrast to previous enthusiastic references to classics of science fiction, such as the movie *Fantastic Voyage*. In fact, several newspapers published pictures from this movie to illustrate the potential medical applications of nanotechnology.

Similarly to previous studies on nanotechnology media representations, this paper is a longitudinal study of nanotechnologies representations in the Spanish national press, and aims to contribute to our understanding of the dynamics of media representations of new technologies. It contributes to the current literature with data from a different country from those usually studied e.g. the United States, UK and the northern European countries in order to assess whether a number of recurrent findings in the literature hold beyond the typical cases. This work also applies a novel combination of text mining techniques, in addition to introducing latent semantic analysis to the study of nanotechnologies media representations, demonstrating the potential of this methodology and critically reflecting on its limitations and possible further applications.

The predominant patterns of media coverage of various novel technologies have been documented in a number of studies. Such studies have highlighted a recurrent pattern in the longitudinal evolution of the salience and nature of coverage. Regarding salience, according to a classic study by Downs (1972), the attention cycle of an issue in the media has five stages, with a gradual decrease in interest as its final stage. A number of studies suggest that coverage of new technologies follows this salience pattern. In addition, studies on media representations of novel technologies have further specified the attention cycle, and made predictions concerning what the initial coverage is likely to focus on and how it can be expected to develop. Nisbet, Brossard and Kroepsch (2003), for example, through a study of newspaper coverage of stem cells in both UK and US, have offered a more detailed picture of the evolution of coverage. According to this study, because of a shift in policy context, when journalists pick up the technological issue after the agenda-building activities conducted by scientists and entrepreneurs, etc., there is a move away from a scientific description of applications and economic benefits to the discussion of risks. Hence, coverage first focuses on scientific and economic opportunities, and later switches to the risks of a new technology.

According to several studies, the same pattern seems to apply also to the representation of nanotechnology. Gorss and Lewenstein (2005), for example, tracked nanotechnology in the American press from 1986 to 2004, and found that media attention to nanotechnology seemed to mimic the coverage of biotechnology in its early stages as a public issue: it started with low levels of attention, which then rose sharply as the issue spread from the elite media to more general outlets. As for biotechnology, the initial coverage of nanotechnology was largely positive and focused on progress or the potential economic benefits, with relatively little discussion of the risks. Nevertheless, according to this research, from the outset nanotechnology coverage does focus somewhat more on risks than was the case

with the coverage of biotechnology, suggesting that issues of public accountability are becoming more salient in the American press.

Similarly, Scheufele and Lewenstein (2005) also made predictions about key themes of coverage: namely a focus on the economic potential or scientific promise of nanotechnology early in the issue cycle. Friedman and Egolf (2005) meanwhile analysed US and UK coverage of nanotechnologies risks between 2000 and 2004 using the most prominent national newspapers and newswires services. They concluded that the UK displayed a slightly more negative coverage than the US and a larger emphasis on societal risks. This is in line with findings from Gaskell, Eyck, Jackson and Veltri (2005), who compared public attitudes towards nanotechnology in the UK and the US and also provided an analysis of nanotechnologies coverage in one British and one American national newspaper (The Independent and The New York Times, respectively). Their research showed that the American newspaper covered the potential benefits of nanotechnology to a greater degree than the British one; it also showed that the coverage of risk, as compared to the benefits, appeared later in both cases.

Although Europe, or at least the UK, displayed overall more negative coverage than US, this does not mean that it was negative overall. Anderson, Allan, Petersen and Wilkinson. (2005) analysed the framing of nanotechnology in the British press, and identified a considerable amount of uncertainty about the nature of nanotechnology, with a tendency to be optimistic and hence underline the potential benefits more than risks. Moreover, the most recurrent news frames were science fiction and popular culture, illustrating a scientific discovery or project, a business story (nanotechnology company in the stock market), and finally Prince Charles statement on nanotechnologies risks and the grey goo scenario. In the few other European countries studied, nanotechnology coverage has been strongly positive. Kjlberg (2009) studied representations of nanotechnology in the Norwegian press and found that a large majority of articles portrayed nanotechnology in a positive light. Two main themes were identified: nanotechnology as important for the future and nanotechnology as under control. Kjrngaard (2010), meanwhile, studied how nanotechnology has been framed in the Danish national newspapers between 1996 and 2006. Overall coverage had a remarkably positive tone, with an overwhelming majority of articles in favour of nanotechnology within a general frame of benefits outweighing risks. Thus, in several European countries and in the US, nanotechnology mass media coverage has been largely positive, and has confirmed the prediction that risks only began to appear in later years, after an initial hype stage.

However, all of the studies described thus far have been from Anglo-Saxon or Northern European countries. Yet, intra-European differences can be substantial in terms of the structure of their public spheres and the actors involved. It is then an open empirical question as to whether the aforementioned pattern in the news coverage applies to other countries. For example, Bauer and Gaskell (2002a) analysed the evolution of the biotechnology controversies in Europe, and noted certain differences across European countries in terms of both public attitudes and mass media coverage. More specifically, Southern European countries exhibited different patterns compared to the rest of Europe, for example in the

cultural factors underpinning their opposition to biotechnology ¹, and in their mass media coverage.

The extant literature suggests then that the Spanish press coverage will display the typical attention cycle of a bell shaped curve of salience: with in terms of second level agenda setting an initial focus on benefits and economic opportunities, followed by an increased attention on risks and controversies. This study tests such predictions by analysing longitudinal data, applying social representations theory as a theoretical framework, in particular the notions of anchoring and codification, which account for differences in the public sphere not only in terms of structural differences but also in terms of cultural repertoires (or social representations). This is operationalized through the application of latent semantic analysis (Landauer, Foltz and Laham, 1998; Foltz, 1996). This methodology will be discussed in the next section together with the theoretical contribution of the social psychological theory of social representations.

3 Theoretical and Methodological Framework

Most of the studies described in the previous section are consistent with second level agenda setting research (Ghanem, 1997; Kim, Scheufele and Shanahan, 2002; McCombs and Ghanem, 2001) that shifts the focus from what the media emphasizes, to how they describe people or objects in the news.

Similarly there is a line of research that concentrates on the idea that in the public sphere there are competing definitions of a new technology, within what is a complex game played for semantic control (Gaskell, Bauer and Durrant, 1998). These studies refer to social representation theory (Moscovici, 2000), which represents a fertile paradigm for the study of the evolution of the biotechnology debate within European public opinion (Bauer and Gaskell, 2008). Through the notions of anchoring and objectification, studies informed by social representations theory have highlighted the emergence in the public sphere of semantic and semiotic configurations relating to a new technology. The purpose of anchoring is classification; it is the process through which the new social object is fully incorporated into pre-existing structures of knowledge, which are simultaneously affected by the new representation (Moscovici, 2000). Another important function of anchoring is to name the new object of representations. Through the process of anchoring, representations become codes of interpretation and action. Objectification, in turn, is the process through which actors seek to make what is abstract concrete by associating it, for example, with metaphors, images and symbols.

Within this context, newspapers' coverage has been analysed using correspondence and cluster analysis computed by text analysis software, which creates semantic maps based on recurrent clusters of words. These semantic maps are

¹ Adopting the label of blue resistance to biotechnology, a traditionalist approach, and green resistance, a modernist approach, Southern (and Catholic) European countries like Italy, Spain and Portugal having relatively more blue resistance than Northern (and Protestant) European countries, in which prevails green resistance (Nielsen, 1997).

often interpreted as core elements of social representations: examples of the use of such maps include Castro and Gomes (2005) on the representation of genetically modified organisms (GMO) in the Portuguese press; Wagner et al. (2002) on biotechnology in the European press; and more recently Gauthier (2009) on social representations of risk in food irradiation in Canada. Thus, rather than classify articles using a favourability index (e.g. Janis and Fadner, 1965) or using a risk-benefit framework, this approach aims to reveal core components of the social representations of new technologies.

Semantic analysis with a network approach has also been applied in order to reveal implicit frames (Hellsten, Dawson and Leydesdorff, 2009). In the latter approach, vocabularies of an article are considered particular instantiations of larger repertoires constituting of a set of vocabularies that recur over time. Identifying these involves looking into the process of codification of an issue in a collection of texts. Semantic analysis of texts is a powerful tool; however, it has its limitations, which emerge from the way it considers meaning generation in humans. Awareness of these shortcomings is necessary in the process of interpretation and particular attention should be paid to the distinction between implicit and explicit meanings. Semantic analysis refers to how meaning units denote one another within a system of relations (Morris, 1955). Therefore, any semantic analysis focuses on the identification of relationships between concepts in terms of their explicit meanings and does not go beyond the explicit meaning of a text (Popping, 2000; Roberts, 2000; Krippendorff, 2003). The outcomes of cluster or multiple-correspondence analyses applied to texts are taxonomical categories. A taxonomy organizes knowledge in terms of a hierarchical structure where items stand for the concepts they denote. The words in a cluster are all deemed equivalent in terms of importance or weight and they are considered in terms of their frequencies and co-occurrences with other words in an ad hoc context constituted by the segmentation strategy (a sentence, a paragraph, etc.). This is also the main limitation in carrying out a latent semantic analysis or LSA (Landauer, Foltz and Laham, 1998). Analysis of this kind does not make use of word order, nor thus of syntactic relations or logic, or of morphology.

Hence, identifying semantic clusters does not fully account for the meaning of a text, which is also determined by its morphology and argumentative structure. In this sense, semantic analysis deals with manifestations of denotative meanings, and not with metaphorical and connotative configurations. The latter are the objects of study of a semiotic rather than semantic analysis (Chandler, 2002) and, according to the previous definitions, such semiotic analysis would reveal the process of objectification of new technology. However, LSA can reveal elements of the semantic space that are not instantiated in any particular text but are part of a longitudinal process of codification that defines repertoires. A semantic longitudinal analysis of a large corpus of articles will reveal semantic clusters that are stable in newspapers codification of a given issue. Another way to describe this is that such analysis reveals the meaning of words as a kind of average of the meaning of all the passages in which it appears, and the meaning of a cluster as a kind of average of the meaning of all the associated words it contains. The semantic categories obtained in the early years of coverage will

be interpreted as elements of a process of anchoring in SRT. Hence, by using LSA, this study aims to specify the processes of anchoring and codification of nanotechnology in the Spanish national press.

4 Methodology

Following from what has been said above, the research questions of this study are the following:

1. Does the Spanish national press follow the same attention cycle of other countries in relation to nanotechnology, being characterized by a bell-shaped salience curve?
2. Does the Spanish national press display an overall positive coverage of nanotechnology?
3. Which themes are present in the representation of nanotechnology in the Spanish national press? And, complementary to the previous question: Does the Spanish press display an initial emphasis on economic potential followed by an increasing salience of risks?

Against the theoretical background laid out above, a quantitative semantic analysis of the anchoring and codification processes of nanotechnology in Spain was conducted. The aim was to highlight the specificity of Spanish media coverage and to test the patterns of coverage as described in the scientific literature. Clustering of elementary contexts (EC) was carried out as an operationalization of identifying taxonomies and classes of a semantic analysis. The software used was T-Lab 7.2. T-Lab software is an all-in-one set of linguistic and statistical tools for text analysis, which can be used in research fields like semantic analysis, content analysis and text mining (Lancia, 2005). T-Lab has tools for: word co-occurrence analysis (computation of word associations, comparisons between word pairs, co-word analysis and concept mapping, sequence analysis, concordances); thematic analysis of the context units (modelling of emerging themes, thematic analysis of elementary contexts, sequences of themes, key contexts of thematic words, thematic classification of documents); and comparative analysis of two or more corpus sub-sets: (specificity analysis, correspondence analysis, multiple correspondence analysis, descending cluster analysis).

5 Corpus

The corpus contains newspaper articles published between January 1997 and August 2010 for a total of 646 articles (32,477 lemmas) from three main Spanish newspapers of different political orientation: El Pais (2,012,000² daily copies on average, left-leaning); El Mundo (1,329,000 daily copies on average, moderate);

²Source of circulation data: Asociacion para la Investigacion de Medios de Comunicacion (AIMC).

and ABC (751,000 daily copies on average, right leaning). Articles were collected from the Lexis-Nexis electronic database of international newspapers. The sampling criteria was the presence of one or more of several keywords such as nanotecnologia, nanociencia, nanoparticulas, nanoestructura and 'manipulacion atomica' and ingeniera atomica'.

The selection of El Pais, El Mundo and ABC does not guarantee a representative sample of Spanish national newspapers in the strict sense as these newspapers are all elite or broadsheet papers. There is a reason for this methodological choice, however: they are each newspapers of record, they are the three best-selling newspapers in Spain, and they represent the three main political views widely defined as centrist, left-leaning and right-leaning (without being any particular political party's newspaper). The common distinction between elite newspapers and tabloid press is very much based on the UK and US media. In Spain, like in other European countries, there are no tabloid newspapers that share the characteristics of, for example, The Sun or The New York Post. In Spain, the best-selling newspapers are what in UK or US are considered elite newspapers. In other words, the selected newspapers are by and large the most influential national newspapers in Spain, not only for their prestige but also because of their wide circulation. Moreover, the scientific literature in public understanding of science suggests that opinion formation in the public sphere on scientific and technological issues is usually led by elite newspapers (Bauer, 2000; Anderson, Allan, Petersen and Wilkinson, 2005).

6 Procedure

All articles were tagged so as to indicate the newspaper of origin and the year and month of publication. The first step in analysing the texts is to perform a lemmatization. Lemmatization involves the reduction of corpus words to their respective headwords (lemmas). In the linguistic dictionaries consulted, every entry corresponds to a lemma that generally defines a set of words with the same lexical root (or lexeme), and that belongs to the same grammatical category (verb, adjective, etc.). As a rule, lemmatization entails that verb forms are taken back to the base form, nouns to the singular form, and so on. T-Lab allows for automatic lemmatization for several languages, including Spanish. Automatic lemmatization was performed with a custom dictionary, thus the vocabulary of the corpus was checked in order to disambiguate homographs. Two or more words are homographs when they have the same graphic form (i.e. they are written in the same way) but have different meanings.

An important decision regarded the text segmentation strategy, and therefore the selection of the unit of context used in the analysis. Three options were available for three different elementary contexts (EC), which are the contexts in which word co-occurrences manifested. The three options that presented themselves were to use sentences, chunks or paragraphs as elementary contexts. T-Lab considers chunks as elementary contexts to be every sequence of words interrupted by a full stop and a carriage return that are fewer than 400 characters

long. A paragraph is defined as an elementary context ending with punctuation marks and impression of the return key to a maximum length of 2000 characters. The middle segmentation option chunk was considered the best unit of analysis as it gave the average length of articles of around 1800 characters. Both sentence and paragraph would have resulted in a too small or a too large unit of analysis respectively.

The analysis proceeds in the following steps. First, a data table composed of context units by lexical units with presence/absence values is constructed. Secondly, a TF-IDF (term frequency-inverse document frequency) normalization and scaling of row vectors to unit length (Euclidean norm)³ is applied, as developed by Salton (1989). T-Lab then clusters context units using as its measure cosine coefficients, and applies the method of bisecting K-means (Steinbach, Karypis and Kumar, 2000; Savaresi and Boley, 2001). For each of the obtained partitions, a contingency table of lexical units by clusters is constructed, and a chi square test is applied to all the intersections of the contingency table. The last step is to perform a correspondence analysis of the contingency table of lexical units by clusters. Because the algorithm used (bisecting K-means) produces a hierarchical clustering, there are several potential solutions in terms of cluster partitions. Thus, another important decision required regards the maximum number of clusters to obtain. This determines the maximum threshold of clusters obtainable from the analysis. However, allowing for, say, 20 clusters does not necessarily produce 20 clusters because the algorithm will always minimize the final number of clusters in its output. The risk is using a threshold too low and therefore to miss clusters, if these are in number higher than the limit; however, carrying out the analysis with different limits and comparing leads to stable outcomes. After testing with two different limits, 10 and 20, the same results were obtained in terms of number of clusters indicating a stable solution. After computing the quantitative semantic analysis, the meaning of the semantic clusters was interpreted by going back to the articles text fragments corresponding to the lemmas of the highest value.

³TF-IDF (term frequency-inverse document frequency) formula is as follows:

$$w_{ij} = tf_{ij} \times idf_i$$
$$w_{ij} = tf_{ij} \times \log \frac{N}{df_i}$$

where:

tf_{ij} = number of occurrences of the i term in j document

df_i = number of documents containing i

N = total number of documents

Term Frequency (tf_{ij}) value can be normalized as follows:

$$tf_{ij} = (tf_{ij}) / \text{Max} (f_{ij})$$

where:

$\text{Max} (f_{ij})$ = the maximum frequency of i (any term) in the j (document)

7 Analysis

As a first step, word associations in the entire corpus were explored focusing on the following terms: nanotecnologia, nanobots, nanociencia, nanoparticulas, aplicacion. Simple associations showed that nanotechnology-related words were mainly associated with words related to the domains of government, business, medical and commercial applications. The next step was to carry out a thematic analysis of elementary contexts as described in the previous section. The output was a representation of the corpus content through a small number of significant thematic clusters. In the third step, thematic clusters were crossed with the main variables (newspapers and year of publication) in order to obtain their relative distributions. The results were a longitudinal dimension of thematic clusters, their relative weight in each year, and their relative importance for each newspaper. After considering the entire corpus, a further analysis focused on the thematic cluster socio-economic impact that was used as a new corpus. On this new corpus a similar analysis was conducted, in order to obtain sub-themes in this semantic cluster. For this new corpus, however, the unit of context was selected differently. Instead of using chunks as the text segmentation unit, the unit of context this time was the sentence. The rationale behind scaling down from the chunk to the sentence was that the new corpus from which the thematic cluster is obtained is constituted by chunks of text rather than entire articles. Hence, with a chunk as unit of analysis, the unit of segmentation for the automatic cluster analysis must be smaller, meaning therefore a sentence rather than a chunk.

Figure 1 shows the distribution of articles on nanotechnology in the selected newspapers between 1997 and 2009. The trend is common to other studies and confirms the attention issue cycle of an initial growing interest with a peak during the period 2006-2008 and a subsequent decline. Considering the first research question (p. 6) regarding the attention cycle, this data suggests a typical bell-shaped curve of salience in the Spanish national press. This is true for all three newspapers, but is limited to one year in the downward trend. Figure 2 presents the output of a thematic analysis of elementary contexts that converged on a structure of five thematic clusters. This part of the analysis relates to the following research questions about the nature of the coverage: does the Spanish national press display an overall positive coverage of nanotechnology?; which themes are present in the representation of nanotechnology in the Spanish national press?; does the Spanish press display an initial emphasis on economic potential followed by an increasing salience of risks? The outcome of the thematic analysis of EC will help test whether the overall orientation of the Spanish national press towards nanotechnology was positive, and will present a more detailed picture of the nature of the coverage in terms of semantic domains in which nanotechnology has been discussed. A three-dimensional representation was selected because of the high number of clusters generated, and in order to better identify their proximity relationships within the factorial spaces. Each cluster received a label based on the qualitative interpretation carried out by going back to elementary contexts grouped in each theme. Clusters were interpreted

Figure 1: Number of articles on nanotechnology by year and newspaper of publication (N=636).

as follows: NanoNatioPolicy, national initiatives (mainly governmental) to invest and promote nanotechnology in Spain; Empresa, private investments and financial market of nanotechnology; NanoEvents, a cluster on nanotechnology events such as conferences, educational programs and innovations; Socioeconomic, impact of nanotechnology on social and economic relations; and Nanoapp, nanotechnology's applications in the production of new materials and medical research.

Considering the relative weight of each thematic cluster in the overall number of elementary contexts, the largest of the three clusters are those on public events on nanotechnologies (NanoEvents 24.9), followed by applications of nanotechnology ('NanoApp,' 23.7) and on their social and economic impact ('Socio-economic', 22.5). The other clusters are in the following order of 'weight:' a cluster on business opportunities associated with nanotechnology ('Empresa,' 14.8); and one concerned with national initiatives and governance of nanotechnology ('NanoNatioPolicy,' 13.4). The relative weight of each thematic cluster in three newspapers⁴ differs between two clusters. In ABC, the theme 'NanoEvents' is more present than El Pais and El Mundo (36.8, 25.3 and 19.2 respectively). More interestingly, in ABC the thematic cluster on the social and economic impact of nanotechnology is almost absent (3.5) compared to El Mundo (25.3) and El Pais (24.6). ABC is the most right-wing of the newspapers selected, hence it would appear that concerns about nanotechnology were more salient in the

⁴Such computation considered the number of EC in each newspaper, which was then was normalized. It uses the number of EC tagged from each newspaper so the percentage is based on number of EC of each cluster in a newspaper divided the total number of EC of a newspaper.

Figure 2: Clusters obtained from the detection algorithm of elementary contexts from the overall corpus (lemmas, N=32,477).

moderate El Mundo or the left-leaning El Pais. Taking into consideration the smaller number of articles on nanotechnology in ABC compared to the other two newspapers, it might not be an issue of political orientation. Another plausible explanation is that ABC is more focused on national issues and less likely to contain discussion on science and technology.

Regarding the distribution in the semantic space of the thematic clusters, two main spaces group the clusters. The first space is shared by three thematic clusters (NanoEvents, NanoNatioPolicy, and Empresa.) These clusters are all linked by the fact that they are themes about economic potential, related governance issues of nanotechnology, and public events that discuss these topics or present public or private initiatives. Their proximity in the factorial space stands for thematic closeness. The thematic cluster NanoNatioPolicy refers to national governance of science and technology, in other words the Spanish governments initiatives on nanotechnology (e.g. El Pais 22/03/2006, 20/02/2002, 12/10/2008; El Mundo 14/06/2006, 28/06/2008, 15/01/2009; ABC 10/09/2009). It denotes news on the creation of research agencies, education programmes, clusters of universities and providing funding opportunities. It also refers to funding opportunities in Europe from supranational entities such as the European Commission and the European Research Council. The thematic cluster Empresa is closely related, and represents news on collaboration between

private business and research universities. Empresa includes descriptions of firms involved in nanotechnology and cases of patents related to nanotech. Its closeness to the previous thematic cluster is due to the frequent mention of government funding opportunities for joint research between firms and universities (e.g. El Pais 21/02/2010, 27/09/2009, 08/04/2007, 7/09/2006, 16/10/2005; El Mundo 28/05/2009, 01/03/2009, 15/09/2007; ABC 7/09/2009; 09/04/2008, 11/12/2007). In the same factorial space, although relatively more distant from the previous two clusters, is a thematic cluster that contains news of public events related to nanotechnology (NanoEvents.) Nanoevents contains both scientific conferences and meetings between research actors, policy-makers and private investors. The latter are often events advertising the promise of economic growth through nanotechnology-based innovations (e.g. El Pais 22/03/2006, 19/09/2001, 22/10/2008; El Mundo 18/07/2008, 26/03/2009; ABC 18/07/2009, 10/10/2007).

The remaining two thematic clusters, applications (NanoApp) and socio-economic impact (Socio-Economic), share one factorial space. The first thematic cluster focuses on applications of nanotechnology (NanoApp), and refers to descriptions about current and potential innovations derived from applying nanotechnology to the production of medical or other goods. It represents the popularizing of nanotechnologies potential in which this technology is often represented as the source of future utopian scenarios (e.g. El Pais 5/08/2001; El Mundo 28/02/2010, 17/01/2005; ABC 04/02/2010, 11/10/2007). A significant amount of attention is paid to medical applications of nanotechnology (e.g. El Pais, 28/09/2008, 03/04/2009) with frequent reference made to the sci-fi movie *Fantastic Voyage* (e.g. El Pais 22/10/2008). The Socio-economic thematic cluster is in the same factorial space as the previous thematic cluster on applications. It also contains concerns over the potential consequences of nanotechnology for the current social and economic order. It entails four subthemes (labelled Progress, Nanorevolucion, HumanEnhanc and Pandora), which will be discussed in more detail in the last part of the next section. Considering the data described above, answers to two of the research questions previously stated can now be provided: nanotechnology coverage in the Spanish national press was overall largely positive, as indicated by the relative weight of thematic clusters; in addition, the automatic thematic analysis identified the semantic domains that characterized the Spanish coverage.

8 Anchoring and Codification

In operational terms, answering the research question on the nature of the coverage relies on the identification of thematic clusters in the early years of coverage. As discussed previously, the current literature suggests an initial emphasis on economic benefits followed by an increasing salience of risk related themes. Within the context of this study, this prediction will be verified or refuted. In other words, in the longitudinal evolution of coverage, are the thematic clusters related to economic benefits initially predominant within the overall corpus? And do those related to risks increase in later years?

Figure 3: Relative weight of thematic clusters in the corpus across years (EC, N=8582)

Considering the longitudinal dimension (Figure 3), it is possible to plot their evolution in terms of their relative weight in each year during the 12-year period 1997-2009. The early years of coverage display a small number of articles, and that must be taken into consideration in the interpretation of the relative weight of thematic clusters. Between 1997 and 2000, the data suggests a substantial presence of themes related to public events (NanoEvents), to nanotechnology applications (NanoApp), and to their socio-economic impacts (SocioEconImpact). In the year 2000, two more themes appeared in the newspapers coverage: one related to government policies (NanoPolicy), and one on private investments and joint collaboration between firms and universities on nanotechnology research (Empresa).

The most interesting result is that the Socioeconomic theme becomes progressively less salient (with some prominence between 2001 and 2003) and after 2007 is not prominent at all. This is interesting because data suggests initial attention should be paid to the social and economic impact of nanotechnology, including articles about its potential controversies, which decrease over time. Contrary to previous studies in which risks of new technologies occurred in the later years of newspapers coverage following an initial overwhelmingly coverage on benefits (e.g. Nisbet, Brossard and Kroepsch, 2003; Gaskell, Eyck, Jackson and Veltri, 2005; Scheufele and Lewenstein, 2005), in the Spanish press data indicates an opposite pattern. More precisely, aspects of danger and risks related to this technology did not gain momentum.

The anchoring of nanotechnology, if we consider the first years of coverage,

Figure 4: Clusters obtained from the thematic clustering of elementary contexts from the socio-economic corpus (N=13759).

did include a controversial account of this technology, together with an account of its potential applications. However, the later codification of nanotechnology marginalized such controversial accounts, and focused more on aspects of governmental policies on nanotechnology innovation (NanoPolicy), on the fostering of collaborations between firms and universities (Empresa), and on the public events that substantially promoted the first two themes (NanoEvents). In later years, the codification of nanotechnology is mostly devoted to governance and business issues and events aimed at promoting these. Attention to the applications of nanotechnology cluster (NanoApp) is rather stable across the period, indicating relatively consistent reporting on scientific and technological discoveries.

Figure 4 presents the thematic sub-clusters derived from the Socioeconomic cluster after using it as a corpus for further analysis. Applying almost the same analysis carried out for the entire corpus, four themes were identified with their relative weight in the new corpus: HumanEnhanc (30.9), Nanorevolucion (26.6), Pandora (23.9) and Progress (18.4). In one factorial space there are two thematic clusters labelled HumanEnhanc and Pandora. The first stands for the use of nanotechnology to overcome human limitations such as diseases, aging and mortality. This theme is largely positive and focuses on potential benefits of nanotechnology (for example El Pais, 8/10/2003; and ABC, 5/06/2008, El Mundo 29/12/2004). Concern is clearly expressed in the other thematic cluster Pandora

present in the same factorial space. It refers to risks concerning a Pandoras Box scenario caused by applications of nanotechnology; it includes an expression of concern on overcoming mortality, and there are also references to nano-bots and nano-killer machines (e.g. *El Mundo* 27/01/2005, 19/05/2003). The latter are so-called technology out of control scenarios, in which nanotechnology is unleashed creating havoc on earth. Prince Charless statement on the risks of nanotechnology (for example *El Mundo* 12/07/2004) also fits into this theme because of the grey-goo scenario. The controversies within the dystopian discourse refer often to two aspects: nano particles as the main agent of risks (for example, *El Pais*, 13/12/2006), and the danger of military applications of nanotechnology (e.g. *El Mundo* 6/10/2003).

A third thematic cluster labelled Nanorevolucion is characterized by issues concerning the social and economic impact of nanotechnology. On the one hand, nanotechnology is portrayed as a revolution and a breakthrough, a radical change in the system of production and socio-economic order of contemporary advanced societies (for example, *El Pais* 18/01/2006 and *El Pais* 19/06/2002, *El Mundo* 05/06/2008), and an opportunity for developing countries (e.g. *El Mundo* 07/01/2005). In this scenario the economic impact is considered especially relevant (e.g. *El Mundo* 12/05/2006) in the sense that there is an identification of nanotechnology as the foundation of the next economic system (e.g. *El Pais* 8/12/2003, *ABC*, 07/09/2009). This underpins the use of the term the new industrial revolution. The last thematic cluster Progress shares the same factorial space of Nanorevolucion. However, this theme emphasizes to a greater extent the positive aspects of nanotechnology in terms of general progress, with a particular stress placed on it being a new energy source, and on innovative nano-engineered materials representing a solution to environmental problems to (e.g. *El Mundo* 17/09/2005 and *El Pais* 13/06/2002).

The final chart (Figure 5) presents the longitudinal evolution of sub-thematic clusters identified within the theme on the socio-economic impacts of nanotechnology. The relative weight of the sub-theme related to risks, Pandora, was highly prominent in the first years of coverage, losing salience afterwards (when, on the other hand, a positive theme like Nanorevolucion gained more space), and finally slightly regaining salience between 2007 and 2009. All other sub-thematic clusters were rather stable in relative weight. Positive sub-themes (HumanEnhanc, Progress and Nanorevolucion) on average constitute over 75 of this theme. Hence, the controversies surrounding nanotechnology constitute only a third of the thematic cluster that deals with the societal impact of nanotechnology. It confirms the findings of a largely positive coverage and a pattern of early prominence of risks obtained from the overall corpus.

9 Discussion

Nanotechnologies coverage in the Spanish national press only partially confirms the findings from previous studies. Regarding the attention issue cycle, Spanish coverage is not an exception to other countries: an initial upward trend is

Figure 5: Relative weight of thematic clusters in the corpus across years (EC, N=8582)

followed by a reverse in salience later on. However, having data only for one year when the downward trend starts does not exclude the possibility that there might be a fluctuation rather than a long-term decrease in salience, although such a decrease is manifested in all three newspapers. Hence, it is not possible to confirm that the attention cycle followed the predicted bell-shaped trend, even though the data and findings are highly compatible with this expectation.

The analysis of thematic clusters indicates an anchoring and a codification process dominated by aspects of economic development and business opportunities, science governance/national policies, public events and description of potential applications of nanotechnology. These dominant thematic clusters in the anchoring and codification of nanotechnology in the Spanish national press confirm the predictions of previous studies. Nanotechnology appeared to be commodified in discursive terms (Best, 1994), expressed mainly in terms of economic relationships. In addition, as in the case of biotechnology as shown by Bauer and Gaskell (2002), the Spanish press devoted attention to the European Union in its double role of regulator and funder. Controversial aspects of nanotechnology were salient in the initial coverage of the Spanish press, becoming rather marginal in more recent years. This is therefore a different pattern from the one suggested by the studies previously discussed (Gaskell, Eyck, Jackson

and Veltri, 2005; Scheufele and Lewenstein, 2005; Anderson, Allan, Petersen and Wilkinson, 2005; Friedman and Egolf, 2005). In Spain, controversies over nanotechnology did not gain momentum or resonate (Gamson and Modigliani, 1989), and most likely were not endorsed by any social actor. One possible explanation is that historically Spanish environmental groups have been not very active on the potential risks of new technologies, and it is only very recently that they have gained influence in the public sphere (Jimnez, 2007).

These findings lend support to agenda building and second level agenda building approaches that implicitly consider the contextual role of the structure of the public sphere and the relevant actors in determining the overall level of coverage. Spain might have a different structure e.g. a lack of strong environmental groups or it may be simply that for this particular issue, the dynamic between the actors involved in the agenda building process is different. As discussed previously, both the social representations and implicit frames approaches can be considered similar to the studies that focus on second level agenda setting. However, taking into account the presence and role of different actors in the shaping of media representations of a new technology will by necessity also require an agenda-building approach (Lang and Lang, 1983). Both approaches can be combined, as shown by Fahmy, Wanta, Johnson and Zhang (2011), who have outlined a second level agenda building model. A study defined in this way will better capture the diversity of the public sphere and will be complementary to a horizontal analysis like the one presented in this study.

Besides the need to expand the theoretical framework used to consider the agenda building activities of actors present in a given countrys public sphere, methodological considerations also arise here concerning the application of quantitative text mining techniques. The methodology employed in this study enabled the answering of the research questions posed, and simultaneously provided the possibility of identifying the nature and evolution of the dominant thematic semantic cluster. There is room, however, to expand on its potential, while remaining aware of its intrinsic limitations. There are well known limits of quantitative analysis techniques in detecting cultural nuances of national coverage, particularly when it comes to figurative language. Positive symbols can play an important role in the collective imaginary. Moreover, images do play an important role in mobilizing enthusiasms and fears see for example the studies on their impact in the case of biotechnology Wagner et al. (2002), and, more related to this study, on nanotechnology as in Landau, Groscurth, Wright and Condit (2008). Complementing a latent semantic analysis by taking into account the figurative dimension will therefore greatly enrich the depth of analysis. But it will also face new methodological challenges too. Nevertheless, quantitative methods of text analysis are continuously moving forward, with some interesting developments.

10 Conclusions

This study has contributed to the study of nanotechnology's mass media representations by exploring the Spanish case, and by assessing whether recurrent patterns in the media coverage of novel scientific and technological issues identified in the scientific literature obtain in this case. At the same time, it has introduced a method of quantitative text analysis informed by latent semantic analysis, the notion of anchoring from SRT (Moscovici, 2000), and the concept of codification (Hellsten, Dawson and Leydesdorff, 2009).

The Spanish case suggests caution in assuming that patterns in mass media coverage of new technologies derived from well-studied countries also apply in other countries, even those within Europe. Intra-European differences are substantial enough that many countries may be only superficially similar. The quantitative methodology applied in this study could be applied in cross-national comparisons at least at the analytical level of identifying semantic domains that could be complemented by more in depth qualitative analysis. This analysis considers only one part of the public representation of nanotechnology in Spain. Other complementary studies should include the actors involved in the public discussion on nanotechnology in Spain. In other words, they should combine a second level agenda setting study like this one with an agenda building study as discussed in the previous section. Moreover, this could be expanded to include all three research areas of a typical social representations study based on the public sphere triangle (Bauer and Gaskell, 2002b): the citizens public perceptions, the mass medias representations, and policymakers discourses.

Further research might also consider different kinds of mass media. Of particular interest in this regard are new media sources, in particular the potential role of social networking sites as conversational hubs. Recently, online media sources have begun to be considered for their potential for a public forum for discussions and conversations among individuals, in the forms of mailing lists, chat rooms and other discussion forums (Triunfol, 2004; Rogers and Marres, 2000). For example, Lewis, Kaufman, Gonzalez, Wimmer and Christakis. (2008) have discussed the potential use of Facebook data in the context of merging virtual and real life variables and the identification of specific network behaviours.

Regarding the future of nanotechnology in Spain, the minor role played by controversial aspects of nanotechnology in mass media coverage is not necessarily good news for its supporters. Although overwhelmingly positive coverage might generate a spiral of silence (Noelle-Neumann, 1990), as dominant press coverage experienced as peer pressure might push dissenters into silence, the eruption of controversy is still possible. Data from a recent Euro barometer survey (Special Euro barometer 341, 2010) indicates that Spanish public attitudes towards nanotechnology tend to be positive. However, there are large percentages of don't know answers, and a consistent minority of worried Spanish citizens. Moreover, exogenous influences on the population pushing it to assume more critical views might come from other European countries or institutions, and lead to the emergence of controversy in the Spanish mass media and public.

11 References

1. Anderson, A., Allan, S., Petersen, A. and Wilkinson, A. (2005). The framing of Nanotechnology in the British Newspaper Press. *Science Communication* 27, 200-220.
2. Bauer, M. W., and Gaskell, G. (2002a). *Biotechnology: The making of a global controversy*. Cambridge: Cambridge University Press.
3. Bauer, M. W., and Gaskell, G. (2002b). Researching the public sphere of biotechnology. In Bauer, M. W., and Gaskell, G. (eds). *Biotechnology: The making of a global controversy* (1-17) Cambridge: Cambridge University Press.
4. Bauer, M. W., and Gaskell, G. (2008). Social Representations Theory: A Progressive Research Programme for Social Psychology. *Journal for the Theory of Social Behaviour*, 38, 4.
5. Berkowitz, D. (1992) Who sets the media agenda? The ability of policymakers to determine news decisions. In Kenamer, J. D. (ed). *Public opinion, the press, and public policy*. Westport, Conn: Praeger.
6. Best, S. (1994). The Commodification of Reality and the Reality of Commodification: Baudrillard, Debord, and Postmodern Theory. In Kellner, D. (Editor). *Baudrillard: A critical reader* (41-68). Oxford: Blackwell.
7. Castro, P. and Gomes, I. (2005) .Genetically Modified Organisms in the Portuguese Press: Thematisation and Anchoring. *Journal for the Theory of Social Behavior* 35, 118.
8. Chandler, D. (2002). *Semiotics*. New York, Routledge.
9. Curtin, P. (1999). Reevaluating Public Relations Information Subsidies: Market-Driven Journalism and Agenda-Building Theory and Practice. *Journal of Public Relations Research* 11(1), 53-90.
10. Downs, A. (1972). Up and down with ecology - the "issue-attention cycle". *Public Interest* 28, 38-50.
11. European Commission Special Eurobarometer 341 (2010) Europeans and Biotechnology. Luxembourg: Publications Office of the European Union.
12. Fahmy, S. S., Wanta, W., Johnson, T. J., and Zhang, J. (2011). The path to war: Exploring a second-level agenda-building analysis examining the relationship among the media, the public and the president. *International Communication Gazette*, 73, 4, 322-342.
13. Foltz, P. W. (1996) Latent Semantic Analysis for text-based research. *Behavior Research Methods, Instruments and Computers* 28, 197-202.

14. Friedman, S. M., Egolf, B. P. (2005). Nanotechnology: Risks and the media. *IEEE Technology and Society Magazine*, 24, 5-11.
15. Gamson W A and Modigliani A. (1989) Media discourse and public opinion on nuclear power: a constructionist approach. *American Journal of Sociology*, 95, 1-37
16. Gaskell, G., Bauer, M. and Durant, J. (1998). The representations of biotechnology: policy, media and public perceptions. In Durant, J., Bauer, M. and Gaskell, G. (Eds,) *Biotechnology in the public sphere: a European source book*. London: Science Museum Press.
17. Gaskell, G., Ten Eyck and T., Jackson, J. and Veltri, G. (2005). 'Imagining Nanotechnology: Cultural Support for Innovation in Europe and the United States'. *Public Understanding of Science*, 14, 81-90.
18. Gauthier, E. (2009). Social Representations of Risk in the Food Irradiation Debate in Canada, 1986-2002. *Science Communication*, 32(3), 295-329.
19. Ghanem, S (1997). Filling the tapestry: The second level of agenda setting. In McCombs, M., Shaw, D.L. and Weaver, D. (eds). *Communication and Democracy: Exploring the Intellectual Frontiers in Agenda-Setting Theory*. Mahwah, NJ: Lawrence Erlbaum, 3-14.
20. Gorss, J.B. and Lewenstein, B.V. (2005). The Salience of Small: Nanotechnology Coverage in the American Press, 1986-2004. Paper presented at the 2005 conference of the International Communication Association.
21. Hellsten, I., Dawson, J., and Leydesdorff, L. (2009). Implicit media frames: Automated analysis of public debate on artificial sweeteners. *Public Understanding of Science*, 19(5), 590-608.
22. Janis, I. and Fadner, R. (1965), The coefficient of imbalance. In: Lasswell, H (Eds), *The Language of Politics*. pp. 153-69. Boston, MA: MIT Press.
23. Jimnez, M. (2007). The Environmental Movement in Spain: A Growing Force of Contention. *South European Society and Politics*, 12(3), 359-378.
24. Kim, S. H., Scheufele, D.A. and Shanahan, J. (2002). Think about this way: Attribute agenda-setting function of the press and the publics evaluation of a local issue. *Journalism and Mass Communication Quarterly* 79 (1), 7-25.
25. Kjærgaard, R. S. (2010). Making a small country count: nanotechnology in Danish newspapers from 1996 to 2006. *Public Understanding of Science*, 19, 80-97.
26. Kjølborg, K.L. (2009). Representations of Nanotechnology in the Norwegian Newspapers Implication for public participations. *Nanoethics*, 3, 61-72.

27. Krippendorff, K. (2003). *Content Analysis: An Introduction to its Methodology*. Beverly Hills, Sage.
28. Landau, J., Groscurth, C. R., Wright, L., and Condit, C. M. (2008). Visualizing nanotechnology: the impact of visual images on lay American audience associations with nanotechnology. *Public Understanding of Science*, 18(3), 325-337.
29. Landauer, T. K., Foltz, P. W., and Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.
30. Lang, G. E. and Lang, K. (1983). *The battle for public opinion: The president, the press, and the polls during Watergate*. New York: Columbia University Press.
31. Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., and Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks*, 30(4), 330-342.
32. McCombs, M. and Ghanem, S. (2001). The convergence of agenda-setting and framing. In Reese, S. D., Gandy, O. H., and Grant, A. E. (eds). *Framing public life: Perspectives on media and our understanding of the social world*. Mahwah, N.J: Lawrence Erlbaum Associates.
33. Morris, C. (1955). *Signs, language, and behavior*. New York, G. Braziller.
34. Moscovici, S. (2000). *Social Representations: Explorations in Social Psychology*. Cambridge: Polity Press.
35. Nielsen, T. H. (1997). Behind the color of No. *Nature Biotechnology* 15, 1320-1.
36. Nisbet, M.C., Brossard, D. and Kroepsch, A. (2003). Framing Science: The Stem Cell Controversy in an Age of Press/Politics. *The Harvard International Journal of Press/Politics*, 8, 36.
37. Noelle-Neumann, E. (1990) The Theory of Public Opinion: The Concept of the Spiral of Silence, *Communications Yearbook*, 14, 25687.
38. Popping, R. (2000). *Computer-assisted text analysis*. London, Sage.
39. Roberts, C. W. (2000). A Conceptual Framework for Quantitative Text Analysis on Joining Probabilities and Substantive Inferences about Texts. *Quality and Quantity*, 34, 259-274.
40. Rogers, R., and Marres, N. (2000). Landscaping climate change: a mapping technique for understanding science and technology debates on the World Wide Web. *Public Understanding of Science*, 9(2), 141.

41. Salton, G. (1989). *Automatic text processing: the transformation, analysis, and retrieval of Information by Computer*, Addison-Wesley, Reading, Massachusetts.
42. Savaresi S.M., Boley D.L. (2001) On the performance of bisecting K-means and PDDP, 1st SIAM Conference on DATA MINING, Chicago, IL, USA, April 5-7, paper n.5, pp.1-14.
43. Scheufele, D. A., and Lewenstein, B. V. (2005). The public and nanotechnology: How citizens make sense of emerging technologies. *Journal of Nanoparticle Research*, 7(6), 659-667.
44. Steinbach, M., Karypis G., Kumar V.(2000) A comparison of Document Clustering Techniques. Proceedings of World Text Mining Conference, KDD2000, Boston.
45. Triunfol, M. L. (2004). Dynamics of List-Server Discussion on Genetically Modified Foods. *Public Understanding of Science*, 13(2), 155-175.
46. Vandermoere, F., Blanchemanche, S., Bieberstein, A., Marette, S., and Roosen, J. (2010). The morality of attitudes toward nanotechnology: About God, techno-scientific progress, and interfering with nature. *Journal of Nanoparticle Research*, 12(2), 373-381
47. Wagner, W., Kronberger, N., Allum, N., De Cheveigne, S., Diego, C., Gaskell, G., Heinssen, M., Midden, C., Odegaard, M., Ohman, S., Rizzo, B., Rusanen, T., and Stathopoulou, A. (2002). Pandoras genes -images of genes and nature. In Bauer, M. and Gaskell, G. (Eds.), *Biotechnology- the making of a Global Controversy*. Cambridge: Cambridge University Press.