HAL
open science

Le chiffre dans le discours politique français contemporain

Cyril Labbé, Dominique Labbé

To cite this version:

Cyril Labbé, Dominique Labbé. Le chiffre dans le discours politique français contemporain. La quantification dans le texte de spécialité, Nov 2013, Brest, France. hal-00949594

HAL Id: hal-00949594

https://hal.science/hal-00949594v1

Submitted on 19 Feb 2014

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

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

La quantification dans le texte de spécialité

Les Nouvelles Journées de l'ERLA No. 14
Université de Bretagne Occidentale - Brest
15-16 novembre 2013

Cyril Labbé
Laboratoire d'Informatique de Grenoble - Université Joseph Fourier
(cyril.labbe@imag.fr)

Dominique Labbé
Laboratoire PACTE (CNRS - Institut d'Etudes Politiques de Grenoble)
(dominique.labbe@umrpacte.fr)

Le chiffre dans le discours politique français contemporain

V. Giscard d'Estaing et les autres présidents

Résumé

Dans leurs discours, les politiques utilisent beaucoup la quantification (nombres et dates). C'est ce que révèle la comparaison des discours tenus par un grand nombre de responsables politiques français quand on les compare à un vaste échantillon du français moderne comportant 30 millions de mots, tous étiquetés. Les chiffres et les dates servent à donner aux discours un ancrage dans le temps, la réalité économique et sociale. Une illustration est donnée avec V. Giscard d'Estaing (1974-1981) comparé aux cinq autres présidents de la République entre 1958 et 2012.

Abstract

In their speeches, the politicians use much quantification (numbers and dates). This is shown by the comparison of speeches by many French politicians to a wide excerpt of modern French of 30 million words, all labeled. Numbers and dates are used to anchor a speech in the present times and in the economic and social reality. An illustration is given with the help of the V. Giscard d'Estaing's presidential speeches (1974-1981) compared to the ones by the other five French presidents between 1958 and 2012.

Comment les locuteurs français utilisent-ils les nombres et, plus largement, la quantification ? Plus précisément, quelle est la singularité de la quantification dans telle ou telle spécialité ? Ces deux questions en amènent logiquement une troisième : la statistique n'est-elle pas la première "spécialité" concernée par la quantification et plus largement par la "politique des grands nombres" (Desrosières 1999) ?

C'est d'ailleurs grâce à la statistique qu'on peut répondre objectivement à la question posée dans ce colloque. Deux exemples permettront de le comprendre. On sait qu'en Bretagne, il y a significativement plus d'agriculteurs et de pêcheurs que dans les autres régions françaises ; ou encore que le climat y serait plus humide.

Ces deux exemples donnent une première définition du principal outil de la quantification : le nombre, c'est-à-dire le résultat d'un compte, l'expression d'une mesure. Le nombre comporte plusieurs unités linguistiques : un ou plusieurs chiffres - exprimant des grandeurs - et une ou des unités désignant les objets décomptés (population active, température).

Bien sûr, à propos de la population active ou de l'hygrométrie, l'intuition va dans le sens d'une singularité bretonne mais il est possible d'aller plus loin grâce au recensement de la population active française et aux relevés pluviométriques dans l'hexagone, mais aussi grâce aux tests statistiques qui permettent d'affirmer que les valeurs observées en Bretagne s'écartent significativement de celles enregistrées dans le reste du pays.

De même, pour juger de la singularité des usages linguistiques dans telle ou telle spécialité, il faudrait disposer de l'équivalent des recensements ou des relevés climatiques, c'est-à-dire de vastes enquêtes sur les pratiques langagières dans la population générale et dans la "spécialité" étudiée, ce qui permettrait de mesurer précisément les différences entre les unes et les autres. Ces vastes échantillons –nommés corpus – pourraient être un outil pour la linguistique appliquée.

Notre exposé évoquera succinctement la constitution de ces corpus, en insistant sur le traitement des chiffres. Ce sera l'occasion de revenir sur le discours politique - dont nous avons déjà présenté les principales caractéristiques (Labbé & Labbé 2011) - et sur le "langage des chiffres en politique" (Bacot & Al 2012).

1. Corpus et méthodes

Depuis plus de 30 ans, une bibliothèque électronique du français moderne est en cours de constitution (voir tableaux en annexe 1). Cette bibliothèque comporte notamment une section consacrée aux présidents de la République française depuis 1958 et jusqu'en 2012. C'est principalement elle qui sera utilisée dans la suite de cette communication.

1.1. Traitements préliminaires

Avant d'entrer dans la bibliothèque, chaque texte est édité (voir un exemple en annexe 2). Il est doté d'une entête – entre <> - indiquant le locuteur, la date, etc. De même, tout ce qui n'a pas été prononcé par le locuteur est également isolé (par exemple dans les entretiens : les questions des journalistes ou mentions telles que "rires", "applaudissements"…)

Beaucoup de ces documents ne se trouvent pas (ou plus) en ligne. Dans ce cas, ils sont scannés puis corrigés. Les textes en ligne demandent aussi un travail important d'édition avant de pouvoir entrer dans la bibliothèque. Ce fut le cas des discours de V. Giscard d'Estaing (qui ont été traités en 2013). Considérons le premier d'entre eux (annexe 2). Le texte original est entièrement en majuscules (c'est le cas de l'ensemble des textes des années 1970 dans la base de données *Vie publique*).

Il a donc fallu écrire un programme qui met ces mots en minuscules tout en conservant la majuscule initiale des noms propres et des débuts de phrase (FRANCE -> France).

Naturellement, l'automate ne peut tout faire et laisse à l'opérateur un nombre important de cas à résoudre. Le plus fréquent est "A" : à (préposition) ou a (verbe avoir) ou A ? Dans le corpus Giscard d'Estaing, il y a 15 142 prépositions 'à', 5 068 'a' (verbe avoir) et 2 'A' majuscule (deux fois à propos des fonctionnaires de catégorie A). Pour mettre en minuscule ce seul mot d'une lettre, il a donc fallu plus de 20 0000 interventions manuelles. Viennent ensuite "DU" (du ou dû ?), DES (des, dés ou dès), "MAIS" (mais ou maïs), "LA" (la, là), "OU" (ou, où) "SUR" (sur, sûr), etc. D'autres problèmes sont moins évidents comme "ELEVES" - élèves ou élevés ? – ou FRANÇAIS (E,S) : le nom de peuple prend une majuscule (les Français), mais pas le nom masculin ou l'adjectif (le *français*, la langue *française*) ? Naturellement, aucun automate ne peut résoudre ces problèmes puisque seul le sens de la phrase permet de trancher entre les différentes solutions…

De plus, les textes en ligne comportent des graphies non-conventionnelles qu'il faut corriger manuellement. Par exemple dans le texte en annexe : CONSEIL_CONSTITUTIONNEL ou EN-PARTICULIER comportent des tirets indésirables.

Une fois ce travail de correction et de standardisation effectué, le texte fait l'objet d'une série d'analyses automatiques dont l'une est spécialement dédiée au traitement des nombres.

1.2. Traitement des nombres

Les nombres posent des problèmes particulièrement redoutables. En voici quelques-uns.

1. Le même nombre peut s'écrire de multiples manières. Par exemple, dans le discours en annexe 2, on trouve '13396203' qui peut aussi s'écrire '13 396 203' – ici les espaces ne doivent pas être considérés comme des séparateurs de mots – voire '13,396,203' (à la mode anglo-saxonne que l'on rencontre souvent dans les textes québécois où la virgule ne doit pas être confondue avec la décimale ni avec le signe de ponctuation…) ou '13 millions 396 mille 203' et enfin on peut aussi écrire : 'treize millions trois cent quatre vingt seize mille deux cent trois'). Naturellement, il faut garder trace de la graphie originale qui doit toujours être retournée au chercheur qui consulte la bibliothèque. Il est tout aussi important que, dans l'analyse '2' et 'deux' ou 'kg', 'Kg' et 'kilogramme(s) ' soient toujours reconnus comme étant les mêmes mots. Pour faire comprendre l'importance de ce point : quelle valeur aurait un recensement dans lequel le même individu serait compté pour trois ou quatre sous prétexte que les enquêteurs l'ont rencontré une fois en costume de ville, une autre fois avec des lunettes de soleil et enfin en robe de chambre… Pourtant, n'est-ce pas ce que font les logiciels d'analyse des données textuelles lorsqu'ils décomptent des "formes graphiques" sans aucune opération de standardisation préalable ?

2. Un nombre, un mot ou plusieurs ? Par exemple, dans le discours en annexe, on trouve : '19 mai 1974' : une seule unité ou trois ? Voire sept ou huit mots (s'il est écrit en toutes lettres) ? Logiquement, il serait important que ce bloc ne soit pas dissocié afin d'être recensé comme tel, mais aussi par exemple afin de pouvoir fournir toutes les formules semblables au chercheur qui s'intéresserait à l'utilisation des dates par les hommes politiques. On remarquera également que le début du nombre peut ne pas être un chiffre. Ainsi dans 'mars 1978', la date commence à 'mars' (à ne pas confondre avec le dieu de la guerre qui prend une majuscule)…

3. Le nombre ne comporte pas seulement des chiffres mais aussi des unités. Dans l'exemple ci-dessus, il s'agit d'une unité de temps (le mois). On peut aussi rencontrer des longueurs (mètre, kilomètre… année-lumière), des masses (gramme, kilogramme, tonne, mégatonne), des volumes (litre, hectolitres, baril…) des puissances (watt, joule…), des monnaies (franc, livre, mark, euro, dollar…), etc. Ces unités sont elles-mêmes susceptibles de s'écrire de multiples manières (F, €, £, $). On a donc dressé la liste de toutes les unités de

mesure – avec leurs variantes orthographiques et abréviations - susceptibles d'être combinées avec des chiffres pour obtenir une quantification…

4. Les ordinaux sont aussi des outils de quantification. Par exemple, dans le discours en annexe, on trouve "le vingtième président de la République", "le premier salut". Vingtième et second doivent être recensés comme des formes particulières de quantification (le classement). Mais alors, il faut les séparer des substantifs homographes et ne pas compter comme nombres : un jeune premier, une première de théâtre, le second (capitaine).

5. les nombres ne sont pas tous des ordinaux ou des cardinaux. Certains mots, bien qu'entrant dans la composition des chiffres, ne sont pas des cardinaux ou des ordinaux : "million", "milliard", "billion", etc. En effet, tous les dictionnaires de langues reconnaissent dans ces mots des substantifs - parce qu'ils ont un genre et qu'ils s'accordent en nombre – et non pas des cardinaux (invariables).

6. La quantification peut être plus ou moins précise (pour un tableau d'ensemble : Charaudeau 1992). Même la quantification chiffrée n'est pas toujours aussi précise qu'on le pense quand on considère la préférence de la plupart des locuteurs pour les chiffres "ronds" (Jansen & Pollmann 2001). Suivant une définition large, dans le texte en annexe 2, il faudra voir de la quantification dans des expressions comme : "ère nouvelle" (2 fois), "longue histoire", "totalité des suffrages", "suffrages égaux", "les moins jeunes", "l'immense rumeur". Des expressions de ce genre signalent en effet une quantité aussi vague soit-elle… C'est pourquoi nous proposons de distinguer la quantification chiffrée (ou explicite), de la quantification non chiffrée (ou implicite). Cette communication traitera surtout de la première.

Ceci n'est qu'un bref aperçu des problèmes posés par la standardisation des graphies et le recensement de la quantification dans les textes en français moderne. Pourtant c'est au prix de conventions rigoureuses que l'on peut obtenir un recensement fiable qui, seul, peut répondre à la question posée.

1.3. Lemmatisation

Les conventions retenues se calquent au plus près sur les usages dans la population générale (Labbé 1990). Au texte original – qui n'est pas modifié – est associé un fichier image (lemmatisé) contenant une ou plusieurs étiquettes pointant vers chaque mot de ce texte original et comportant son entrée de dictionnaire. Les nombres sont traités de la manière suivante. Nous utilisons les deux exemples tirés de la première allocution de V. Giscard d'Estaing le jour de son investiture (tableaux 1 et 2).

- 'Le' est reconnu comme n'étant pas un nom propre – ce qu'il peut être par exemple dans 'Le Corbusier' ou 'Le Drian' – mais comme un article doté d'une majuscule car il est placé en début de phrase. Dans le fichier image, son lemme est donc "le, article". En effet, c'est à cette entrée qu'on le trouve dans les dictionnaires de langue.
- Une balise, pointant entre 'le' et '19' indique le début du nombre et une autre balise, placée après '1974', borne la fin du nombre. Comme il s'agit d'une date, il est logique que le mois (mai) figure entre les deux balises. Il en est ainsi de toutes les unités comme les mètres, les tonnes, les kilomètres, les heures, les euros, etc.
- '19' est reconnu comme la combinaison de deux adjectifs numéraux. Certains dictionnaires comportent une entrée à 'dix-neuf' (au même titre que 'dix-sept' et 'dix-huit'). Cependant, ces ouvrages n'en comportent pas pour 'vingt-neuf', 'trente-neuf'… 'cent neuf', etc.), ce qui nous a conduit à considérer qu'il s'agit de deux mots. Dans le traitement de ces nombres,

on a également renoncé aux traits d'union dans un souci d'harmonisation et de simplification : '90' reçoit trois étiquettes (quatre vingt dix).

- '1974' est doté de cinq étiquettes car il n'a pas d'entrée dans les dictionnaires de langue et doit donc être décomposé en autant d'éléments.

Tableau 1. Etiquetage de "le 19 mai 1974".

Texte original	Etiquettes :	
	lemme	catégorie grammaticale
Le	le	article
	<Nombre 19 mai 1974>	balise
19	dix neuf	adjectif numéral adjectif numéral
mai	mai	nom masc.
1974	mille neuf cent soixante quatorze	adjectif numéral adjectif numéral adjectif numéral adjectif numéral adjectif numéral
	<Fin nombre>	balise

Remarque : le texte original n'est pas modifié ('Le' conserve sa majuscule, 19 et 1974 sont toujours écrits en chiffres) mais le fichier image permet de les retrouver sous leur entrée du dictionnaire, c'est-à-dire à leur place dans le lexique du français, et tels que les ont entendus les auditeurs du discours.

Tableau 2. Etiquetage de : "13396293 (hommes et femmes)"

mot	Etiquettes :	
	Lemme	catégorie grammaticale
13396203	<Nombre 13396203> treize million trois cent quatre vingt seize mille deux cent trois <Fin nombre>	balise adjectif numéral nom masculin adjectif numéral adjectif numéral adjectif numéral adjectif numéral adjectif numéral adjectif numéral adjectif numéral adjectif numéral adjectif numéral balise

L'automate place une balise pointant vers le nombre – dans laquelle il conserve la graphie originale – et le convertit en lettres suivant les conventions du français, puis il signale la fin du nombre par une autre balise. Ici, il a consulté la liste des unités conventionnelles et n'ayant pas trouvé dans cette liste "hommes et femmes", il ne les a pas inclus dans le nombre.

Ces étiquettes permettent une étude de la quantification explicite.

2. Le poids des chiffres

L'importance de la quantification explicite dans un texte peut être estimée par la densité des numéraux et ordinaux (chiffres). Le tableau 3 ci-dessous indique ces densités dans les différents corpus composant la bibliothèque avec quelques détails pour la section consacrée au discours politique. Etant donné la faiblesse relative de ces densités, elles sont exprimées non pas en pourcentage mais en "pour mille mots".

Afin d'identifier les principales caractéristiques des corpus, on donne également les proportions occupés par les substantifs et leurs satellites (groupe du nom) et par ceux des verbes (groupe du verbe). Le groupe du nom comporte les substantifs, les adjectifs, les déterminants et les prépositions. Le groupe du verbe comprend, outre les verbes, les pronoms, les adverbes et les conjonctions de subordination. Certes, le partage n'est pas absolu : on trouve des adverbes dans le groupe du nom (notamment devant l'adjectif) ; il y a des prépositions associées au verbe, etc. Mais dans tous les corpus en langue française, de nettes corrélations existent entre les éléments constitutifs de ces deux groupes.

Tableau 3. Le poids des chiffres dans l'ensemble de la bibliothèque du français contemporain (pour mille mots)

	densité des chiffres	densité du groupe du nom	densité du groupe du verbe
Discours politique	21.3	654,8	343,3
Présidents français	*19,4*	*616,2*	*381,2*
Premiers ministres	*20,5*	*714,5*	*284,6*
Campagne 2012	*23,6*	*648,0*	*349,5*
Littérature	7,8	579,6	416,3
Français oral	14,9	441,7	544,1
Presse	60,4	772,1	225,9
Sciences	121,0	826,9	155,5
Ensemble	20,8	616,2	379,4

Le tableau se lit de la manière suivante : sur mille mots utilisés dans le discours politique, 21,3 en moyenne sont des chiffres ; 654,8 (soit près des deux tiers appartiennent au groupe du nom) et 343,3 (soit environ le tiers) appartiennent au groupe du verbe (le total est inférieur à 1 000 car les locutions et les mots étrangers ne sont pas comptés).

Deux tendances apparaissent. D'une part, le genre semble un facteur déterminant. L'oral est dominé par le groupe du verbe ; l'écrit par le groupe du nom. Le discours politique appartient manifestement à l'écrit, même si une bonne partie de ces textes sont des transcriptions de l'oral (conférences de presse, entretiens télévisés, etc.).

D'autre part, quant aux chiffres, la littérature ou l'oral en utilisent relativement peu (et tous deux sont également caractérisés par une densité du groupe du nom inférieure à la moyenne). A l'inverse, l'information ou la communication scientifique sont de gros utilisateurs de chiffres mais aussi de tous les composants du groupe du nom. On peut donc supposer que la quantification est d'abord l'apanage des textes tendant à la transmission d'informations, de connaissances et que le substantif (et ses satellites) sont les principaux vecteurs de cette transmission. Le discours politique semble à mi chemin et les différentes sections de cette partie de la bibliothèque paraissent assez proches (avec toutefois, pour les Premiers ministres, une préférence plus marquée pour le nom et une densité un peu plus importante des nombres dans le discours électoral).

Un discours orienté vers la connaissance comporte plus de nombres et de substantifs (ainsi que les unités qui leur sont associées). En revanche, des propos spontanés et prononcés dans une situation d'interlocution – donc dominés par la subjectivité ou l'intersubjectivité -, sont dominés par le verbe (et les pronoms, les adverbes, les conjonctions de subordination…) et comportent peu de nombres.

Cependant, ces moyennes ne cachent-elles pas des différences individuelles ? Le tableau 4 apporte un élément de réponse en ce qui concerne les présidents français.

Tableau 4. Densité des chiffres (ordinaux et cardinaux) chez les 6 présidents (ordre chronologique)

Présidents	Densité	Rapport aux 5 autres (%)
C. De Gaulle	17.22	-11,7
G. Pompidou	15.63	-20,3
V. Giscard d'Estaing	**25.76**	+40.9
F. Mitterrand	**28.01**	+50.7
J. Chirac	12.16	-38.7
N. Sarkozy	17.85	-18,1
Total présidents	19.40	-

Pour l'ensemble des présidents, la densité des nombres est de 19,40 pour mille mots. Mais la seconde colonne du tableau indique que les différences entre les présidents sont très importantes. Voici un exemple du calcul de ce rapport aux autres pour de Gaulle :

- dans le corpus constitué par les textes des 5 autres présidents, la densité des chiffres est de 19,51 pour mille mots ;

- dans le corpus de de Gaulle elle est de 17,22 pour mille mots

Le rapport 17,22/19,51 = 0 ,883 montre que, là où les 5 autres utilisent en moyenne 100 nombres, de Gaulle en emploie seulement 88…

Dans l'ensemble, c'est sur cette catégorie – et sur celle des noms propres - que l'on enregistre les variations les plus fortes. Pour autant peut-on considérer ces variations comme significatives ? Et, dans l'affirmative, comment les expliquer ?

3. Les écarts sont-ils significatifs ?

Puisqu'il est question de quantification, une question est évidente : peut-on se tromper en affirmant que V. Giscard d'Estaing se singularise pour ce qui concerne l'utilisation des chiffres ? Pour répondre, il faudrait disposer de l'équivalent, pour les pratiques langagières, du recensement de la population française ou des relevés pluviométriques dans l'hexagone que nous évoquions au début de cet exposé. Faute de disposer de ces étalons de comparaison, on doit se contenter de comparer les présidents entre eux, sans pouvoir tirer de conclusions générales.

Un test statistique permet de choisir – avec un risque d'erreur qu'il s'agit de minimiser - entre deux hypothèses contradictoires :

- H_0 : la densité d'emploi des nombres par V. Giscard d'Estaing ne diffère pas significativement de celle de tous les autres, ou encore : les écarts ne sont probablement pas "anormaux". Cela équivaut à la situation suivante : le corpus des présidents est une "urne" et le corpus "VGE" un "échantillon" extrait, sans remise, de cette urne ;

- H_1 : V. Giscard d'Estaing présente une propension à utiliser les chiffres différente des autres. Le corpus "VGE" ne peut être considéré comme un "échantillon" du discours présidentiel.

Accepter une hypothèse (avec un certain risque d'erreur) ne signifie pas qu'elle est "vraie" mais simplement que l'on peut écarter l'hypothèse contraire (Desrosières 1988). Par exemple, accepter H_0 (le phénomène est le même chez tous les présidents) signifie que les différences constatées sont considérées comme des fluctuations propres à tout phénomène naturel (rejet de H_1). A l'inverse, accepter H_1 revient à dire que l'écart entre les observations et les valeurs attendues est si fort que l'on peut rejeter H_0 avec un risque d'erreur minime et chiffré.

Voici le calcul qui permet de valider l'une ou l'autre des deux hypothèses et d'estimer le risque d'erreur. Soit :

— l'urne (C) composée de N_c mots. Ici les 3 824 965 mots contenus dans le corpus des présidents ;

— l'échantillon (B) tiré de l'urne (avec $B \subset C$), échantillon composé de N_b mots (les 566 519 mots du corpus VGE) ;

— une catégorie grammaticale i (les adjectifs numéraux) avec F_{ic} occurrences dans C et F_{ib} dans B (75 263 dans le corpus présidents et 14 595 dans VGE) ;

Si les mêmes lois de composition sont à l'œuvre dans l'urne C et dans l'échantillon B, alors l'effectif théorique de la catégorie grammaticale i dans l'échantillon (E_{ib}) sera le nombre de ses occurrences dans l'urne pondéré par le taux de sondage U (rapport entre la taille de B et celle de C) :

$$E_{ib(u)} = F_{ic} * U \quad avec \;\; U = \frac{N_b}{N_c} \qquad E_{ib(u)} = 75\,263 \frac{566\,519}{3824\,965} = 11\,147$$

Le nombre d'occurrences constatées (F_{ib} = 14 595) est supérieur à celui attendu (E_{ib} = 11 147), peut-on affirmer que ce suremploi chez V. Giscard d'Estaing est caractéristique par rapport aux autres ? Pour répondre à cette question, il faut considérer la probabilité de l'événement observé F_{ib} par rapport à l'événement attendu (E_{ib}). Cette probabilité est la combinaison de deux événements :

— le nombre de possibilités différentes de choisir N_b mots dans un total de N_c :

$$C_{N_c}^{N_b} = \frac{N_c!}{N_b!(N_c - N_b)!} = \begin{bmatrix} N_c \\ N_b \end{bmatrix}$$

— le nombre de possibilités différentes de choisir F_{ib} mots dans un total de F_{ic} :

$$C_{F_{ic}}^{F_{ib}} = \frac{F_{ic}!}{F_{ib}!(F_{ic} - F_{ib})!} = \begin{bmatrix} F_{ic} \\ F_{ib} \end{bmatrix}$$

La probabilité composée de ces deux événements suit une loi hypergéométrique dont les paramètres sont : F_{ic}, F_{ib}, N_b, N_c:

$$P(X = F_{ib}) = \frac{\begin{bmatrix} F_{ic} \\ F_{ib} \end{bmatrix} \begin{bmatrix} N_{ic} - F_{ic} \\ N_{ib} - F_{ib} \end{bmatrix}}{\begin{bmatrix} N_c \\ N_b \end{bmatrix}} \quad (1)$$

Appliquée à V. Giscard d'Estaing, comparé à l'ensemble des présidents, la formule (1) donne $P(X = 14\,595) < 1.0^{e-12}$ ou encore 0.000000000001. C'est-à-dire que l'on a moins d'une chance sur mille milliards de se tromper en considérant l'écart comme significatif…

Le même calcul est appliqué à chaque président, considéré par rapport à l'ensemble et à chacun des autres. Tous les écarts sont significatifs à des seuils d'erreur inférieurs à 1‰.

Il est donc possible d'affirmer que, en l'état des corpus disponibles, deux présidents se singularisent par un suremploi remarquable des chiffres : V. Giscard d'Estaing et F. Mitterrand et que les autres se caractérisent par une sobriété plus ou moins grande. La palme revient à J. Chirac, puis l'on trouve : G. Pompidou, C. de Gaulle, N. Sarkozy.

A priori, la fonction présidentielle ne change pas entre 1958 et 2012 et les situations d'énonciation ne sont guère différentes. Dès lors, comment expliquer ces écarts considérables ? Cette singularité peut-elle être reliée à d'autres dans le vocabulaire et dans le contenu du discours ? Pour répondre à ces questions, on examine d'abord le vocabulaire caractéristique de chaque président.

4. La quantification dans le vocabulaire caractéristique

Le calcul du vocabulaire caractéristique est présenté dans : Labbé & Labbé 1994 ; Monière, Labbé & Labbé 2005 ; Monière & Labbé 2012. Le raisonnement est le même que celui présenté ci-dessus. Les effectifs de chacun des vocables dans le corpus VGE (l'échantillon) sont comparés à ceux de ce même vocable chez l'ensemble des présidents (l'urne). Pour chaque vocable, l'indice calculé à l'aide de la formule (1) permet de choisir entre les deux hypothèses contradictoires. Plus l'indice est faible plus le vocable sera considéré comme caractéristique de V. Giscard d'Estaing. Les mêmes calculs portent également sur les catégories grammaticales et sur les phrases les plus caractéristiques.

4.1. Les vocables

Le calcul montre que V. Giscard d'Estaing emploie beaucoup plus *France* et *Français(e,es)* que les autres. Le verbe qui singularise le plus ce président est *falloir* (*il faut*) suivi de *savoir*. Viennent ensuite : *croire, souhaiter, concerner, connaître, conduire, répondre, considérer, indiquer, rappeler, prévoir, observer, penser*… A part *conduire* (la politique de la France), ces verbes ne sont pas à proprement parler des verbes d'action. En revanche, beaucoup sont des quasi-synonymes de *savoir* (*croire, connaître, prévoir, observer*…).

A l'inverse voici les verbes qu'il utilise significativement moins que les cinq autres : *vouloir, travailler, aller, engager, devoir, faire, dire, assumer, penser, oublier, refuser, changer*…

Les listes complètes sont plus longues mais ces extraits sont significatifs. Les combinaisons de verbes sont encore plus parlantes. Ces "modalités" verbales ont été présentées en 2010 devant ces mêmes journées de l'ERLA (Labbé & Labbé 2010).

Les modalités préférées de V. Giscard d'Estaing sont la nécessité (*falloir*) combinée à la connaissance (*il faut savoir, croire, connaître, considérer, indiquer*). "Il faut savoir" est donc la formule favorite de ce président. Si l'on songe que la quantification est le principal vecteur de la connaissance, on comprend la manière dont le président construit ces propos : présentation à l'auditoire de la *situation de la France* – étayée par beaucoup de nombres - puis conclusions concernant la meilleure politique possible, compte tenu de cette situation.

Les modalités évitées sont celles de la volonté (*vouloir faire*) et de l'obligation juridique, morale ou politique exprimée par *devoir faire*. Autrement dit, V. Giscard d'Estaing *indiquait* à ses auditoires les éléments *connus* qui devaient *conduire la politique* et non pas une volonté, une morale ou le droit…

Les outils de la quantification viennent au premier plan des substantifs privilégiés par le président. Outre *franc, dollar* et *deutschemark, centaine, millier, million et milliard*, on trouve dans l'ordre : *problème, heure, situation, nombre, question, année, gouvernement, politique, vie, an, fait, action, prix, mois, cours, hausse, chiffre, inflation*, etc.

La liste des adjectifs caractéristiques dévoile une pensée qui fonctionne par couples d'oppositions dont beaucoup servent à classer ou à quantifier : *national/étranger* ; *français/étranger* ; *intérieur/extérieur* ; *jeune/âgé* ; *normal/excessif* ; *supérieur/inférieur* ; *positif/négatif* ; *rapide/lent* ; *haut/bas* ; *faible/fort* ; *maximum/minimum* ; *prochain/lointain* ; *actuel/ passé* ; *antérieur/ultérieur* ; *brusque/progressif*, etc.

Il y a également des familles de qualificatifs dont certains peuvent être rapprochés de la quantification : *continu*, *régulier*, *uniforme*, *répétitif*… *quotidien*, *mensuel*, *annuel*, etc.

Enfin, les substantifs et adjectifs entrent dans des associations presqu'insécables comme '*économique et social*'. Ces syntagmes répétés (Pibarot et Al. 1988) permettent d'identifier des thèmes chers à ce président. Outre *il faut savoir (*et *il faut que les Français sachent*), la *situation économique* (et *sociale*), on trouve : *énergie* et *nucléaire* (substantif et adjectif), *travail* (et *travailleur*) *manuel*, *cours* du *franc* ou du *pétrole*, *condition* de la *femme*, etc.

4.2. La préférence pour le nom

Le même raisonnement est appliqué à toutes les catégories grammaticales (tableaux 5 et 6). Ici les cinq autres présidents sont l'étalon de comparaison (colonne A).

Tableau 5. Densités des catégories grammaticales chez V. Giscard d'Estaing comparé aux cinq autres présidents (‰)

Catégories	A (Autres présidents) ‰	B (V. Giscard d'Estaing) ‰	(B-A)/A %
Verbes	159.92	150.16	-6.1
Formes fléchies	*99.58*	*96.03*	*-3.6*
Participes passés	*22.70*	*23.65*	*+4.2*
Participes présents	*2.17*	*2.47*	*+13.6*
Infinitifs	*35.46*	*28.02*	*-21.0*
Noms propres	25.51	18.27	-28.4
Noms communs	174.06	186.75	+7.3
Adjectifs	54.45	58.46	+7.4
Adj. participe passé	*5.30*	*5.12*	*-3.4*
Pronoms	128.77	117.48	-8.8
Pronoms personnels	*66.92*	*61.45*	*-8.2*
Pronoms démonstratifs	*20.67*	*17.21*	*-16.7*
Pronoms possessifs	*0.34*	*0.23*	*-30.5*
Pronoms indéfinis	*4.22*	*2.68*	*-36.5*
Pronoms relatifs	*26.96*	*26.29*	*-2.5*
Déterminants	178.33	193.28	+8.4
Articles	*128.84*	*133.65*	*+3.7*
Nombres	*18.28*	*25.76*	*+40.9*
Possessifs	*14.10*	*12.89*	*-8.6*
Démonstratifs	*7.56*	*10.47*	*+38.6*
Indéfinis	*9.55*	*10.51*	*+10.1*
Adverbes	71.08	58.81	-17.3
Prépositions	147.91	156.52	+5.8
Conjonctions	57.33	57.91	+1.0
Coordination	*30.77*	*32.77*	*+6.5*
Subordination	*26.56*	*25.14*	*-5.3*

Le test statistique présenté ci-dessus révèle que, sauf pour les pronoms relatifs, tous les écarts sont significatifs (avec moins de 1% de chance d'erreur). Pour les conjonctions, l'écart est significatif si l'on sépare la coordination (sur-employée) et la subordination (sous-employée).

Au total, V. Giscard d'Estaing privilégiait les mots appartenant au groupe du nom (et ses satellites) au détriment de ceux du groupe du verbe (tableau 7).

Tableau 7. Poids des groupes du nom et du verbe chez V. Giscard d'Estaing comparé aux 5 autres présidents

Groupes de catégories	A (Autres présidents) ‰	B (V. Giscard d'Estaing) ‰	(B-A)/A %
Groupe verbal	386.31	351.58	-9.0
Groupe nominal	611.04	646.05	+5.7

Dans le détail, l'écart positif maximal est enregistré sur les nombres (+40,9) puis les adjectifs démonstratifs (*ce* chiffre, *cette* politique…). V. Giscard d'Estaing *indique* aux *Français* les chiffres *qu'il faut savoir*. Les dates ancrent le discours dans le temps ; les autres nombres offrent des représentations des phénomènes sociaux et économiques, etc.

Chez V. Giscard d'Estaing, cette préférence pour les nombres est à mettre en relation avec la forte densité des participes présents. Cressot (1963) a signalé les caractéristiques particulières du participe présent, forme verbale la plus proche de l'adjectif : "Cette forme a pris au XIXe siècle un développement considérable, surtout à partir de Flaubert. Les écrivains qui attribuent aux choses une vie et une volonté secrètes, ont compris l'utilité de l'adjectif verbal pour leur expression dynamique du monde, et la possibilité d'atténuer, grâce à lui, la note trop éclatante des adjectifs en –eur et en –teur ". Ici les "choses animés d'une vie secrète" ce sont les évolutions de l'économie et de la société française telles qu'elles sont mesurées par une batterie d'indicateurs chiffrés.

A l'inverse, les mots à majuscules (patronymes, toponymes, sigles des organisations…) sont les plus significativement sous-employés, avec les pronoms indéfinis et démonstratifs. Autrement dit, il y a chez ce président une très forte réticence à personnaliser et, spécialement à nommer les acteurs de la vie politique. Or, ces mots à majuscules sont des sortes d'interfaces entre le discours et la réalité extérieure à celui-ci. Ils assurent l'ancrage des propos dans l'espace géographique, économique, social. Ils sont également très présents dans le discours polémique parce qu'il est centré sur la dénonciation du ou des adversaires et par une intense personnalisation du propos.

Si V. Giscard d'Estaing utilisait plus de nombres que ses prédécesseurs, cela suffisait-il à compenser la faible personnalisation, l'abstraction relative de ses propos et à lui assurer l'image d'un politicien pragmatique à laquelle il tenait tant ?

4.3. Un discours dépersonnalisé

Les pronoms suivent les mouvements affectant la densité des verbes en les amplifiant. Chez Giscard d'Estaing, c'est particulièrement vrai des pronoms indéfinis (*certain*) et possessifs (*mien*, *nôtre*) mais aussi des principaux pronoms personnels (tableau 6)

Dans un discours dominé par les chiffres, la première personne (du singulier et du pluriel) est moins présente. A l'inverse, la troisième personne (*il* et *ils*) supplante cette première personne. On remarque avec Benveniste (1970) que cette troisième personne est par excellence la "non-personne", spécialement dans les emplois impersonnels (*il faut*…) et l'on se souviendra que la faible personnalisation est une caractéristique du discours politique (Labbé & Labbé 2011) que V. Giscard d'Estaing amplifie encore.

Tableau 6. Poids des principaux pronoms chez V. Giscard d'Estaing comparé aux 5 autres présidents

Pronoms personnels	A (Autres présidents) ‰	B (V. Giscard d'Estaing) ‰	(B-A)/A %
je	18.69	16.25	-13.1
moi	1.07	0.51	-52.0
mien	0.07	0.04	-43.4
il (& elle)	13.01	16.12	+24.0
lui (& elle)	0.69	0.63	-9.0
se	6.45	6.03	-6.5
sien	0.06	0.04	-36.1
on	7.79	4.48	-42.4
nous	8.94	6.98	-22.0
nôtre	0.13	0.13	-2.5
vous	6.46	6.52	+0.8
vôtre	0.08	0.03	-58.8
ils (& elles)	2.89	3.10	+7.1

Plus généralement, on notera aussi que certaines caractéristiques se retrouvent dans le discours électoral où la polémique se cache, plus ou moins bien, derrière un discours apparemment dépersonnalisé et informatif du moins tant que le discours négatif (polémique) ne l'emporte pas sur l'explication (Savoy 2010, Monière & Labbé 2013).

4.4. Phrases caractéristiques

Le logiciel relit l'ensemble des textes de V. Giscard d'Estaing et attribue à chaque phrase un score en fonction du nombre de vocables caractéristiques qu'elle contient. Les phrases dont le score est le plus élevé synthétisent le mieux les thèmes favoris et le style du locuteur. Elles sont un peu l'équivalent des citations illustratives d'un article de dictionnaire. La phrase ci-dessous est la plus caractéristique du président (elle contient 91 vocables significativement sur-employés par lui). Sans surprise, elle est marquée par un haut degré de quantification et s'adresse à la profession la plus favorable au président (elle a été prononcée dans son allocution d'ouverture du salon de l'agriculture le 3 mars 1975).

> "Le gouvernement n'a pas attendu d'y être contraint pour prendre des mesures de soutien des revenus de l'agriculture ; après s'en être entretenu avec les dirigeants qualifiés du monde agricole, il a pris, vous le savez, la semaine dernière, un certain nombre de décisions que je rappelle et qui consistent à ajouter à la progression des revenus entraînés par les décisions de prix de Bruxelles un effort national complémentaire qui prendra la forme de deux versements de deux primes ; une prime de 160 francs par tête de bétail pour les 45 premières vaches de chaque troupeau - plus exactement pour les 15 premières bêtes de chaque troupeau - et d'autre part une prime spéciale agricole de 1200 francs versée à chaque exploitant à l'exception de ceux qui ont les revenus les plus élevés ; cette contribution budgétaire représente 2 milliards 700 millions de francs et elle permettra d'assurer aux agriculteurs, d'après les calculs les plus précis, la parité de l'évolution de leurs revenus en 1975 par rapport aux autres catégories de Français, c'est-à-dire en fait, la garantie de leur pouvoir d'achat ; je voudrais dire aussi que les dispositions ont été prises par le ministre de l'agriculture et par le ministre des finances pour que ces primes puissent être versées très rapidement - quand je dis très rapidement je veux dire dans un délai maximum de deux mois et si possible dans les six prochaines semaines, puisqu'il s'agit d'un complément de revenu, il convient de faire en sorte que les agriculteurs puissent en bénéficier le plus rapidement possible."

Un détail prouve qu'il ne s'agit pas du texte rédigé à l'avance mais de celui qui a été effectivement prononcé et transcrit à partir de l'enregistrement. En effet, le président annonce une prime par tête de bétail mais il se trompe sur le plafond de cette aide et se reprend : "pour les 45 premières vaches de chaque troupeau - plus exactement pour les 15 premières bêtes de chaque troupeau". Ce lapsus montre évidemment les dangers de l'abus des chiffres : celui de commettre des erreurs ou de noyer son auditoire sous les détails, risques que ne court pas l'orateur qui s'en tient aux formules générales (comme J. Chirac).

Le logiciel isole également plusieurs phrases construites sur le thème de la compétition entre pays et sur leur classement (c'est-à-dire une quantification ordinale). Voici la plus caractéristique qui compare le développement économique à une course cycliste :

> "A l'heure actuelle, il y a un grand problème, dans la vie politique de la France et la vie sociale de la France, c'est cette inflation car, du sort de cette inflation, dépend le fait que la France rejoindra, comme elle doit le faire, le groupe des pays les plus avancés et les plus actifs du point de vue industriel, c'est-à-dire le groupe des Etats-Unis, l'Allemagne et le Japon, qui est actuellement en tête du développement économique mondial ou qu' au contraire elle rejoindra le peloton moins brillant et moins heureux de ceux qui connaissent actuellement des difficultés économiques" (25 août 1976).

Enfin le procédé de quantification le plus courant chez V. Giscard d'Estaing consiste à effectuer un rapprochement entre plusieurs grandeurs afin d'éclairer son auditoire. Naturellement, ces rapprochements ne sont pas choisis au hasard ! En voici deux très révélateurs, à propos de la crise pétrolière qui suit la révolution iranienne. Les passages sont cités intégralement afin d'illustrer le procédé et son véritable but.

> "Pour la France, l'augmentation des prix du pétrole - la dernière augmentation a été décidée le 26 juin 1979, il y a moins de trois mois - va représenter un débours supplémentaire de 30 milliards de francs, c'est-à-dire que pour avoir la même quantité de pétrole, il va falloir que nous payions à l'étranger 30 milliards de francs de plus.
> Il est très difficile, à partir des finances personnelles, d'imaginer ce que sont trente milliards. Voici deux exemples.
> Je m'adresse aux agriculteurs, cela représente la même situation que si on venait de l'extérieur moissonner la totalité des céréales en France, récolter la totalité des plantes fourragères et des betteraves et qu'on emportait le tout sans payer, la totalité de la moisson annuelle et des autres cultures fourragères et industrielles emportées sans payer !
> Pour les travailleurs de l'industrie, ceci représente à peu près la même chose que si l'on était venu prendre la totalité de la production annuelle de voitures de la régie Renault et qu'on les ait emportées sans payer à l'étranger. C'est la même chose" (17 septembre 1979)

> "Effectivement, ces événements ont des conséquences tout à fait sérieuses sur le plan économique et social. Elles tiennent essentiellement à deux causes : la première, c'est le renchérissement du prix du pétrole, avec tout ce qu'il entraîne. Notre facture pétrolière va représenter cette année presque autant que le total de l'impôt sur le revenu payé par les Français. Cet argent allant à l'étranger, c'est comme si les Français payaient l'équivalent de deux fois l'impôt sur le revenu. Ceci est considérable" (26 février 1980)

Dans ce type d'explication, le président restait apparemment toujours dans le registre de la comparaison. Comme l'indiquent Dumarsais (1730) et Fontanier (1827), la comparaison s'affiche explicitement grâce à des formules de liaison comme : "cela représente la même situation que si", "ceci représente à peu près la même chose que si". Mais, chez les hommes politiques, la métaphore – c'est-à-dire l'effacement du "comme si" - n'est jamais loin de la comparaison. Par exemple, dans cette phrase : "la totalité de la moisson annuelle et des autres cultures fourragères et industrielles emportées sans payer" !

On remarque enfin que ces phrases sont construites sur des quantifications apparemment rigoureuses mais que le choix des exemples est hautement symboliques : le pain, l'automobile, l'impôt sur le revenu, tout cela est emporté sans payer par l'étranger !

Conclusions

Le recensement des chiffres utilisés par V. Giscard d'Estaing révèle ses deux dimensions préférées : le temps (*heure, jour, mois, année, an, siècle* mais surtout *histoire*) et les *faits* quantifiables spécialement en unités monétaires… En ce qui concerne la première de ces deux dimensions, on peut citer la double métaphore qui conclut son premier discours ("écrire l'histoire", "arriver dans un monde nouveau") :

> "Voici que s'ouvre le livre du temps avec le vertige de ses pages blanches. Ensemble comme un grand peuple uni et fraternel abordons l'ère nouvelle de la politique française".

Chez le vingtième président de la République, la quantification servait surtout à mettre en valeur l'orateur et à capter l'adhésion par des voies plus proches de la rhétorique classique que de la science moderne. Toutefois, la prégnance du modèle pédagogique est certaine. Elle se traduit par des sortes de leçons assez dépersonnalisées avec relativement moins de verbes, de pronoms et d'adverbes que chez les autres hommes politiques.

En effet, l'homme politique a deux registres à sa disposition : le pédagogique et le polémique. Le premier utilise nettement plus la quantification que le second. La densité relative des chiffres est même ce qui démarque le plus clairement ces deux registres. Le style pédagogique est dépersonnalisé, peu verbal et dominé par les mots appartenant au groupe du nom qui est le vecteur privilégié des notions et des informations factuelles. Ce modèle domine les propos de V. Giscard d'Estaing. A l'inverse, la polémique est fortement personnalisée - à l'aide des pronoms mais aussi des noms propres - elle privilégie le verbe, spécialement les modalités de la volonté, du devoir et du jugement, ainsi que les verbes de la pensée et de l'action.

De manière plus générale, la bibliothèque électronique du français moderne montre que la quantification est l'apanage de la science et que celle-ci est fortement dépersonnalisée. C'est pourquoi la quantification est également très abondante dans tous les textes tournés vers la transmission d'informations et notamment dans la presse ou dans le texte pédagogique. A l'inverse, la littérature ou l'oral (authentique), en utilisent peu car ils sont dominés par la subjectivité, les sentiments et par d'autres dimensions étrangères à la communication des connaissances.

Enfin, les disciplines littéraires comme la linguistique, la lexicologie ou la stylistique se sont détournées des recensements statistiques qui sont leur "ventre mou", selon l'expression de G. Molinié (1986). Nous espérons avoir suggéré combien la lexicométrie – c'est-à-dire l'alliance de la linguistique appliquée avec la statistique et l'informatique - offre aux sciences humaines des outils simples permettant de dépasser la simple intuition pour aboutir à des connaissances solides. Encore faudra-t-il disposer de véritables corpus du français contemporain, c'est-à-dire de vastes échantillons représentatifs des pratiques langagières dans la population générale. Car sinon, comment juger de la singularité du segment spécifique ou du locuteur que l'on souhaite étudier ?

Bibliographie

Bacot, Paul, Desmarchelier, Dominique & Rémi-Giraud, Sylvianne (2012) : "Le langage des chiffres en politique", *Mots. Les langages du politique*, 100,

Banks, David (1998) : "Vague quantification in the scientific journal article", *ASp*, 149-22, 17-27.

Benveniste, Emile (1970) : "La nature des pronoms", *Problèmes de linguistique générale*, Paris, Gallimard, I, 251-257.

Charaudeau, Patrick (1992) : *Grammaire du sens et de l'expression*, Paris, Hachette, 237-278.

Cressot, Marcel (1963) : *Le style et ses techniques*, Paris, PUF.

Desrosières, Alain (1988) : *La partie pour le tout : comment généraliser ? Cinq contributions à l'histoire de la statistique*, Paris, Economica.

Desrosières, Alain (1993) : *La politique des grands nombres*, Paris, La découverte.

Dumarsais, César Chesneau (1730) : *Des tropes*, Réédition Paris, Flammarion, 1988.

Fontanier, Pierre (1827) : *Les Figures du discours*, Réédition (avec une préface de Gérard Genette), Paris, Flammarion, 1968.

Jansen, C. J. M. & Pollmann, M. M. W. (2001) : "On round number : Pragmatic aspects of numerical expressions". *Journal of Quantitative Linguistics*, 8/3, 187-201.

Labbé, Cyril & Labbé, Dominique (1994) : Que mesure la spécificité du vocabulaire ? Grenoble : CERAT, décembre 1994 et juin 1997. Reproduit dans *Lexicometrica*. 3, 2001.

Labbé, Cyril & Labbé, Dominique (2010) : "La modalité verbale en français contemporain", *Communication aux XIe Journées de l'ERLA*, Brest, 19 novembre 2010.

Labbé, Cyril & Labbé, Dominique (2011) : "Existe-t-il un langage propre à la politique ?" *Communication aux XIIe Journées de l'ERLA*, Brest, 18-19 novembre 2011.

Labbé, Dominique (1990) : *Normes de saisie et de dépouillement des textes politiques*, Grenoble, Cahiers du CERAT.

Labbé, Dominique & Monière, Denis (2013) : *Votez pour moi ! La campagne présidentielle de 2012*, Paris, l'Harmattan.

Molinié, Georges (1986) : *Eléments de stylistique française*, Paris, PUF.

Monière, Denis, Labbé, Cyril & Labbé, Dominique (2005) : "Les particularités d'un discours politique : les gouvernements minoritaires de Pierre Trudeau et de Paul Martin au Canada", *Corpus*, 4, 79-104.

Monière, Denis & Labbé, Dominique (2012). "Le vocabulaire caractéristique du Premier ministre du Québec J. Charest comparé à ses prédécesseurs", In Dister, Anne, Longrée, Dominique & Purnelle, Gérald (éds) : *Proceedings of the 11th International Conference on Textual Data Statistical Analysis*, Liège, LASLA - SESLA, 2012, 737-751.

Pibarot, André, Picard, Jacques & Labbé, Dominique (1998) : "Les syntagmes répétés dans l'analyse des commentaires libres", In Mellet, Sylvie (ed) : *4e Journées d'analyse des données textuelles*. Nice, 1998, 507-516.

Savoy, Jacques (2010) : "Discours électoral et discours présidentiel", In Bolasco, Sergio & Al. (Eds) : *Proceedings of 10th International Conference Statistical Analysis of Textual Data*, Rome, Edizioni Universitarie di Lettere Economia Diritto, Vol 2, 827-838.

Annexe 1

Bibliothèque électronique du français moderne (1er novembre 2013)

	Longueur (mots)	Vocabulaire
Discours politique	11 529 763	42 885
Présidents français (1958-2012)*	3 824 965	23 602
Premiers ministres canadiens (1867-2012)	1 098 161	13 514
Premiers ministres québécois (1867-2012)	2 993 823	22 458
Premiers ministres français (1945-2012)	288 526	7 952
Littérature (XVIIe –XXe siècles)	10 903 628	56 192
Romans et nouvelles	6 202 751	48 365
Théâtre	2 571 497	15 551
Poésie	675 187	18 810
Correspondance	345 542	11 070
Romans policiers	548 682	17 274
Presse	2 939 632	58 690
Economie	645 341	13 964
Hydrologie	129 173	7 197
Français oral	2 978 122	18 429
Total	29 674 341	99 921

* Corpus des présidents : tableau ci-dessous

	Mandat	Nombre d'interventions	Nombre de mots	Vocabulaire
Charles de Gaulle*	1958-1969	79	201 927	6 551
Georges Pompidou**	1969-1974	122	216 809	7 651
Valéry Giscard d'Estaing*	1974-1981	171	566 519	9 194
François Mitterrand*	1981-1995	68	324 552	8 201
Jacques Chirac*	1995-2007	81	224 326	6 392
Nicolas Sarkozy***	1997-2012	726	2 290 832	19 022
		1 247	3 824 965	23 602

* Corpus en cours de constitution. Limité pour l'instant aux interventions radio-télévisées et conférences de presse.
** Corpus en cours de constitution. La plupart des textes ont été communiqués par D. Mayaffre.
*** Corpus exhaustif.

(…)

Annexe II.
Discours d'investiture de V. Giscard d'Estaing (27 mai 2013)

Original sur viepublique.fr
MESSIEURS LES PRESIDENTS, MESDAMES, MESDEMOISELLES, MESSIEURS, DE CE JOUR, DATE UNE ERE NOUVELLE DE LA POLITIQUE FRANCAISE. CECI N'EST PAS SEULEMENT DU, MONSIEUR LE PRESIDENT DU CONSEIL_CONSTITUTIONNEL, A LA PROCLAMATION DU RESULTAT QUE VOUS VENEZ DE RAPPELER ET DONT, PAR RESPECT POUR LA FRANCE ET POUR SA LONGUE HISTOIRE, JE MESURE L'HONNEUR. CECI N'EST PAS SEULEMENT DU AUX 13396203 FEMMES ET HOMMES QUI M'ONT FAIT LA CONFIANCE DE ME DESIGNER POUR DEVENIR LE VINGTIEME PRESIDENT DE LA REPUBLIQUE FRANCAISE. CECI EST DU EN REALITE A LA TOTALITE DES SUFFRAGES DU 19 MAI 1974. CES SUFFRAGES EGAUX SELON LA REGLE DEMOCRATIQUE QU'IL S'AGISSE DE CEUX DES FEMMES ET DES HOMMES, DES JEUNES ET DES MOINS JEUNES, DES TRAVAILLEURS ET DES INACTIFS, ET QUI SE SONT PRONONCES CHACUN A LEUR MANIERE ET SELON LEUR PREFERENCE EN TEMOIGNANT LEUR VOLONTE DE CHANGEMENT. J'ADRESSE LE PREMIER SALUT DU NOUVEAU PRESIDENT DE LA REPUBLIQUE A CEUX QUI DANS CETTE COMPETITION ASPIRAIENT A LE DEVENIR ET QUI AVAIENT LA CAPACITE DE LE FAIRE ET NOTAMMENT M. FRANCOIS MITTERRAND ET M. JACQUES CHABAN-DELMAS. AINSI C'EST MOI QUI CONDUIRAI LE CHANGEMENT, MAIS JE NE LE CONDUIRAI PAS SEUL. SI J'ENTENDS ASSUMER PLEINEMENT LA TACHE DE PRESIDENT, ET SI J'ACCEPTE, A CET EGARD, LES RESPONSABILITES QU'UNE TELLE ATTITUDE IMPLIQUE, L'ACTION A ENTREPRENDRE ASSOCIERA LE GOUVERNEMENT DANS SES INITIATIVES ET LE PARLEMENT DANS SON CONTROLE ET DANS SES DROITS. JE NE LE CONDUIRAI PAS SEUL PARCE QUE J'ECOUTE ET QUE J'ENTENDS ENCORE L'IMMENSE RUMEUR DU PEUPLE FRANCAIS QUI NOUS A DEMANDE LE CHANGEMENT. NOUS FERONS CE CHANGEMENT AVEC LUI, POUR LUI, TEL QU'IL EST DANS SON NOMBRE ET DANS SA DIVERSITE, ET NOUS LE CONDUIRONS EN-PARTICULIER AVEC SA JEUNESSE QUI PORTE COMME DES TORCHES LA GAIETE ET L'AVENIR. MESSIEURS LES PRESIDENTS, MESDAMES, MESDEMOISELLES, MESSIEURS, VOICI QUE S'OUVRE LE LIVRE DU TEMPS AVEC LE VERTIGE DE SES PAGES BLANCHES. ENSEMBLE COMME UN GRAND PEUPLE UNI ET FRATERNEL ABORDONS L'ERE NOUVELLE DE LA POLITIQUE FRANCAISE

Texte de la bibliothèque électronique du français moderne

<Valéry Giscard d'Estaing - président de la république française - allocution à l'occasion de sa prise de fonction - palais de l'Elysée - 27 mai 1974>
<Texte déchargé sur le site www.vie-publique.fr - Mise en minuscule - correction et standardisation orthographiques - balisage et lemmatisation : Dominique Labbé (pacte-iep Grenoble) 2013>

Messieurs les présidents, mesdames, mesdemoiselles, messieurs, de ce jour, date une ère nouvelle de la politique française. Ceci n'est pas seulement dû, monsieur le président du conseil constitutionnel, à la proclamation du résultat que vous venez de rappeler et dont, par respect pour la France et pour sa longue histoire, je mesure l'honneur. Ceci n'est pas seulement dû aux 13396203 femmes et hommes qui m'ont fait la confiance de me désigner pour devenir le vingtième président de la République française. Ceci est dû en réalité à la totalité des suffrages du 19 mai 1974. Ces suffrages égaux selon la règle démocratique qu'il s'agisse de ceux des femmes et des hommes, des jeunes et des moins jeunes, des travailleurs et des inactifs, et qui se sont prononcés chacun à leur manière et selon leur préférence en témoignant leur volonté de changement. J'adresse le premier salut du nouveau président de la république à ceux qui dans cette compétition aspiraient à le devenir et qui avaient la capacité de le faire et notamment monsieur François Mitterrand et monsieur Jacques Chaban-Delmas. Ainsi c'est moi qui conduirai le changement, mais je ne le conduirai pas seul. Si j'entends assumer pleinement la tâche de président, et si j'accepte, à cet égard, les responsabilités qu'une telle attitude implique, l'action à entreprendre associera le gouvernement dans ses initiatives et le Parlement dans son contrôle et dans ses droits. Je ne le conduirai pas seul parce que j'écoute et que j'entends encore l'immense rumeur du peuple Français qui nous a demandé le changement. Nous ferons ce changement avec lui, pour lui, tel qu'il est dans son nombre et dans sa diversité, et nous le conduirons en particulier avec sa jeunesse qui porte comme des torches la gaieté et l'avenir. Messieurs les présidents, mesdames, mesdemoiselles, messieurs, voici que s'ouvre le livre du temps avec le vertige de ses pages blanches. Ensemble comme un grand peuple uni et fraternel abordons l'ère nouvelle de la politique française.