

FINDING SOCIAL AND POLITICAL PREJUDICES IN PUBLIC DISCOURSE?  
A CASE STUDY ON THE 2007 AND 2012 PRESIDENTIAL FRENCH CAMPAIGNS

Baider Fabienne  
University of Cyprus  
[fabienne@ucy.ac.cy](mailto:fabienne@ucy.ac.cy)

## Abstract

In this research we suggest that working on a journalistic corpus with specific softwares can help studying linguistic patterns and choices which are made on the basis of political affiliation or gender stereotypes. The software SEMY for instance gives semantic profiles semi-automatically, ANTCOnc gives useful KWIC abstracts and TERMOSTAT works on discourse specificities. Using all of these tools we found convergent striking asymmetries between female and male candidates in journalistic discourse (however conditionally) as far as our corpus dedicated to the 2007 and the 2012 presidential campaigns are concerned. Social gender' (i.e. stereotypical expectations about who will be a typical member of a given category) and / or political favoritism affect the representation of leadership in discourse and may affect in turn the readership, hence the electorate.

## Introduction

Work analyzing ideology in discourse (Baider & Constantinou 2013, Kolh 2011, Reutenauer *et al.* 2010; Valette 2010; Mussof 2008) showed the indexicality (indirectness) of the construction of prejudices in discourse i.e. *the lexical and stylistic choices* can give a racist or sexist semantic prosody to a text without the text being explicitly racist or sexist. Therefore *lexical specificities* characterizing a discourse can be a mean to detect ideological position. The purpose of this article is to show how indeed softwares such as ANTCOnc, SEMY or TERMOSTAT (Drouin 2003, Lesage *et al.* 1993) can help in finding the ideological positioning of discourses. We present different works based on journalistic corpora (2007 and 2012 election campaigns in France), and show that the results found in a semi-automatic way allow indeed a fairly good comparison of political profiles and campaigns. However the results and the interpretations have to be confronted with a close manual reading and with a close knowledge of the social context in order:

1. To verify whether the findings are plausible;
2. To decide between different interpretations of the lexical findings.

### 1. Lexical and semantics asymmetry in journalistic discourses

#### 1.1 Theoretical framework

In this paper we consider language as a form of social practice and we base our analysis on the following three assumptions which guide works in critical discourse analysis (Fairclough, 1989, 22):

- Language is a part of the society and not somehow external to it;
- Language is a social process;

- Language is a socially conditioned process, conditioned that is by other (non-linguistic) parts of society).

Our analysis takes into account the social reality of French politics, the French political culture and journalist practices since Fairclough (1995) defines CDA as the systematic exploration of :

often opaque relationships of causality and determination between (a) *discursive practices, events and texts*, and (b) wider *social and cultural structures, relations and processes*; (Fairclough1995, 132).

The aim of our discourse analysis is then to uncover the “implicit ideologies in texts” and to “unveil the underlying ideological prejudices and therefore the exercise of power in texts” (Widdowson, 2000), to point out ideology and its effects of domination” (Fairclough, 1995, 20) while working on written discursive practices:

(...) to investigate how such practices, events and texts arise out of and *are ideologically shaped by relations of power and struggles over power*; and to explore how the opacity of these relationships between discourse and society is itself a *factor securing power and hegemony* (Fairclough 1995, 133).

This means that the macro context where the discourse is set in its social reality has to be taken into consideration after having studied the micro level i.e. the text itself, collocations, usages of moods and modality, etc.

As far as presentation of data is concerned, we will follow the three dimensions as described in Fairclough (1989, 26-27) when studying and presenting our data:

- first we *describe* the data in order to identify the formal properties of the text and patterns in the text, this being done with the help of data mining tools such as TERMOSTAT;
- then we *interpretate* the data and identify the relationship between the journalistic texts and the social interactions during the campaign, which implies a knowledge of the political context;
- finally we suggest an *explanation* by identifying the relationship between those social interactions and the broad social context, in particular in this paper the gendered dimension of politics. We conclude on its potential social effects.

In van Dijk (2006 *et passim*) the overall strategy of ideology consists in a dichotomous positioning between several entities:

- positive presentation /action of our preferred candidate, of the party, our electorate, etc., emphasizing their good things, de-emphasizing their bad things,
- negative / action of the other champion candidate, the other party the other electorate, etc., emphasizing their bad things and de-emphasizing their good things.

Among different linguistic devices being used to construct such discourse, rhetorical structures can include figures such as comparison, metaphors, metonymies, irony, etc. as we observed in our data, especially in the 2012 corpus.

### 1.2 Previous research on the indexicality of ideological discourse

Vallette & Rastier (2006, 4) identified racist content in online documents by observing language choices. For instance as far as morpho-syntax is concerned, racist texts were characterized by an intensive use of evaluative adverbs and negative adverbs (nothing, never, much, too much, etc.); an intensive use of first person plural pronoun in order to build an opposition between <us> and the <other(s)>, a predominant usage of pejorative morphemes. Recurring themes were reflected in the lexical choice and the most common metaphors such as metaphors equating the Other as a human subspecies, an animal (*rat, monkey, cockroach*, etc.) or a disease. The coexistence of several of these phenomena, and of others not mentioned here, had for results the building of a text which has the illocutionary force and the impact of a racist text without being racist explicitly.

In the same line of thought, previous research on sexism in discourse noted that language choices can indeed build semantic oppositions and discourse asymmetry between the described entities without doing so in an obvious manner (Ochs 1992, Kitzing 2007, Goodwin 2011). Previous studies (Lazar 2005, Macaulay and Brice 1997, Penelope 1990) have shown that in many English-language textbooks most of the agents or doers are men and boys “whereas women and girls, if present at all, are often themes to which things happen” (Eckert and McConnell-Ginet 2003, 73).

Similarly Michard (1988) and Michard and Ribery (2008), working from scientific texts in French found that writing sexist texts is most of the time done in an indirect fashion. For instance female human beings were built as "non-agent" entity (2008, 174-184) by associating the nouns referring to women:

- with *automatic processes* as the ones describing *instruments* or *machines* or construction as *non-animated* because of the choice of verbs or nouns referring to them, hence the lack of agency in discourse of the entity described;
- with a constant evaluation of the entities constructed as non-agent in contrast with another entity. They noted a high number of adverbs and adjectives and the use of negative modifiers such as *few, a little, a few, less* collocated with the NPs under investigation, hence the lessening the autonomy of these nouns.

Combining those syntactic and lexical findings, Michard and Ribery concluded that a power relationship can be established discursively between the described entities. Indeed such an asymmetry may explain the functions and the qualities assigned to one or the other person within concrete social situations. The same discursive asymmetry was also noted in the discourse analysis relating to politicians (Kahn 1994, 1996; Sreberny-Mohammadi & Rossi 1996; Roncarolo 2000; Coulomb - Gully 2009 and 2012 ; Bertini 2007). Discourse practices and lexical choices play a critical role in the construction of the entities described as incompetent or in the opposite as being the perfect leader. These practices are too often perpetuating gender and ethnic stereotypes.

What is the most important to our work in these previous studies is the conclusion that ideological positioning is built indexically: how can we analyze (semi)- automatically such indirect positioning in discourse? The studies listed and described above point all to *lexical specificities* characterizing a discourse taking an ideological position. Therefore in order to assess the lexical environment of each NP understudy, we suggest using tools which can process a considerable amount of data, allowing quickly the identification of lexical choices and therefore help tremendously a quantitative and a qualitative reading.

### 1.3 Lexical specificities and semantic recurrence

In the recent years, several free softwares have been made available to extract lexical and syntactic patterns characteristics of the data under investigation.

We chose the software TERMOSTAT (Drouin 2003, Lesage *et al.* 1993)<sup>1</sup> to examine both, the most *frequent* and the most *specific* lexical categories for each NP under study (Aubry, Hollande and Sarkozy), criteria which are the basis of our hypothesis. We can then decide which categories are the most relevant to a qualitative reading. Moreover, the concordancer ANTCO<sup>2</sup> provides us with paragraphs where the NP of each candidate appears in the article. A manual and selective reading is then possible. It also provides the most frequent collocates and this allows a quick reading of discursive stereotypes.

With TERMOSTAT any corpus studied is compared with a ‘neutral’ journalistic discourse called ‘Reference corpus’ (28 million of occurrences from *Le Monde* (2002) integrated to the Internet platform (Drouin 2003)<sup>3</sup> ; this allows a thorough quantitative reading of a corpus and enables comparison between corpora. Thanks to its calculation regarding the degree of specificity, TERMOSTAT gives several lists of terms considered the most salient in our data, whether because they are the most frequent or whether they are the most specific to the discourse. Frequency and specificity are the most used criteria for corpus analysis, as previous studies on ideology and discourse have noted (Michard and Ribery 2008, Valette and Rastier 2006). Especially lexical units such as derogative adverbs and negative adjectives were thought as the most relevant to ideological loading (Reutenauer *et al.* 2010 ; Valette 2010) and were more likely to be found when referring to “dominated entity”. TERMOSTAT identifies such grammatical categories and we can quickly see which nouns, adjective or adverbs are the most used in the data.

On the other hand ANTCO enabled us to carry out a qualitative analysis since we can isolate blocks of text where the NPs are present. In particular it allow us to check whether some NP are more often associated than other NPs with evaluative lexical units (Michard & Ribéry and Valette’s hypothesis) when looking at the categories <collocates> and <cluster>. Quickly we can verify a higher number of adverbs, adjectives and negative modifiers such as *un peu*, *quelques*, *moins* collocated with NP under investigation. As well the important function Key Word in Context (Kwic) allows us to see all the sentences in which the NPs under investigation are used. A manual reading can then point out with which *syntactic processes* these NPs are associated the most (Subject of a sentence for instance) and the agency of a unit (among other thematic roles, cf. Siewierska 1991). Therefore, with these abstracts we can observe whether NPs are constructed in discourse as a non-animated entity or at least as less in control of the verb (if they are they most of the time Subjects of a verb but non – Agents for instance).

Indeed in our previous research we suggested that considering those discourse practices, i.e. focusing in particular on noun phrases (NPs) referring to female and male politicians, could be a way of assessing the ‘linguistic profile’ of politicians in journalistic discourse and in particular their agency. Agency is fundamental in

---

<sup>1</sup> [http://olst.ling.umontreal.ca/~drouinp/termostat\\_web](http://olst.ling.umontreal.ca/~drouinp/termostat_web).

<sup>2</sup> <http://www.antlab.sci.waseda.ac.jp/software.html>

<sup>3</sup> See Lebart & Salem (1994) for explanation on the two criteria used to quantify the specificity of a term, one of them being the “standardized view of the frequency of the lexical units”.

discourse, since to be elected one could assume that to be perceived and constructed as a leader (associations with dynamism, initiative, courage, etc.) means to be constructed and perceived as an Agent, in control of the environment. This semantic value has also a pragmatic outcome: the described entity will be more likely perceived as the person to be elected.

The purpose of the next section is to show how indeed softwares such as SEMY (Grzesitchak Mick 2007)<sup>4</sup> or TERMOSTAT (Drouin 2003, Lesage et al 1993) can help in finding the ideological positioning of discourses in journalistic corpora, by identifying either the semantic recurrent elements or the specificities of the data.

## 2. Journalistic discourse, gender ideology and ‘semantic’ profiling

Several studies have analyzed the descriptions of French politicians and their actions in written and oral corpus. Most of them raised the issue of the role of gender (the gender of the entity described as well as the gender of the entity writing the text) in these descriptions. Thus Caillat (2012), Rabatel (2012) and Devriendt (2012) showed how the reported speech in journalistic discourse was adapted differently when comments were referring to a female candidate or to a male candidate in the 2012 presidential campaign. Similarly, our own work examined how female and male politicians are constructed in the French newspapers.

### 2.1 *The semantic profiling of a noun*

Our data for the year 2007 were collected over the 12 months before the French presidential elections of May 2007 for the candidates Royal, Bayrou and Sarkozy in the newspaper *Le Monde*. We collected 330 articles (roughly 100 for each candidate) focusing on one of the candidate i.e. containing the name of one of the candidate in the title. The sum of the words was comparable for each candidate since we had a total of 19.000 words for each candidate (see for the description of the data and the results Baider 2010 and Baider& Jacquey in print).

The software SEMY on the basis of the corpora we feed the program proposed lists based on the semantic recurrence in each textual file. These recurrences are based on the semantic feature which is found as the most salient in the data (see Giora 2003 for notion of saliency). We grouped the semantic features into a semantic field in order to summarize the semantic profile for each name. To do this classification we worked from the thesaurus CRISCO (<http://www.crisco.unicaen.fr>), in which near synonyms are listed. The example for the verbs *viser* (to aim) is given below. The verb was found as a recurrent semantic feature in Sarkozy’s file. We can then gather *viser* with the other verbs listed below.

To insert graph CRISCO

The verb *vouloir*, when found in the data given by SEMY, will be gathered in the same semantic field as *souhaiter* and *chercher*: these three verbs were for instance very much present in Sarkozy data analysis, and much less so in Royal data. This

---

<sup>4</sup> Developed at the ATILF at <http://www.atilf.fr/>

grouping, we suggest, allows us to describe then the semantic profile of the names Royal, Sarkozy and Bayrou in the journalistic discourse under study.

## 2.2 Gendered stereotypes?

Regarding the name Royal, the most salient dimension of the candidate personality was the relational aspect: 50% of the semantic features listed by SEMY in the context of the name Royal belonged to the field of COMMUNICATION (such as *to call, to declare, to announce, to express, to suggest*) or to the field of POLITICAL CAMPAIGN (*to vote, to prefer, to elect and to appoint*). In the case of her male opponents the semantic profiling was quite different.

Regarding the name Sarkozy, the most salient dimension was the field of ACTION (*to sign, to hire, to act, to do, to depart*), and the field of OBJECTIVES (*to aim, to search, to covet, to want, to consider, to reckon, to wish*). The semantic profile of the name Bayrou emphasizes as well ACTION towards a goal, but in a more positive fashion than in the case of the name Sarkozy (*to progress, to consolidate, to continue, to want, to offer, to make*); it emphasized also REBELLION (*to be unfavorable, to be against, to denounce, to disapprove, etc.*).

If we compare the three semantic profiles, we find our results to be similar with the conclusions of previous studies, even though these studies were carried out within a qualitative approach. Namely we identified the gender stereotype defining the female candidate and the male candidates to be the opposition ACTION vs. COMMUNICATION, reminding us of the cliché of women being good in relating to people and men being better in taking steps and making decisions.

These results are explained either by the discourse of the candidates themselves or by the perception the journalists have of the politicians. We cannot determine the source of the difference but the difference is there.

From a semantic point of view, SEMY has then revealed what we call “the deep structure” of the meaning of the names Bayrou, Royal, Sarkozy as constructed in discourse. However the software TERMOSTAT (Drouin 2003, Lesage *et al.* 1993) reveals another facet of the meaning by using not the semantic features but the *specificities* of words found in the context of each name. As well it shows how different these specificities are for each name. This is the topic of the next section.

## 3. Journalistic discourse, gender and the Specificity criteria

### 3.1 TERMOSTAT platform

The TERMOSTAT platform extracts from data made of simple and complex lexical units their specificity relatively to a reference corpus integrated into the platform (Drouin 2003). The reference corpus is a collection of articles in Le Monde (2002) which gather about 8 million words. TERMOSTAT is based on research results obtained by Lebart and Salem (1994), the criteria used to quantify the specificity of a term being the *test-value*, which is a standardized view of the frequency of the lexical units (Drouin 2003, 3). The platform through the calculation of specificity selects the most frequent terms in the submitted corpus compared to the most common terms in other texts gathered in the reference corpus. We then get a list with the most salient terms and other information relating to the linguistic profile of the file such as the various ‘physical appearances’ of this expression in the corpus (plural or feminine forms for nouns for instance), and finally the grammatical category of the lexical unit.

For instance if we look at the file for the noun Royal in Table 1 in section 3.2, the most frequent terms will be the terms found in articles focused on political campaign (*candidate, first round, vote, campaign, party nomination, meeting*, etc.) and they are similar for all candidates. The specificities are the words *socialist, socialist primaries*, etc. which indeed identify the political characteristics of the candidate immediately.

### 3.2 Specificities for the nouns Bayrou, Royal and Sarkozy

Tables 1, 2 and 3 present the specificities for each Noun Phrase. We reduced the lexical fields to the 35 most specific expressions in each corpus. We divided into three areas each list: the field of campaigning (1<sup>st</sup> column), the field of politics (2<sup>nd</sup> column) and the socio-cultural context (3<sup>rd</sup> column).

Expressions relating to campaigning and politics in the first and second columns classify the candidates in their respective political affiliations. We can then observe very clearly, the profile of the socialist candidate Royal (*socialist party, socialist candidate, socialist Members, socialist primary*, etc.). For the noun Sarkozy, his profile is also clearly delineated as right - wing (*Gaullist accents, UMP*, etc.). As for the noun Bayrou, his political affiliation is about being in the center (*centrist party*), but the specificities shows as well this placement being somewhat problematic, ‘stuck’ between the two major and competitor parties (*motion of socialist censorship, opponents on the right*) since he seems to have to contend with the demands and the pressure of his rivals (this was not found for the major parties, UMP and Socialists).

#### To insert Table 1

The specificities of the name Royal as shown in Table 1 reminds us of the semantic profile of COMMUNICATION for the noun Royal (*blog, televised debate, agitated debate*) and of CAMPAIGNING (*primary vote, voting intention, opinion survey*) found with SEMY. However we also identify another element which gives the ‘spirit’ of her campaign, the obsession of winning at all costs (*opiniomane*), and a duty to win.

If we compared further with the results obtained with SEMY, we can note an axiological value of these specificities, a value which was not clearly delineated with the previous software. The positive expressions (*majority approval, creating a citizen jury, important support, expressed support*) are less frequent than the negative expressions (*opinioname, eclectic support*). Actually quite a few phrases infer a difficult and turbulent campaign (*crossfire, anti-Ségolène festival, agitated debate*) and are particularly pejorative (*opiniomane* and *anti- Ségolène festival*). A great instability and antagonism characterize the specificities of the noun Royal. This will not be the case for the name Bayrou, nor for the name Sarkozy.

#### To insert Table 2

Pejorative terms such as *exhausted Republic, months of disorder* or *Bayroumania* are somewhat reminiscent of the turbulent energy we identified in the Royal corpus. However this negative aspect is largely compensated by the presence of positive and reassuring phrases such as *enthusiastic people, quiet centrist, independent strategy, extraordinary congress*. Moreover expressions such as *political fighter* and *opponents on the right* seem to explain the turbulence and define it as a particularly combative energy. The expressions *extraordinary congress, month of disorder, political fighter* and *original* seem to characterize an unexpected candidate or at least a non -

conformist type of presidential candidate. Finding expressions such as *Andalusian rhythms* and *techno-disco song* seem even incongruous in the context of a presidential campaign strengthening the original side of the centrist candidate. The reference to *civil unions* and *citizens' forums* also could identify Bayrou campaign strategy: to be seen as the man able of rallying everybody (as his centrist affiliation could suggest as well). Therefore the name Bayrou identified by the profile PROGRESS in SEMY is here profiled as the CONSENSUS choice, given the will of bringing citizens of opposite parties (left and right) together.

### Insert Table 3

Regarding the corpus Sarkozy, the political opposition against the candidate is denoted by the expression *anti-Sarkozy front*, a precarious and even a violent situation is inferred as well with the expressions *months of separation*, *broken wire*, and *tampered files*. These expressions could *prima facie* indicate a negative axiological value of the data. However, the accumulation of expressions denoting control (*master on board*, *quiet break-up*, *rapid information*, *idea of strength*, *courage*, *a kind of virility*) and demonstrating support (*sarkozystes*, *sympathizers*, *supporters*), results in a more balanced evaluation: the opposition and the difficulties exist, however it is possible to counter these problems thanks to the *supporters* and given the *courage* mentioned above. This confirms the semantic profiling of ACTION found with the software SEMY. However TERMOSTAT allows another point of view as this is action *against* a strong opposition; it also gives details on who are the targets of these actions. Indeed the list above includes the presence of phrases such as *aggressors of policemen*, *concerted immigration* and *juvenile offenders*, terms that make more precise the OBJECTIVES found by SEMY, namely tackling insecurity. These objectives also reveal the political strategy of Sarkozy (and his program), based on social concerns about violence and immigration, topics relating more to the National Front agenda than to the traditional UMP agenda. This strategy is actually summarized by the expression found in the 2<sup>nd</sup> column *turn to the right* of his politics. We also find another aspect of Sarkozy campaign strategy, an aspect which caused much stir in socio-political studies (Achin and Dorlin 2008): Royal and Sarkozy used and exploited politically their gendered personalities and this for the first time in the history of French politics hence the reference to <virility>.

### 3. AntConc and agency of an entity

#### 3.1 Kwic abstracts

With the concordance ANTCONC we worked on KWIC abstracts such as these ones listed below:

Ségolène Royal a haussé le ton, hier, et a accusé, à demi-mots,  
Comme Ségolène Royal, qui espère retourner l'opinion en sa faveur,  
un manque de relais au sein du PS, Ségolène Royal a ouvert le débat fin mai.  
Ségolène Royal, mitterrandiste à l'ère primaire  
Non ! Juste un discours de Ségolène Royal, candidate à la primaire socialiste,  
Le successeur de François Mitterrand s'appelle Ségolène Royal», lâche-t-il,

To evaluate agency, we focused on the syntactic functions of the names. Eckert and McConnell-Ginet pointed out that any NP in the position of subject in a sentence



will be associated with power and control (2003: 52). We also evaluated the semantic functions, in order to examine the actual *causality* of each NP, and restricted our study to Agent and non-Agent categories (Dowty 1991; Siewierska 1991). This evaluation has to be done manually (Baider 2009.)

### 3.2 Results

We found that the syntactic choices made in journalistic discourse when referring to the presidential candidates supported a partial dichotomy with the nouns Sarkozy and Bayrou functioning more often as a Subject as the noun Royal. As a matter of fact, if, in the same sentence both names of the candidate are found, the noun *Royal* will be found in 75% of the cases in a clause embedded to the clause in which the noun *Sarkozy* is the subject, as shown in the *Appendice 2*. This contrastive study of syntactic choices used to describe a female and a male candidate confirms a perception of the journalists conforming to stereotypes: the noun of the male candidate is, syntactically, in control of the verb, while the noun of the candidate is its subordinate. A difference was even greater as far as the semantic functions were concerned. The name of male candidate is constructed first as an ‘Agent’ which confirms the findings with SEMY and TERMOSTAT regarding the lexical semantics of the verbs used about him. On the contrary, the name of the female candidate played first the role of an aim (*for Royal*), a place (*chez Royal*) or a relation (*de Royal*) which we gathered in the category ‘Cause’. Moreover, the noun Royal was described twice as much as Sarkozy as an ‘Experiencer’ i.e. as a person who feels or a person who does not have the control on the action described in the sentence (Baider 2009).

However we could not find any of this asymmetry with another corpus gathering on the smaller candidates such as Voynet or Bové, or Buffet and Besancenot who had no chance to win. Several explanations were then suggested.

We suggested that when the stakes are not as high as a presidential election, reproducing gender stereotypes in the political sphere did not seem to be on the journalists’ agenda *i.e.* Buffet (the then communist party leader) was granted more agency than her young opponent Besancenot (the then extreme left leader). Parity can be afforded when there is not much to lose or to gain<sup>5</sup>. As well the Sarkozy and Royal candidatures were well-known to have been a somewhat exceptional case, in the sense that as French voters and French commentators noted, Royal and Sarkozy had played for the first time in the French political history on sexual stereotypes (Gingras 1997), (Achin & Dorlin 2008), (Coulomb-Gully 2009, 2012).

Therefore the French press reports *specifically* on Royal (*vs.* Sarkozy and Bayrou) during the 2007 presidential campaign may have been driven by and perpetuated a normative view of the sexes “constructed against the backdrop of strong social stereotypes and ascriptions about gender” (Bergvall 1999: 283). However explanation for these results could be as much the responsibility of the politician herself, and therefore her linguistic profile could also be considered as an exception but not the rule given our other results.

## 4. Discussion and comparison with 2012

---

<sup>5</sup> A tactic which is also followed by some parties when drawing their list of candidates for elections: women are present in region where stakes are not high or where they are most likely to lose (Sineau 2010).

In order to test the validity of our previous hypothesis and to decide among the different explanations put forward, we analyze the 2011 socialist primaries. Focusing on Aubry and Hollande shall we find similar statistics when referring to male and female politicians as we did in the case of Royal vs. Sarkozy?

When studying the two main socialist opponents Aubry and Hollande in the journalistic discourse during the socialist primaries, their linguistic profiles during the 2012 presidential campaign (Baider & Jacquey in press) confirmed the following hypothesis which could be drawn from the data given in this paper for the 2007 campaign: female and male politicians will be treated on the same footing in the French press as long as the stakes are not high; clearly once they get higher such as being potentially elected president, female politicians face negative discrimination whether in the journalistic discourse or in the opinion polls.

In the case of the 2012 campaign and before the Strauss Khan fiasco, we observed the name Aubry to be described extremely positively. Actually after a close reading of newspapers discourse with TERMOSTAT and with the KWIC abstracts, *no* blatant gender stereotyping could be witnessed in the six newspapers we took for building our corpus<sup>6</sup>. In the opposite, we observed reverse stereotypes.

The discourse described the name Aubry as a dynamic and organised manager, and the name Hollande as an embattled politician busy dealing with conflict. These topoi described in 2007 Sarkozy and Royal respectively. However we noted the shift in the linguistic portraits during the summer 2011, once it was clear that both Aubry and Hollande were favourites for representing the party against the incumbent president. Despite his lack of personality and his lack of professional experience as described in the press, Hollande became the consensus of the left, when three months before Aubry was the 'rassembleuse'. Suddenly *le vote Hollande c'est pour éviter la fragmentation* 'voting for hollande will avoid fragmenting', *il est mieux à même de rassembler la gauche* 'he is better suited (than her) to unite the left'. One sentence of our September corpus says it all: '*Hollande a plus de chances de gagner contre Sarkozy mais Aubry le mériterait plus*'. Hollande is more likely to beat Sarkozy but Aubry deserves it more'.

But why would Hollande be more likely to beat Sarkozy? Nothing in our data from March till May can explain that Hollande was better suited for the job than Aubry. Nothing in our data can explain the lead he always had in the opinion polls either. Nor his personality, nor his experience, nor his strategy was described as better than Aubry's. In the contrary as the many abstracts show all along our 2012 study (Baider & Jacquey in print). We could only conclude that Aubry seemed to be more suited for the job in the discourse describing both candidates but she lost the candidacy; and we can only add that Hollande, as an astute politician, saw his tactic of working on a personality opposite of Sarkozy's paying at the end (*il a déjà la tête dans la présidentielle* 'he already has his sights on the presidential campaign'). However we also observed that despite such a positive ethos, Aubry never lead in the polls. By contrast, despite a catastrophic portrait, Hollande lead in the polls, after Strauss Khan withdrawing from the race.

---

<sup>6</sup> We investigated six different newspapers to have a broad political spectrum and to cover big regions in France. We included the traditional *Libération*, *Le Monde* and *Le Figaro* newspapers since they are the most read whether on line or on paper. We also included *La Croix*, *Ouest-France* and the *Est-Républicain* for having data on regional press. Therefore 120 articles had been collected for each candidate who run in the campaign (1. 670 570 words in total).

Both female contenders Royal and Aubry had a long political career behind them; Aubry's game was the opposite of Royal's, playing the card of toughness and was portrayed as competent. Her adversary Hollande was also the opposite of what was Sarkozy since he was described with having no charisma, nor experience. But both female politicians lost. Maybe again, the fear of electing an atypical member of the French political class, a woman, although powerful, could then also explain both failures (Gerber 2009, Adler 1993, Lafrance *et al.* 1997).

### **Concluding remarks**

In concluding this study, we hope to have shown that the conclusions drawn from qualitative studies may be supplemented, refined and nuanced by working the same data using specific computer softwares. We have seen that SEMY based on definitions of the digital version of *Le Trésor de la langue française* allowed a reading taking into account the most salient semantic features found in the contextual environment of the names Royal, Bayrou and Sarkozy. This software gives what has been described as "an unprecedented glimpse into semantic organization" (Valette 2009). Then the linguistic profile for each candidate found with SEMY has been clarified and supplemented by the use of TERMOSTAT which works with specificities. Even though these observational tools must always be improved, since the lists they provide require manual work, we can only recommend adding these tools when working on textual meaning, provided that the researcher has a good knowledge of the context in which the data have been produced. We would then conclude with the three following recommendations:

- Manual studies are essential for syntactic-semantic research since for the moment they are impossible automatically ; such qualitative reading also allow the researchers to know very well their data, a prerequisite for deciding on which interpretation would be the best suited. ANTCOINC provides the KWIC abstracts which facilitate greatly this evaluation;
- SEMY reveals recurring semantic features which cannot be detected with a superficial reading: it brings to light isotopies (Rastier 1987) which are the semantic characteristics recurrent in a text;
- TERMOSTAT reveals how these isotopies are realized and how different they are for each candidate.

The three methodologies therefore appear complementary to the analysis of textual meaning. Finally to draw broader conclusions from these data and as far as gender is concerned, we have to examine the relationships between the other two dimensions of a communicative event, the discourse practices (*i.e.* here the processes of text production in the French press) and the socio-cultural practices (*i.e.* here the French society, its social and cultural activities of which these communicative events are part).

### **References**

Achin, Catherine & E. Dorlin (2008). Nicolas Sarkozy ou la masculinité mascarade du président, *Raisons politiques* (31) 19-46.

Adler, Laure (1993). *Les femmes politiques*. Paris: Seuil.

Baider Fabienne & Evelyne Jacquey (2010). Substantive Dis-Embodiement, Syntactic Embedment. *Feminism, Femininity and Gendered Discourse*, (Eds. J. Holmes & M. Marra). Cambridge Scholar Press. 145-167.

Baider Fabienne. (2009). Lexical change, discourse practices and the French press: Plus ça change, plus c'est la même chose? *Language Variation – European Perspectives II*, (Eds. S. Tsiplakou, M. Karyolemou et P. Pavlou), 27–46. John Benjamins: Amsterdam.

Baider Fabienne & Evelyne Jacquey (in print). Data mining for Tracking Gender Bias: a Curse or a Blessing? The 2012 presidential duos. *Ecological and Data-Driven Perspectives in French Language Studies* (eds H. Tyne *et al.*), Cambridge Scholar Press.

Baider Fabienne & Constantinou Maria (in print). Far right parties' discourse in Greece and France, to be published in the 2014 *Year book in Corpus linguistics*. (Ed. J. Tromero). Springer.

Bertini, Marie.-Josèphe. (2002) *Femmes, le pouvoir impossible*, Paris, Pauvert/Fayard.

Coulomb-Gully, Marlène (2009). Présidentielles 2007 : Média genre et politique. *Mots* 90 (1) 5-11.

Coulomb-Gully, Marlène (2012). *Présidente : le grand défi !* Éditions du Seuil, Paris.

Caillat Dominique (2012). Le recours au discours rapporté dans le débat présidentiel de l'entre-deux tours de 2007 : des usages sexués?, Paper given at the *Vème colloque du groupe de recherches Ci-dit*, 14-16 June 2012, Stockholm University, Sweden.

Devriendt Emilie (2012). Du DR dans deux séquences médiatico-discursives de la pré-campagne présidentielle française 2012: candidats vs. candidates, une loi du genre? Paper given at the *Vème colloque du groupe de recherches Ci-dit*, 14-16 June 2012, Stockholm University, Sweden.

Drouin, Patrick (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology* 9 (1) 99-117.

Eckert, Penelope & Sally McConnell-Ginet (2003). *Language and gender*. Cambridge: Cambridge University Press.

Fairclough Norman (1989). *Language and Power*. London: Longman

Fairclough Norman (1995). *Media Discourse*. London: Edward Arnold.

Gerber, Gwendolyn L. 2009. Status and the gender stereotyped personality traits: Toward an integration. *Sex Roles* 61: 297–316

Gidengil Elizabeth & Everitt Joanna (2003). Talking Tough: Gender and Reported

Speech in Campaign News Coverage. *Political Communication* 20, 209-32.

Giora Rachel (2003). *On our mind*. Oxford University press.

Grzesitchak Mick, Evelyne Jacquey & Mathieu Valette (2007). *Systèmes complexes et analyse textuelle : Traits sémantiques et recherche d'isotopies*, ARCo'07 – Cognition, Complexité, 227-235.

Goodwin M. H. (2011). Engendering Children's Play: Person Reference in Children's Conflictual Interaction." In *Conversation and Gender*. S. A. Speer and E. Stokoe, eds. 250-271. Cambridge: Cambridge University Press.

Kahn, Kim Fridkin (1994). The distorted mirror: Press coverage of women candidates for statewide office. *Journal of Politics* 56 (1) 154-173.

—. 1996. *The political consequences of being a woman: How stereotypes influence the conduct and consequences of political campaigns*. Columbia University Press: New edition.

Kanter, Rosabeth M. (1977). *Men and women of the corporation*. New York: Basic Books.

Kitzinger C. (2007). Is 'woman' always relevantly gendered? *Gender and Language* 1(1): 39-49.

Kohl, Diane. (2011). The Presentation of "Self" and "Other" in Nazi Propaganda. *Psychology & Society* 7 – 26.

Lafrance Marianne, Hiram Brownell & Eugene Hahn (1997). Interpersonal verbs, gender and implicit causality. *Social Psychology Quarterly* 60 (2): 138-152.

Lazar, Michelle M. (ed.) (2005). *Feminist critical discourse analysis: Studies in gender, power and ideology*. Palgrave Macmillan.

Lesage, René; Price, Wilson; Bissonnette, Carole & Patrick Drouin (1993). Enquête sur l'état d'utilisation des outils automatisés d'aide à la rédaction dans les organisations. *META* 38 (2) 367-389.

Macaulay, Monica & Coleen Brice (1997). Don't touch my projectile: Gender bias and stereotyping in syntactic examples. *Language* 73 (4) 798-825.

Michard, Claire (1988). Some socio-enunciative characteristics of scientific texts concerning the sexes, the nature of the right. *A feminist analysis of order patterns*, (ed. G. Seidel), 27-61. Amsterdam: John Benjamins.

Michard, Claire & Claude Ribéry (2008 [1987]). *Sexisme et sciences humaines Pratique linguistique du rapport de sexage*. Lille: Presses Universitaires du Septentrion.

- Mussolff, Andreas. (2008). What can critical metaphor analysis add to the understanding of racist ideology? Recent studies of Hitler's anti-Semitic metaphors. *Critical Approaches to Discourse Analysis across Disciplines* 2 (2):1-10.
- Ochs, Elinor. (1992). Indexing Gender. *Rethinking Context*. A. Duranti and C. Goodwin (eds). 335-358. Cambridge, Cambridge University Press.
- Penelope, Julia (1990). *Speaking Freely: Unlearning the Lies of the Fathers' Tongues*. New-York: Pergamon Press.
- Rabatel, Alain (2012), 'La parole des politiques soumise à contre-enquête' représente-t-elle de la même façon la parole des hommes et des femmes politiques ?'. Paper given at the *Vème colloque du groupe de recherches Ci-dit*, 14-16 June 2012, Stockholm University, Sweden.
- Reutenauer Coralie, Mathieu Valette & Evelyne Jacquy (2010). De l'annotation sémique globale à l'interprétation locale : environnement et image sémiques d'« économie réelle » dans un corpus sur la crise financière, *Cognitica – Actes du colloque de l'Association pour la Recherche Cognitive Arco '09 : Interprétation et problématiques du sens* (9–11 November 2009, Rouen), 29-39.
- Reutenauer, Coralie (2009). Analyse et modélisation sémantique à partir de ressources lexico-sémantiques : résumé, *Texte !* (XIV),1, <http://www.revue-texto.net/index.php?id=2093>.
- Roncarolo, Franca (2000). A woman at the Quirinal: Thanks, but no thanks. The social construction of women's political agenda in the 1999 Italian Presidential Election. *European Journal of Women's Studies* 7 (1) 103-126.
- Sineau, Mariette (2010). La Force du nombre : femmes et démocratie présidentielle. Ed. de l'Aube : Paris.
- Sreberny-Mohammadi, Annabelle & K. Ross (1996). Women MP's and the media representing the body politic. *Women in Politics* 49 (1) 103-117.
- Valette, Mathieu (2010). Propositions pour une lexicologie textuelle, *Zeitschrift für Französische Sprache und Literatur* 37, 171-188.
- Valette, Mathieu & François Rastier (2006). Prévenir le racisme et la xénophobie – propositions de linguistes. *Les langues modernes* 2, 68-77.
- van Dijk, Teung. (2006). Discourse and manipulation. *Discourse & Society* 17(2): 359-383.
- Widdowson, Henry. G. (2000). On the limitations of linguistics applied. *Applied Linguistics* 21 (1): 3-5.