

Vote for me. Don't vote for the other one

Edward Arnold ^{a*} and Dominique Labbe ^b

^a*Trinity College, Dublin, Republic of Ireland;* ^b*PACTE – CNRS, Grenoble, France*

(Received 18 December 2014; Accepted 5 January 2015)

The French presidential election takes place in two ballots. The second round opposes the two leading candidates at the end of the first. Between the two ballots, since 1974, the two finalists take part in a TV debate along the lines of the US presidential debates. This presentation analyses the texts of these six debates (136,000 words). A library of more than 6000 political texts – and nearly 13 million words – provides some benchmarks.

This paper presents the statistical indices proposed for the analysis of the communication within a situation of interaction. These indices are derived from theories concerning the presentation of actants in the speech, the expression of the speaker's subjectivity and the speech modalization. The application of these indices allows to bring a new perspective on these debates and it defines, for each of these indices, its scope, limitations and possible improvements.

The first part analyses the tendency of the speakers to personalize. These indices are broken down into the following dimensions: the relative importance given to the speaker, to the other and to the real message recipients (the listeners). The second part measures the fundamental choice in favour of the verb and, within this part of speech, between the accomplished ones (verbs to be and to have) and modal verbs (possible, desirable, obligation, knowledge). Finally, the greater or lesser density of the negation highlights the real scope of discourse.

The study leads to interesting conclusions about electoral discourse and the evolution of French political discourse over the last 40 years. Finally, it emphasizes the usefulness of vast corpora of texts and of lexicometry for language studying and teaching.

Keywords: systemic-functional grammar; enunciation theory; French political discourse; TV debates; personalisation

The language of power depends on the power of language
(Michael A. K. Halliday. Introduction: How Big Is a Language? On the Power of Language, 2006)

1. Introduction

The computer gave a new impetus to applied linguistics (Halliday 2006), especially to corpus linguistics (McEnery and Wilson 2003; Biber, Conrad, and Reppen 1998). Indeed, computers have allowed the researcher to establish large corpora that are the main tools for these applied linguistic studies and discourse analysis (for an overall presentation: Schifffrin, Tannen, and Hamilton 2003). This is especially the case for spoken corpora, which are prominent tools in the study of “real” languages (Adolphs and Carter 2013). These spoken corpora are much

*Corresponding author. Email: ejarnold@tcd.ie

A first version of this paper was presented at the 59th conference of the International Linguistic Association. Paris, France, 22–24 May 2014.

more difficult to establish than the written ones (Crowdy 1993; Nelson 1997; Du Bois and Alii 2000–5; Douglas 2003). They are also much more difficult to analyse.

1.1 *What utterance does?*

Until now, the most comprehensive theory is Halliday's systemic-functional grammar, especially his notion of cohesion (Halliday and Ruqaiya 1976). This theory makes it possible to analyse what utterances do and how they function (Halliday 1994).

This paper presents some statistical indices that are useful in the analysis of a communication within a situation of interaction between two "co-actants" (Halliday 1994). These indices are also derived from French theories concerning the presentation of actants in speech (Amossy 2010; Charaudeau 1992b), the expression of the speaker's subjectivity (Benveniste 1956, 1958; Dubois 1969, Kerbrat-Orecchioni 1981) and speech modalisation (Benveniste 1965; Gross 1999). These theoretical propositions can be tested on large corpora in order to define the scope and limitations of these theories and to suggest possible improvements to them.

The main purpose is to answer a complex question that all users of the spoken language corpora confront: some singularities that are observed may come from characteristics of the spoken language itself or, conversely, they could be explained by the personalities of the speakers, their persuasion strategies and the specific circumstances of the uttering.

In order to assess these explanations, the different speakers should have been placed in similar enunciative contexts so that the influence of such contexts are neutralised. Such laboratory conditions are very difficult to find, but the French presidential election offers an excellent case study.

1.2 *The presidential election debates*

The French presidential election takes place in two ballots. The second round opposes the two leading candidates at the end of the first. Between the two ballots, since 1974, the two finalists take part in a TV debate along the lines of the US presidential debates.¹ This paper analyses the texts of these six debates which comprise 136,000 words (Appendix 1). A library of more than 6000 political texts – and nearly 13 million words – provides some benchmarks. This gives the opportunity to analyse the "confrontation of political discourses" (Dupuy and Marchand 2011; Burger, Jacquin, and Micheli 2011).

Every form of oral communication needs to be replaced in its "context" for it to be evaluated. In the context of presidential debates, the situation of uttering has not changed, strictly speaking, since 1974: two individuals confront each other in a studio with two journalists who are responsible for ensuring that the two candidates have exactly the same speaking time. Even if the institutional framework has changed slightly (after 2002, the presidential term in office was reduced from 7 to 5 years and general elections were scheduled after presidential elections) the context has not changed fundamentally since 1974.

This being said, the political situation between the two rounds is always different. Is the outgoing president a candidate (as in 1981, 1988 or 2012)? Which candidate was the front-runner? Who will get the transfer votes from other first-round candidates (the second-round candidates are primarily trying to secure those votes)?

In other words, not only can the differences between different speakers come from their personalities, from their personal conception of politics, from their programmes, but also from the electoral context between the two rounds.

Comparison standards are provided by other sections of our Electronic Library of Modern French (Appendix 2), which includes 6000 political texts, comprising 12.5 million words (in French). We will also use data from British politics (Arnold 2005, 2008).

The following section analyses the tendency of the speakers to personalise. These indices are broken down into the following dimensions: the relative importance given to the speaker, to the other and to the real message recipients (the listeners, electors to convince). The second part measures the fundamental choice in favour of the verb and, within this part of speech, between the accomplished ones (verbs *to be* and *to have*) and modal verbs (possible, desirable, obligation, knowledge). Finally, the greater or lesser density of the negation highlights the real polemical scope of discourse.

The study leads to interesting conclusions about electoral discourse and the evolution of French political discourse over the past 40 years. Finally, it emphasises the usefulness of vast corpuses of texts and of applied statistics for language studying and teaching.

2. Pronouns and personalisation

Has the speaker chosen to personalise the comments or, conversely, to depersonalise them? Dubois and Dubois (1969) suggest that one needs to concentrate on the relative density of personal pronouns in order to calculate a *global personalisation index*.

2.1 Global measure

If, as Halliday (1994) suggests, analysis is restricted to the first two personal pronouns, the personalisation index can be formulated as follows:

$$\frac{\text{Number of personal pronouns}}{\text{Total number of words}}$$

Similarly, indices for references to the speaker (*I, je* in French), to his or her opponent (*you, vous* in French) can be calculated (Table 1). A third person is present in the debates (*we, nous*), that is to say, the speaker and other persons more or less specified.

Table 1. Relative density of total number of personal pronouns, first and second persons (per thousand words: ‰).

| Date | Candidates | Personalisation | | | |
|---------|----------------------|-----------------|---------------------|-------------------------|------------------------|
| | | index (‰) | je (‰) (<i>I</i>) | vous (‰) (<i>you</i>) | nous (‰) (<i>we</i>) |
| 1974 | V. Giscard d'Estaing | 84.9 | 32.3 | 21.4 | 6.5 |
| | F. Mitterrand | 90.2 | 29.6 | 24.0 | 5.9 |
| 1981 | V. Giscard d'Estaing | 82.3 | 25.2 | 20.0 | 9.6 |
| | F. Mitterrand | 80.4 | 34.3 | 12.3 | 3.4 |
| 1988 | F. Mitterrand | 85.3 | 30.8 | 16.3 | 6.0 |
| | J. Chirac | 85.7 | 33.4 | 18.5 | 9.2 |
| 1995 | J. Chirac | 77.6 | 27.8 | 11.7 | 11.6 |
| | L. Jospin | 74.0 | 30.9 | 8.7 | 5.6 |
| 2007 | S. Royal | 71.1 | 28.9 | 11.8 | 4.4 |
| | N. Sarkozy | 77.8 | 26.5 | 13.4 | 4.4 |
| 2012 | N. Sarkozy | 85.8 | 20.8 | 20.9 | 8.7 |
| | F. Hollande | 78.6 | 29.5 | 25.3 | 8.5 |
| Average | | 80.8 | 28.7 | 17.2 | 6.7 |

Density is expressed in terms of per thousand words (%). In 1974, for example, V. Giscard d’Estaing used on average around 85 personal pronouns (exactly 84.9) per 1000 tokens. This is 5% more than the average number of all the debates (average given on the last line of the table). Is this figure significant? For the political discourse section of the library as a whole it is 57.3‰. Compared to this standard, presidential debates have 48% more pronouns. This is a much higher density and is highly significant from a statistical point of view. For a discussion of these statistical tests see Labbé and Labbé (2013b). The co-actants in debates are described by Halliday (1994). See also Gee (1999) for the use and analysis of “I-statements”.

Can this significant level of personalisation be explained by the intrinsic characteristics of the face-to-face debate, or is it a feature of electoral campaigns? In the case of France, two rounds of the presidential election give a sound basis for comparison (Labbé and Monière 2008, 2013; Monière and Labbé 2010). The average personalisation index for the discourse of the main candidates was 64.8‰ in 2007 and 69.6‰ in 2012. It can be concluded, with a very low risk of being contradicted, that the context of the face-to-face debate encourages a degree of personalisation, which is clearly much higher than during electoral campaigns where the level of personalisation already exceeds “normal” ones. Do these averages show that the candidates made consciously different communication choices? To ascertain whether this is the case, the average has been used as a reference point and the individual density of each speaker has been converted into an index. Consider, for example, the case of V. Giscard d’Estaing (in 1974):

$$\frac{84.9}{80.8} \times 100 = 105.1.$$

In other words, in 1974, V. Giscard d’Estaing used 5% more personal pronouns than the average of all debates. For his opponent in the debate, the variation is +11.6%. If we accept Dubois’ analysis, it can be concluded that the first face-to-face debate was much more “tense” than the average of the following debates.

The results of this calculation are reproduced in Figure 1. The horizontal axis represents the average. Any value above the average indicates an overpersonalisation and vice versa.

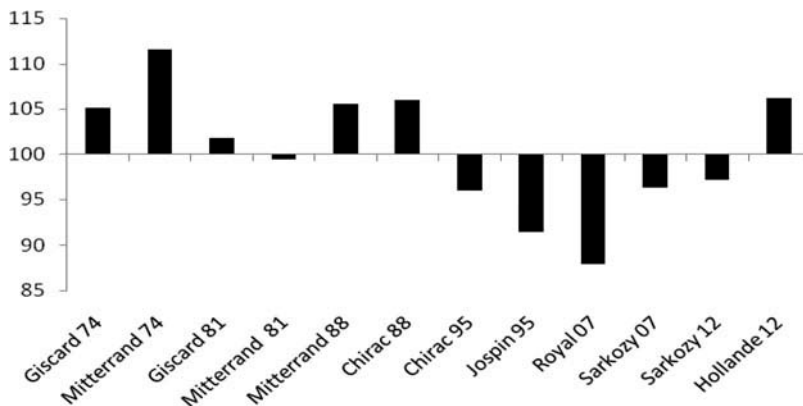


Figure 1. Propensity of each candidate to personalise discourse (average 100).

Two observations need to be made. First, with the exception of the last election, the two opponents seem to have made divergent choices. Some debates were particularly personalised: 1974 and 1988, and to a lesser extent 2012. Conversely, two debates were more impersonal (1995 and 2007). In other words, the choices could have been more a result of the political situation than the product of the speakers' personality or strategies. For example, F. Mitterrand and V. Giscard d'Estaing had a more personalised discourse in 1974 than in 1981 (Labbe 1981). This is equally the case for J. Chirac when the debates of 1988 and 1995 are compared. Only in 2012 (with the exception of a very small blip in 1981) can divergent choices be observed: N. Sarkozy opted for less personalisation (similar to his level of 2007), whereas F. Hollande returned to the levels of the debates of 1974 or 1988. Second, even if the differences between the speakers appear not to be very significant, from a statistical point of view, the gap between the extremes of each scale is statistically significant. There is a gap of 28% between F. Mitterrand in 1974 and S. Royal in 2007.

The personalisation index thus gives significant information that can be analysed by breaking it down according to discourse actors: the speaker, the addressee(s) (details in Table 1).

2.2 The First Person

The First Person is not just a word. It is a family of words. The first-person singular comprises not only "je" (*I*) but also "j'", "me" and "m'" (*me*), "moi", "mien" (ne,s). Possessive adjectives need to be added to this list: "mon, ma, mes" (*mine*). If the study is limited to pronouns, the average for all debaters is 28.7‰, which means that nearly three words out of a thousand are first-person pronouns (mainly "je"). Is this frequency normal in French politics? The following points of comparison are helpful in this respect:

French presidents (1958–2012): 19.5‰
 2007 presidential election: 19.6‰
 2012 presidential election: 20.9‰

With respect to these references, during the debates, candidates have over-used the first person (+45%). And it should be added that the French politicians use the first person twice as much as politicians from North America when they speak French (Canada and Quebec).

So during televised debates, speakers use the first person much more than they would in other situations, even electoral ones. However, in this case, choices are clearly very different (Figure 2, using the same principles as Figure 1).

F. Mitterrand is the champion of the "je": 34.3% in 1981 and 30.8% in 1988. He used it 60% more than N. Sarkozy did in 2012, the candidate who used it the least. In other words, when N. Sarkozy used the first person 100 times (during the last debate in 2012) F. Mitterrand used it 142 times (in 1981). With the exception of 1981 and 2012, the gap between the two candidates is not very significant (around 10%). The level of 2012 is very similar to that of 1981. The situations in 1981 and 2012 were quite similar. The fact that the outgoing president (V. Giscard d'Estaing and N. Sarkozy, respectively) had to defend his mixed record and deliberately remained vague on his future programme led to a relatively infrequent use of "je". The challengers opposite the outgoing president (F. Mitterrand in 1981 and F. Hollande in 2012) asserted themselves all the more so as a result of this electoral advantage.

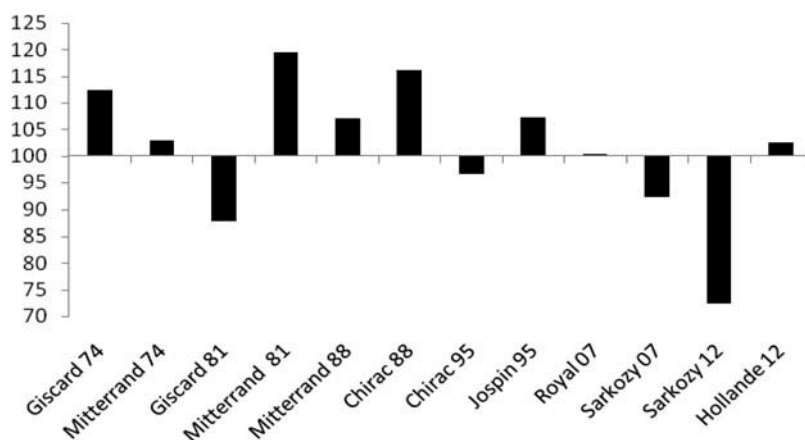


Figure 2. Propensity to use the first person (mean of debaters = 100).

These often very different choices have to be related to another fundamental question: How do the candidates address their adversary?

2.3 The other

There are three main ways of addressing somebody. One can use “vous” (you) or the person’s name preceded or not by “Monsieur” (Mr) or “Madame” (Madam). The person can also be referred to indirectly by “il” (he) or “elle”. In automatic scanning it cannot be distinguished from “il faut” (we must) and third persons. Consequently, the count has to be manual. The total of Table 2 indicates how much of the discourse is devoted to the opponent. The relative weighting of these three ways of referring to an opponent are

Table 2. The three ways to refer to an opponent (per thousand words).

| Date | Candidates | Vous (you) | Monsieur + nom (Mr + name) | Il (he) | Total |
|-----------|----------------------|------------|-------------------------------|---------|--------------|
| 10-mai-74 | V. Giscard d’Estaing | 21.40 | 12.3 | 1.35 | 35.05 |
| | F. Mitterrand | 23.96 | 6.34 | 1.41 | 31.71 |
| 05-mai-81 | V. Giscard d’Estaing | 19.99 | 5.96 | 1.18 | 27.13 |
| | F. Mitterrand | 12.34 | 6.95 | 1.66 | 20.95 |
| 28-avr-88 | J. Chirac | 16.30 | 8.05 | 1.93 | 26.28 |
| | Mitterrand* | 18.53 | 8.91* | 0.00 | 27.44 |
| 02-mai-95 | J. Chirac | 11.71 | 6.68 | 0.00 | 18.39 |
| | L. Jospin | 8.74 | 8.17 | 2.12 | 19.03 |
| | S. Royal | 11.80 | 0.76 | 0.00 | 12.56 |
| 02-mai-07 | N. Sarkozy** | 13.38 | 6.69** | 1.17 | 21.24 |
| 02-mai-12 | N. Sarkozy | 20.94 | 12.43 | 1.05 | 34.42 |
| | F. Hollande | 25.34 | 2.19 | 0.13 | 27.66 |
| Moyenne | | 17.22 | 7.12 | 1.00 | 25.34 |

*monsieur le premier ministre.

**madame Royal.

Table 3. Relative weighting of three ways of referring to an opponent (as a % of the total – maxima and minima values are in bold).

| Date | Candidats | Vous (you) | Monsieur + nom (Mr + name) | Il (he) | Total |
|-----------|----------------------|---------------|-------------------------------|-------------|-------|
| 10-mai-74 | V. Giscard d'Estaing | 61.1 | 35.1 | 3.9 | 100.0 |
| | F. Mitterrand | 75.6 | 20.0 | 4.4 | 100.0 |
| 05-mai-81 | V. Giscard d'Estaing | 73.7 | 22.0 | 4.3 | 100.0 |
| | F. Mitterrand | 58.9 | 33.2 | 7.9 | 100.0 |
| 28-avr-88 | J. Chirac | 62.0 | 30.6 | 7.3 | 100.0 |
| | Mitterrand* | 67.5 | 32.5 | 0.0 | 100.0 |
| 02-mai-95 | J. Chirac | 63.7 | 36.3 | 0.0 | 100.0 |
| | L. Jospin | 45.9 | 42.9 | 11.1 | 100.0 |
| 02-mai-07 | S. Royal | 93.9 | 6.1 | 0.0 | 100.0 |
| | N. Sarkozy** | 63.0 | 31.5 | 5.5 | 100.0 |
| 02-mai-12 | N. Sarkozy | 60.8 | 36.1 | 3.1 | 100.0 |
| | F. Hollande | 91.6 | 7.9 | 0.5 | 100.0 |
| Moyenne | | 68.0 | 28.1 | 3.9 | 100.0 |

shown in Table 3. The last column of this table measures the propensity to refer to an opponent. This propensity can be broken down into a direct reference to the opponent (vous), and indirect reference (il) or a mark of courtesy (monsieur). The weighting of these three aspects is shown in Table 3.

The use of “vous” directly refers to the opponent. It is generally accompanied by a question – or a condemnation – and introduces a maximum level of tension into the exchange. The most significant use of this form was during the debate between F. Hollande and N. Sarkozy in 2012. In other words, the two finalists in 2012 made the same choice to directly address their opponent. Conversely, in 1995, L. Jospin endeavoured to create the greatest difference between himself and J. Chirac by avoiding addressing his opponent directly. The latter used the direct form of address far less than he had against F. Mitterrand 7 years previously.

The second form of address (*Monsieur* or *Madame* and the name of the opponent) has a number of advantages; notably, it establishes a distance with the other and gives more weight to one’s discourse. The most enthusiastic advocate of this approach was N. Sarkozy in 2012, who tried to keep his opponent at a semantic distance, but was addressed directly by F. Hollande (in bold in the first column). There were 7.3% cases of “Monsieur”, followed by “Hollande” in most cases. In 2007, he used the same tactic with S. Royal and addressed her as “Madame” 10 times more than S. Royal used “Monsieur” with him. In 1988, F. Mitterrand never directly used his opponent’s name, and addressed him as “Monsieur le Premier Ministre” (Mr Prime minister). The following is the most memorable exchange:

M. CHIRAC.- Permettez-moi juste de vous dire que, ce soir, je ne suis pas le Premier ministre et vous n’êtes pas le Président de la République, nous sommes deux candidats, à égalité, qui se soumettent au jugement des Français, le seul qui compte, vous me permettrez donc de vous appeler monsieur Mitterrand.

M. MITTERRAND.- Mais vous avez tout à fait raison, monsieur le Premier ministre.²

There is a third way of addressing an opponent by using the third person: *il* (he) or *elle* (she). As we can see from the third column of [Tables 2](#) and [3](#), this third option is quite marginal (4% of references to the opponent on average). It was used significantly by L. Jospin in 1995 against J. Chirac (1 out of 10 times) and also by F. Mitterrand in 1981 (almost 8 times out of a 100). Indeed, both candidates had decided to avoid addressing their opponent directly and to use his name as little as possible. In both cases there is a deliberate choice to choose a communication strategy that focuses on the programmes and ideas of the opponent more than on his own character. Finally, it is important to note that certain uses of the pronoun *on* (somebody) pushes this logic to extremes: “quelqu’un que vous connaissez et qu’il n’est pas nécessaire de nommer” (“Someone you know and who doesn’t need to be named”).

Of course, this message is so ambiguous that it becomes difficult to understand. For example, “Je voudrais qu’on m’explique – c’est-à-dire que l’autre candidat m’explique...” (F. Mitterrand 1981)...³. This explains why this way of referring to the opponent is almost never used, the *on* (somebody) being almost always a familiar form or a quasi-impersonal form of “nous” (we).

[Table 3](#) shows that all the debaters (except L. Jospin in 1995) clearly favour a direct form of address. This is the main characteristic of the “French-style debate”. On the contrary, in North America, the debaters only directly address each other in exceptional circumstances. The journalists ask the questions, and the speaker refers to his opponent by name (Savoy 2010a, 2010b). The sum of these three forms of address (final column of [Table 2](#)) gives an index that reflects a propensity to talk about the opponent, to directly address or criticise him. This index varies between 12.6‰ (S. Royal in 2007) and 35‰ (V. Giscard d’Estaing in 1974 and N. Sarkozy in 2012). [Figure 3](#) summarises the different choices made by candidates (as previously, the average is 100).

All the differences are significant. V. Giscard d’Estaing (1974) or N. Sarkozy (2012) referred 2.8 times more to their respective opponent than S. Royal did in 2007. It can be said with very few risks of error that S. Royal chose to ignore her opponent as much as possible, whereas V. Giscard d’Estaing and N. Sarkozy made the opposite choice to focus the discussion on their opponent.

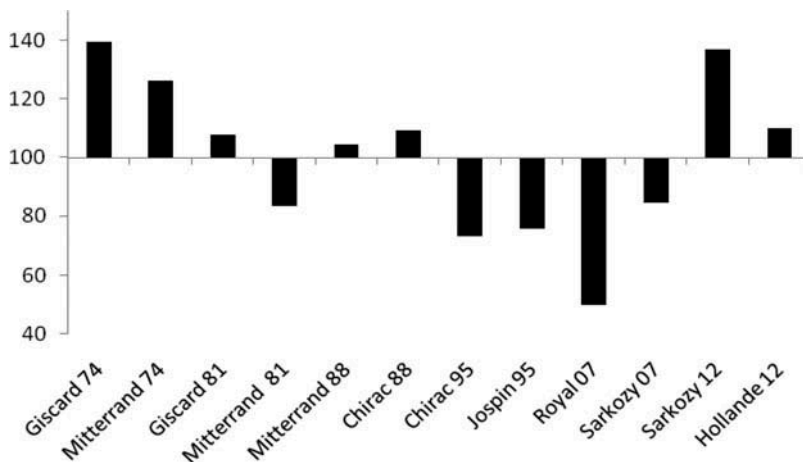


Figure 3. Distance in relation to the average propensity to refer to an opponent (average = 100).

The first debate ever (1974) was the confrontation during which the candidates spoke the most about their opponent. It was undoubtedly excessively so, as the same two candidates were much more reserved 7 years later in 1981. F. Mitterrand even consciously established a significant distance between the candidates. The two most ‘reserved’ debates in this respect were in 1995 (25% less than the average) and in 2007. The presence of the other opponent in a candidate’s discourse is highly revealing of a choice of communication strategy. This can be labelled the “propensity to challenge the other” or, as Dubois puts it, the “indice de la tension interpellative” (the index of interpellative tension). In turn, this index leads to another question: Did the candidate prefer to criticise their opponent or did they make the choice of self-valorisation?

2.4 Criticism or self-valorisation?

The answer can be found in the following equation:

$$\frac{\text{Auto} - \text{references (first person)}}{\text{References to the other}}$$

If the result exceeded 100 (a positive index), the speaker has devoted more time to self-valorisation, which can be summarised as “vote for me”. If the result is less than 100 (a negative index), the speaker has used most of his or her speaking time to say “don’t vote for the other”. Figure 4 presents the results of the calculation for the whole corpus. The horizontal axis is set at the point of equilibrium (100 or, in other words, the point at which there are as many references to oneself as to one’s opponent). The majority of the results are above this axis, which indicates that the majority of the debaters had a greater propensity to speak about themselves than about their opponent. However, this was not the case for the first and the most recent debate (1974 and 2012); during which, both candidates used the same technique (criticism of the other).

During the first debate in 1974, the candidates chose to criticise each other, and 10% more of the time was devoted to criticising and challenging directly the other candidate than it was to promoting their own candidacy and policies.

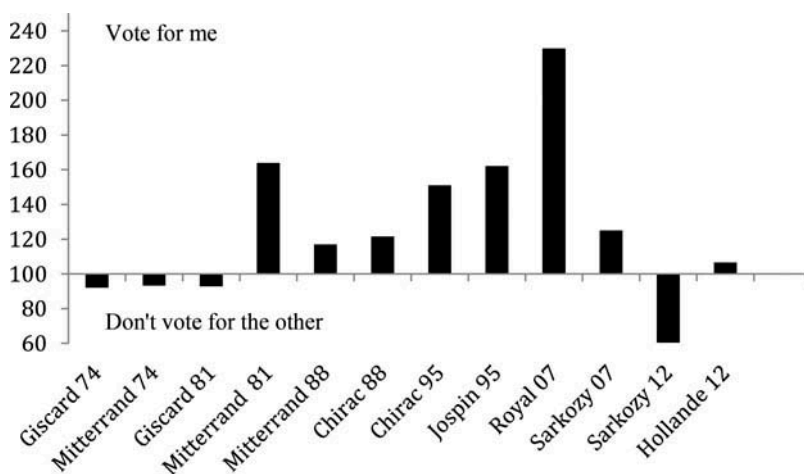


Figure 4. Relationship between self-valorisation and criticism of the other (balance at 100).

F. Mitterrand reversed this trend in 1981, and tried as much as possible to avoid addressing his opponent directly or using his name (the figure of 164 indicates that the propensity to talk about himself exceeded by 64% the criticisms made about the opponent). This tactical choice was continued up until 2012. During this election, the two candidates had clearly chosen the option of criticism of the other. N. Sarkozy had a level of self-promotion 40% below the level of criticism of the other. F. Hollande spent as much time criticising the outgoing president as he did talking about himself. During this debate, viewers heard more criticism and attacks than it did talk on policies.

These debates conclude with each candidate directly addressing the viewers and delivering a final appeal for their vote. The rest of the time the viewers are certainly the real target audience, but indirectly so through the use of the pronoun “nous” (we) (inclusive tension) or with the proper nouns “Frenchmen/women”. This inclusive tension is at its highest at the two extremities of the period studied. V. Giscard d’Estaing in 1974 and 1981, and N. Sarkozy and F. Hollande in 2012, all chose to criticise their opponent, as if this tactical choice involved calling upon the viewers to witness and weigh up these accusations.

3. Verbs in political discourse

The pronominal system is one dimension of verbal structures (as opposed to noun groups). The analysis of verbs is consequently a natural complement to that of pronouns, and is also significant. According to standard theory, and notably Dubois, the verb is the node of the French clause. Dubois distinguishes two levels. On the one hand, the verb is the main vector of dynamics between the subject, its discourse and the intended target of the discourse. If this initial theory is correct, a high density of verbs is a useful indication of “verbal tension”. On the other hand, in French, verbs indicate that which is accomplished (to be/to have) or what is incomplete (to do or to speak). Charaudeau (1992a) developed this idea by distinguishing between the stative verbs and those that express a process that can be subdivided into “actions” (when an actant is the agent of a process) and “facts” when there is nobody at the origin of the process.

This theory has never been substantiated, as in the French language it is difficult to identify all verb forms as there are many different conjugations, and the most frequent verbs are also substantives: *être* (to be or a human being), *avoir* (to have or an asset) *pouvoir* (to be able or the power), *devoir* (to must or the duty), *savoir* (to know or the knowledge), etc. These numerous homographies make any analysis or inventory by computer impossible. Tagging each of these words in the Digital Library of Modern French would overcome these difficulties and would test these theories.

3.1 Verb density

Table 4 gives the density of verbs (first column) and indicates variations between speakers.

In all the debates, one finds an average of 174 verbs per 1000 words. Most candidates are very close to this average. The minimum is 158‰ (L. Jospin in 1995) and the maximum 182‰ (F. Hollande in 2014). So, there are slight but nevertheless significant distances in first column ($\pm 6.5\%$). The density indicates a marked preference for using verbs in the political text section of the Digital Library of Contemporary French. In oral French, the average density of verbs is 192‰ (this section of the Library comprises more than 400 interviews and responses to open questions in four opinion polls). However,

Table 4. Relative density of verbs, negative constructions, past participles and auxiliaries (per thousand words: extreme values in bold).

| Date | Candidates | Total verbs | Etre+avoir (%)* | Passé (%) |
|---------|----------------------|--------------|-----------------|-------------|
| 1974 | V. Giscard d'Estaing | 171.1 | 26.5 | 18.4 |
| | F. Mitterrand | 180.0 | 22.1 | 19.0 |
| 1981 | V. Giscard d'Estaing | 171.1 | 25.6 | 21.7 |
| | F. Mitterrand | 169.9 | 22.4 | 15.8 |
| 1988 | F. Mitterrand | 181.1 | 23.5 | 25.7 |
| | J. Chirac | 174.8 | 28.5 | 21.9 |
| 1995 | J. Chirac | 172.6 | 27.6 | 13.8 |
| | L. Jospin | 158.3 | 27.4 | 17.2 |
| 2007 | S. Royal | 170.0 | 24.9 | 17.3 |
| | N. Sarkozy | 177.1 | 28.6 | 13.0 |
| 2012 | N. Sarkozy | 176.0 | 28.4 | 21.4 |
| | F. Hollande | 181.8 | 29.7 | 18.1 |
| Average | | 173.7 | 26.4 | 18.5 |

*Without auxiliaries.

unlike political debate, these interviews are real and spontaneous oral productions. On the one hand, politicians are professional speakers who are used to expressing themselves in public in a formal register. On the other hand, the debaters have carefully prepared their arguments beforehand and recite them during the debates. Furthermore, political discourse has the same lexical and syntactical characteristics as written French (Labbé and Monière 2008). In political discourse as a whole, verb density is at 148‰. When measured against this base line, presidential debates use 18% more verbs, which cannot be a coincidence. If we accept Dubois' hypothesis of verbal tension, is this high incidence a characteristic of the specific context of face-to-face oral exchange, or is it typical of electoral discourse as a whole?

In the last two presidential election campaigns in France, the following average verbal densities can be observed: 2007, 158‰; 2012, 165‰.

Two conclusions can be drawn from this high level of verbal tension. First, election campaign discourse uses more verbs than the rest of political discourse. Second, in this respect, the number of verbs used in the 2012 campaign discourse was clearly higher than in 2007. A closer analysis of Table 4 shows that the last debate reproduced and even exceeded the levels of 1974.

3.2 *To be and to have*

French enunciation theory suggests that, in this language, *to be* and *to have* should be distinguished from all the other verbs. In the totality of the debates, nearly a quarter of the uses of *to be* is as an auxiliary (followed by a past participle) as are 61% of the uses of *to have*. In addition, if the debaters are taken as a whole, *to be* or *to have* (used as non-auxiliaries) represent on average more than one verb out of four (26.2% to be precise). This score varies from 22.1% with F. Mitterrand in 1974 to 29.7% with F. Hollande in 2012. In other words, the density in F. Hollande's discourse is a third higher than it was for Mitterrand in 1974. This is a highly significant difference.

The interpretation of these results is problematic as it does not appear to correlate with any other explanations given by grammatical theory, specifically as regards candidates'

personalisation of discourse, self-promotion or criticism of opponents. In addition, the specific situation of each candidate does not seem to have an effect on frequency of verbal use. Those candidates who had previously taken part in several debates, despite being in very different electoral situations, subsequently made the same choices. Indeed, globally speaking, lexical and verbal choices change very little from one debate to another with V. Giscard d'Estaing, F. Mitterrand, J. Chirac or N. Sarkozy. These options seem to reflect an individual stylistic choice and a more or less clear personal preference to describe what the candidate claims to have accomplished. F. Hollande, N. Sarkozy or J. Chirac is more likely to use these two verbs, whereas F. Mitterrand or S. Royal is reluctant to use them and prefer other verbs.

In short, “to have” and “to be” are the two simplest verbs in the French language. In addition, they allow speakers to structure their discourse logically, as if what they are saying is in the natural order of things. For those who use these verbs the most, this tendency is generally linked with short sentences and a relatively limited vocabulary. Conversely, F. Mitterrand, and to a lesser extent S. Royal, uses “to have” and “to be” less frequently, and employs a more formal register, constructing longer sentences and developing an apparently more abstract or complex discourse.

3.3 Modalisation

In an article published in 1965, Benveniste proposed that the term of “modality” be used for any construction where a verb – the modal auxiliary – is combined with an infinitive verb, for example, *to want to do* (*vouloir faire*) (Benveniste [1966] 1970). Despite very few studies on these constructions, they are to be found at the heart of French language. In most corpora they are more numerous than auxiliary + past participle constructions (Labbé and Labbé 2013a).

According to Benveniste, two verbs *to can* and *to want* (*pouvoir* and *vouloir*) are modal auxiliaries by nature. He claims that this function of modalisation has been extended to other verbs such as *to desire*, *to wish*, *to must*, *to know* (*désirer*, *espérer*, *falloir*, *vouloir*, *savoir*). To these can be added two other French “pseudo-auxiliaries”: *aller* and *venir*, which do not have an equivalent in English and are translated by: *to be going to* and *to have just done*. These two modal auxiliaries are frequently used in debates to convey the future or immediate past: *I am going to answer*, *I have just done* (*je vais répondre*, *nous allons faire*... *je viens de dire*, *nous venons de faire*, etc.).

Dubois claims that these modal auxiliaries “indicate an attempt to control the debate and the level of tension facing the opponent” (p. 107). Thus the density of modal verbs in the discourse or the propensity to modalise utterances should be labelled “modalisation tension” (*tension modalisatrice*), and is related to the total number of words contained in the corpora.

This index is reproduced in the first column of Table 5. The auxiliaries also indicate the nature of this density: will/desire (*vouloir*), moral or legal imperatives (*devoir*), possibility (*pouvoir*), necessity (*falloir*), knowledge (*savoir*). Table 5 indicates the frequency of these auxiliaries observable in each candidate’s discourse, that is to say, their preference for one or the other of these modalities.

Paradoxically, J. Chirac is responsible for both the highest and the lowest density of modal auxiliaries. In 1995 against L. Jospin, the debate was centred on “necessity” (*falloir*). In 1988, against Mitterrand, the modal verbs conveyed will and desire (*vouloir*). Conversely, in 2012, the candidates focused on what their opponent was going to do, and consequently why voters should not elect him.

Table 5. Density of modal auxiliaries (per thousand words) and frequency of main auxiliaries by individual candidates.

| Date | Candidates | Modal. tension (%) | Possibility pouvoir | Will vouloir | Necessity falloir | Going to aller | Duty devoir | Do faire |
|------|-------------------------|-----------------------|------------------------|-----------------|----------------------|-------------------|----------------|-------------|
| 1974 | V. Giscard d'Estaing | 30.2 | 1 | 2 | 3 | 4 | 6 | 5 |
| | F. Mitterrand | 27.5 | 1 | 2 | 4 | 5 | 3 | 6 |
| 1981 | V. Giscard d'Estaing | 29.9 | 1 | 2 | 3 | 5 | 4 | 6 |
| | F. Mitterrand | 32.2 | 1 | 2 | 4 | 6 | 3 | 7 |
| 1988 | F. Mitterrand | 31.2 | 3 | 1 | 2 | 4 | 5 | 6 |
| | J. Chirac | 21.7 | 2 | 1 | 4 | 5 | 3 | 6 |
| 1995 | J. Chirac | 38.1 | 4 | 2 | 1 | 3 | 5 | 6 |
| | L. Jospin | 24.8 | 1 | 2 | 3 | 5 | 4 | 7 |
| 2007 | S. Royal | 30.2 | 1 | 2 | 4 | 3 | 5 | 6 |
| | N. Sarkozy | 36.6 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2012 | N. Sarkozy | 29.3 | 2 | 3 | 4 | 1 | 5 | 6 |
| | F. Hollande | 29.3 | 1 | 3 | 5 | 2 | 4 | 6 |
| | Average | 30.1 | 1 | 2 | 3 | 4 | 5 | 6 |

These results are quite similar to average densities recorded in other corpora of French political discourse.

In presidential campaign speeches, the average density of modal auxiliaries is 32‰ in 2007 and 33‰ in 2012. This density varies in the discourse of French 5th Republic presidents, from 22‰ (C. de Gaulle) to 33‰ (N. Sarkozy). There has apparently been a regular increase since G. Pompidou (23‰) up until F. Mitterrand (29‰) and J. Chirac (32‰). This reveals a more generalised process of personalisation and increasing density in political discourse since 1958. Indeed, this average level and tendency are also present in “general policy declarations” – these *déclarations de politique générale* are similar to speeches from the throne (Queen’s Speech) by English prime ministers – by the heads of French government since 1945 (average 27‰).

The high density of modal verbs would seem to be a characteristic of French political discourse. It is less present in the discourse of Quebec prime ministers (20‰) and even less marked in English-speaking Canada (16‰). This is also the case in oral and Literary French (Cyril and Dominique 2013a).

The majority of debaters deviate very little from the mean, but some individuals show considerable distances; Figure 5 shows such cases.

These differences cannot be ascribed to the “personal style” of the candidates since one of them (J. Chirac) can be found at both ends of the scale. His propensity to modalise increased by 75% between 1988 and 1995, as if his announced defeat encouraged him to take a low profile in 1988, and, conversely, when victory seemed certain in 1995 he radically changed his discourse. Others, such as F. Mitterrand, seemed to be more stable in their values, and also did not hesitate to modalise more when the situation was favourable (1981 and 1988 for F. Mitterrand, 2007 for N. Sarkozy).

The orientation of the propensity to modalise completes this index. For most debaters, modal verbs of possibility and wish/will are preponderant. Conversely, modals of knowledge (*savoir*) are absent or highly marginal for all of them. There have been some exceptions to this preponderance of possibility and will/desire. The 1988 debate was

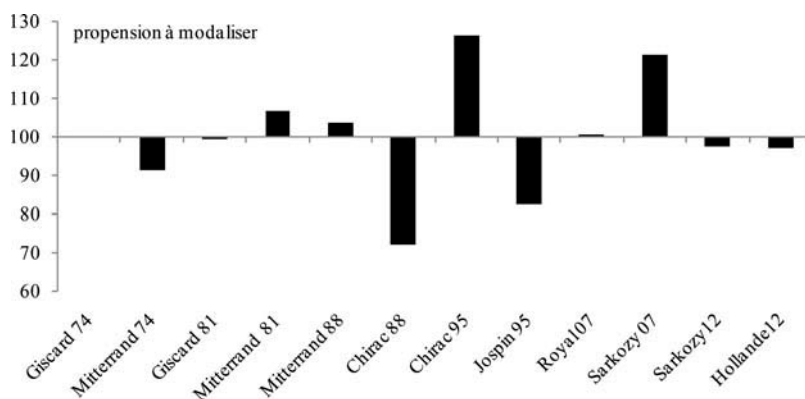


Figure 5. Propensity to modalise (in relation to the mean).

dominated by the will (*vouloir*) modal auxiliaries (by both F. Mitterrand and J. Chirac). In 1995, J. Chirac mainly used present and future forms of obligation and necessity (*il faut* and *il faudra*). In 2012, N. Sarkozy chose to criticise what his opponent (F. Hollande) “was going to do” (*va faire*). F. Hollande used the same tactic almost as intensively (*aller* is used just behind *pouvoir* in his use of modal auxiliaries). *Aller* is actually a way of expressing a future fact, but which has been given an even greater sense of immediacy by attaching a present clause to it. So “il va faire” (“he is going to do”) implies “immediately after being elected”, and it is for this reason that electors shouldn’t give him their vote!

3.4 Density of the negation

A negative construction indicates the reiteration of a proposition to which the speaker is opposed. A high density of negatives is indicative of a discourse that is structured against the opponent, that is to say, an essentially polemical communication (Labbé 2005). This density can be measured by comparing the frequency of “not” (*ne... pas*) and “no more” (*ne plus*)... to the total number of verbs. The result can be called a negativity or polemical index.

Table 6 expresses this dimension (from the least to the most polemical).

The last line indicates that on average more than 1 out of 10 verbs is in a negative construction (11.3% to be precise). However, variations around this mean can be considerable. The smallest indices are with F. Mitterrand (1974) and J. Chirac (1988). The greatest values are to be found with the debate between N. Sarkozy and S. Royal in 2007, and F. Hollande in 2012. If one takes the lowest density as a benchmark (last column of Table 6), it can be observed that N. Sarkozy exceeds this comparison by 71% against S. Royal (in 2007) and by 53% against F. Hollande (in 2012). F. Hollande himself uses negative constructions 44% more than the average, and V. Giscard d’Estaing 25% more in 1974.

There is an almost perfect convergence between these negative verbal constructions and the self-valorisation index discussed earlier. The most negative candidates are also those who have made the choice of criticising their opponent rather than put forward their own candidacy through self-valorisation. This is a constant characteristic of N. Sarkozy, and was the choice made by F. Hollande when he confronted N. Sarkozy in 2012. This

Table 6. Density of negative constructions (100 verbs ordered by increasing density).

| Candidates | Negative construction (% verbs) | Index (mean = 100) | Index (Mitterrand 1974 = 100) |
|-----------------|------------------------------------|-----------------------|----------------------------------|
| Mitterrand 1974 | 9.1 | 81 | 100 |
| Chirac 1988 | 9.1 | 81 | 100 |
| Mitterrand 1988 | 9.8 | 87 | 108 |
| Mitterrand 1981 | 10.0 | 89 | 110 |
| Chirac 1995 | 10.6 | 94 | 117 |
| Jospin 1995 | 10.7 | 95 | 118 |
| Giscard 1981 | 11.0 | 97 | 120 |
| Royal 2007 | 11.2 | 100 | 123 |
| Giscard 1974 | 11.4 | 101 | 125 |
| Hollande 2012 | 13.1 | 116 | 144 |
| Sarkozy 2012 | 13.9 | 123 | 153 |
| Sarkozy 2007 | 15.6 | 138 | 171 |
| Average | 11.3 | 100 | 124 |

confirms the statement that the 2012 debate was the most negative (or polemical) of all the debates.

4. Conclusions

The French presidential face-to-face debates in a television studio give precious information to researchers of language or communication. It is beyond the scope of this study to perform a complete analysis of these debates. However, the aim has been to test certain hypotheses of systemic-functional grammar and enunciation theory. As such, a number of statistical variables seem useful for the analysis of political discourse, as they help to classify speakers and reveal choices of political communication and even personal characteristics and style. A more or less intense tendency to personalise is linked to self-valorisation and devalorisation of the opponent. This reveals a fundamental choice between two different registers: the explanative genre and the polemical one. The density of verbs and negativity, and the tendency to modalise further nuances this analysis. More research on these variables is necessary to assess their real importance. Large digital databases with standardised and tagged texts provide large corpora and indispensable reference standards for these applied studies. These digital libraries will also be precious tools for the study and teaching of languages.

For political science, this initial inventory leads to interesting conclusions. Candidates for elections have an initial choice to make. They can underline their own qualities and contributions, and show how they could resolve the problems facing the country. Conversely, they could focus on criticising their opponents by claiming that they are unsuited to lead the country, and that to elect them would be catastrophic. All electoral discourse contains a mixture of both approaches, but depending on what dimension the candidate favours, the tone of the politician's discourse will be very different. If the first approach is chosen, the discourse is relatively non-conflictual and less personalised, and dominated by explanations and a defensive approach, in other words, by an attempt to be pedagogical. The second approach will imply a personalised and polemical discourse. In 2007, N. Sarkozy had made that choice but S. Royal chose not to engage with this tactic. In 2012, the two candidates chose to favour the polemical approach and provoked a

“spiral of negativity”, which led to a “rhetoric of invective” (Labbé and Monière 2013). These negative campaigns are not exclusive to France. They have been present in North American politics for the last 20 years or so (Hansen and Pedersen 2008; Monière 2012).

Acknowledgements

The programmes used for this article and the Digital Library of Modern French were prepared with the help of Cyril Labbé (Institut de Mathématiques Appliquées – Université Joseph Fourier – Grenoble). The corpora of political discourse used were compiled with the help of Denis Monière (Université de Montréal).

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. Coullomb-Gully (2009) for France and Savoy et al. (2010a, 2010b) for the United States (bibliographic references are at the end of the paper).
2. “M. CHIRAC. – May I just say that tonight I am not the Prime Minister and you are not the President of the Republic. We are two candidates on an equal footing, who will be judged by the French people, the only judgment that counts. Allow me then to call you Monsieur Mitterrand.
M. Mitterrand. – But you are absolutely right, Monsieur Prime Minister”.
3. “I would like someone to explain to me – that is to say, that the other candidate explains to me...”.

Notes on contributors

Edward Arnold (1965-) completed his PhD at Institut d’Etudes Politiques de Grenoble. He is Assistant Professor in French and European Studies since 1995 and is Director of Centre for European Studies, Trinity College, University of Dublin, for the period 2011–15. His research interests are discourse analysis and history of ideas.

Dominique Labbé (1947-) completed his PhD at Fondation Nationale des Sciences Politiques de Paris 1975. He has been working with Laboratory PACTE (Centre National de la Recherche Scientifique, University of Grenoble) since 1976. His research interests are statistics applied to language. He published more than 200 scientific publications. His last book was *La Champagne Présidentielle de 2012* published in 2013 by L’Harmattan, Paris (co-authored by Denis Monière).

ORCID

Edward Arnold  <http://orcid.org/0000-0002-6002-2258>

Dominique Labbe  <http://orcid.org/0000-0003-3445-1857>

References

- Adolphs, S., and R. Carter. 2013. *Spoken Corpus Linguistics*. New York: Routledge.
- Amossy, R. 2010. *La présentation de soi, ethos et identité verbale*. Paris: PUF.
- Arnold, E. 2005. “Le discours de Tony Blair (1997-2004).” *Corpus* 4: 55–77.
- Arnold, E. 2008. “Le sens des mots chez Tony Blair (people et Europe).” In *9^e Journées internationales d’analyse statistique des données textuelles (Lyon, 12–14 Mars 2008)*, edited by S. Heiden and B. Pincemin, 109–119. Vol. 1. Lyon: Presses universitaires de Lyon.
- Benveniste, E. 1956. “La nature des pronoms.” *Reproduit dans Benveniste* 1966: 251–257.
- Benveniste, E. 1958. “De la subjectivité dans le langage.” *Reproduit dans Benveniste* 1966: 258–265.

- Benveniste, E. 1965. "Structure des relations d'auxiliarité." *Reproduit dans Benveniste* 1970: 177–193.
- Benveniste, E. [1966] 1970. *Problèmes de linguistique générale*. Paris: Gallimard (rééd. 1980).
- Biber, D., S. Conrad, and R. Reppen. 1998. *Corpus Linguistics*. Cambridge: Cambridge University Press.
- Burger, M., J. Jacquin, and R. Micheli, eds. 2011. *La parole politique en confrontation dans les médias*. Bruxelles: De Boeck.
- Charaudeau, P. 1992a. "Faits de discours." In *Grammaire du sens et de l'expression*, 133–155. Paris: Hachette.
- Charaudeau, P. 1992b. "La modalisation et les modalités énonciatives." In *Grammaire du sens et de l'expression*, 569–629. Paris: Hachette.
- Coullomb-Gully, M., dir. 2009. 2007. Débats pour l'Elysée. *Mots* 89. mars.
- Crowdy, S. 1993. "Spoken Corpus Design." *Literary and Linguistic Computing* 8-4: 259–266.
- Douglas, F. M. 2003. "The Scottish Corpus of Texts and Speech: Problems of Corpus Design." *Literary and Linguistic Computing* 18 (1): 23–37. doi:10.1093/lilc/18.1.23.
- Du Bois, J. W., and Alii. 2000–5. *Santa Barbara Corpus of Spoken American English, Parts 1–4*. Philadelphia, PA: Linguistic Data Consortium.
- Dubois, J. 1969. "Énoncé et énonciation." *Langages* 4: 100–110. doi:10.3406/lgge.1969.2511.
- Dupuy, P.-O., and P. Marchand. 2011. "Confrontation et positionnement dans les duels de l'entre-deux tours: une approche lexicométrique." In *La parole politique en confrontation dans les médias*, edited by M. Burger, J. Jacquin, and R. Micheli, 129–147. Bruxelles: De Boeck.
- Gross, M. 1999. "Sur la définition d'auxiliaire du verbe." *Langages* 135: 8–21.
- Gee, J. P. 1999. *An Introduction to Discourse Analysis: Theory and Method*. London: Routledge.
- Halliday, M. A. K., and R. Hasan. 1976. *Cohesion in English*. London: Longman.
- Halliday, M. A. K. 1994. "So You Say 'Pass'. Thank You Three Muchly." In *What's Going on Here: Complementary Studies of Professional Talk*, edited by A. D. Grimshaw. Norwood: Ablex. Reprinted in Halliday, M. A. K. "Linguistic Studies of Text and Discourse." In *Collected Works of M. A. K. Halliday*, edited by W. Jonathan, 228–254. Vol. 2. London: Continuum.
- Halliday, M. A. K. 2006. "Computational and Quantitative Studies." In *Collected Works of M. A. K. Halliday*, edited by J. J. Webster. Vol. 6. London: Continuum.
- Hansen, K. M., and R. T. Pedersen. 2008. "Negative Campaigning in a Multiparty System." *Scandinavian Political Studies* 31-4: 408–427. doi:10.1111/j.1467-9477.2008.00213.x.
- Kerbrat-Orecchioni, C. 1981. *L'énonciation de la subjectivité dans le langage*. Paris: A. Colin (réédition 2002).
- Labbé, C., and D. Labbé. 2013a. "La modalité verbale en français contemporain. Les hommes politiques et les autres." In *La modalité, le mode et le texte spécialisé*, edited by B. David, 33–61. Paris: L'Harmattan.
- Labbé, C., and D. Labbé. 2013b. "Le chiffre dans le discours politique français contemporain. V. Giscard d'Estaing et les autres présidents." *Communication aux XIVe Journées de l'ERLA*. Brest. November 15–16.
- Labbé, D. 1981. "Le débat Giscard d'Estaing-Mitterrand." *Revue Française de science politique* XXXI-5-6 (Oct.–Dec. 1981): 951–981.
- Labbé, D. 2005. "Le général de Gaulle en campagne." In *Aspects linguistiques du texte de propagande*, edited by D. Banks, 213–233. Vol. 2005. Paris: L'Harmattan.
- Labbé D., and D. Monière 2008. "Des mots pour des voix : 132 discours pour devenir président de la République française." *Revue Française de science politique* 58 (3): 433–455. doi:10.3917/rfsp.583.0433.
- Labbé, D., and D. Monière. 2013. *La campagne présidentielle de 2012. Votez pour moi !* Paris: L'Harmattan.
- McEnery, T., and A. Wilson. 2003. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Monière, D. 2012. *La spirale de la négativité dans les campagnes électorales canadiennes de 2008 et de 2011. Journée d'étude : Comment convaincre ? Analyse scientifique de la campagne électorale 2012*. Grenoble: Institut d'études politiques de Grenoble. March 9.
- Monière, D., and D. Labbé. 2010. "Quelle est la spécificité des discours électoraux? Le cas de Stephen Harper." *Revue canadienne de science politique* [Canadian Journal of Political Science] 43-1: 69–86.
- Nelson, G. 1997. "Standardizing Wordforms in a Spoken Corpus." *Literary and Linguistic Computing* 12-2: 79–85.

- Savoy, J. 2010a. "Discours électoral et discours présidentiel." In *Proceedings of 10th International Conference Statistical Analysis of Textual Data*, edited by B. Sergio, et al., 827–838. Vol. 2. Rome: Edizioni Universitarie di Lettere Economia Diritto.
- Savoy, J. 2010b. "Lexical Analysis of US Political Speeches." *Journal of Quantitative Linguistics* 17-2: 123–141. doi:[10.1080/09296171003643205](https://doi.org/10.1080/09296171003643205).
- Schiffrin, D., D. Tannen, and H. E. Hamilton, eds. 2003. *The Handbook of Discourse Analysis*. Oxford: Blackwell.

Appendix 1

Table A1. The six televised debates held during the two rounds of the French presidential elections.

| Date | Candidates | Length (word tokens) | Vocabulary (word types) |
|---------------|----------------------|----------------------|-------------------------|
| 10 May 1974 | V. Giscard d'Estaing | 10,408 | 1164 |
| | F. Mitterrand | 8515 | 1255 |
| 5 May 1981 | V. Giscard d'Estaing | 11,906 | 1442 |
| | F. Mitterrand | 9640 | 1425 |
| 28 April 1988 | J. Chirac | 9770 | 1328 |
| | Mitterrand | 9820 | 1444 |
| 2 May 1995 | J. Chirac | 10,337 | 1430 |
| | L. Jospin | 12,248 | 1580 |
| 2 May 2007 | S. Royal | 11,776 | 1460 |
| | N. Sarkozy | 12,851 | 1533 |
| 2 May 2012 | F. Hollande | 15,509 | 1671 |
| | N. Sarkozy | 15,283 | 1623 |
| Total | | 138,063 | 5247 |

Appendix 2

Table A2. Digital library of modern French (Bibliothèque Electronique du Français Moderne) (1 March 2014).

| | Length (word tokens) | Vocabulary (word types) |
|--------------------------------------|----------------------|-------------------------|
| Political discourse | 11,529,763 | 42,885 |
| French presidents (1958–2012) | 3,824,965 | 23,602 |
| Canadian prime ministers (1867–2012) | 1,098,161 | 13,514 |
| Quebec prime ministers (1867–2012) | 2,993,823 | 22,458 |
| French prime ministers (1945–2012) | 288,526 | 7952 |
| Presidential campaign (2007) | 809,384 | 8091 |
| Presidential campaign (2012) | 1,773,808 | 13,652 |
| Debates | 138,063 | 5247 |
| Literature (17th to 20th century) | 10,903,628 | 56,192 |
| Novels and short stories | 6,202,751 | 48,365 |
| Theatre | 2,571,497 | 15,551 |
| Poetry | 675,187 | 18,810 |
| Correspondence | 345,542 | 11,070 |
| Crime fiction | 548,682 | 17,274 |
| Press | 2,939,632 | 58,690 |
| Sciences | 774,514 | 18,523 |
| Oral French | 2,978,122 | 18,429 |
| Total | 29,674,341 | 99,921 |