# Trois débats et une élection. Débats à l'occasion de l'élection présidentielle américaine de 2008 [Obama-McCain]

Erin MacMurray

Sorbonne nouvelle – SYLED – 19 rue des Bernardines – 75005 Paris – France

TEMIS R&D – 164 rue de Rivoli – 75001 Paris – France

## Résumé

Les affrontements télévisés entre les candidats aux élections américaines constituent un moment important dans la course à la présidence. En 2008, le monde entier avait les yeux tournés vers Barack Obama et John McCain, deux candidats opposés par leur profil, leur âge et leur programme politique. Cette étude tente de mettre en évidence les différences perceptibles dans le discours de chaque candidat. Les méthodes textométriques, appliquées à un corpus aligné des trois débats de septembre à octobre, soulignent les spécificités du vocabulaire de l'un et de l'autre.

## Abstract

Televised debates during American presidential elections can be a crucial turning point. In 2008, the whole world was focused on Barack Obama and John McCain, two candidates different by their background, age and political agenda. This study attempts to shed light on a number of differences observed in the vocabulary chosen by either candidate. Textometric methods, applied to an aligned corpus of the debates from September to October, help bring out the specific vocabulary used by both candidates.

**Keywords:** textometry, vocabulary differences, repeated segments, parallel corpora, political discourse

## 1. Contexte de la recherche

Les *face-à-face* des candidats à la présidence américaine constituent un moment crucial dans la campagne électorale. Ces débats sont suivis aux Etats-Unis et certainement au-delà des frontières du pays. La diffusion et la traduction des affrontements entre candidats sont assurées dans le monde entier [1].

Lors des élections 2008, Barack Obama et John McCain, deux candidats que tout oppose du fait de leurs origines, leur âge, leur programme politique, s'affrontent dans trois débats de fin Septembre

[1] Aujourd'hui, presque 50 ans après les premiers débats télévisés entre John F. Kennedy et Richard Nixon, la diffusion télévisée est bien ancrée dans la course aux élections présidentielles. L'importance de ces débats est telle qu'il a fallu créer une commission pour encadrer les questions posées et le temps de parole. Depuis 1987 c'est la *Commission on Presidential Debates* qui est en charge cet encadrement dans le but de fournir « les informations les plus pertinentes possibles aux téléspectateurs ainsi qu'à ceux qui écoutent ». Cette commission a également la responsabilité de transcrire les débats afin d'en garder une trace écrite officielle. Ces transcriptions nous permettent de constituer une base sur laquelle il est possible de faire des analyses textométriques.

à fin Octobre. Ces oppositions sont-elles perceptibles au niveau de leur discours ? De quelle manière chaque candidat a-t-il exprimé ses positions lors des trois débats ? C'est à ces questions que nous tentons d'apporter quelques éclaircissements en utilisant l'approche textométrique.

### *1.1. La campagne électorale*

Devant la difficulté d'assumer le bilan des huit ans de présidence de George W. Bush, les deux candidats, cherchent à s'en distinguer par deux stratégies différentes. Afin de donner un nouveau visage au parti républicain, McCain tente de s'éloigner de la position de Bush sur les questions sociales tout en préservant son électorat ultra conservateur. Il doit réussir simultanément une nette rupture avec la politique de Bush pour convaincre les indécis et remporter les élections. Obama joue une carte différente, soulignant la responsabilité du parti républicain dans l'ensemble de la politique des huit dernières années. Il tente manifestement de profiter de sa jeunesse et de son origine « ethnique », face à un McCain vieillissant et dont le parti reste associé aux échecs des républicains.

### *1.2. Profil des candidats* [2]

| | *Obama* | *McCain* |
|---|---|---|
| **Age** | 48 | 72 |
| **Race** | white and black | white |
| **Religion** | Christian<br>United Church of Christ | Christian<br>Baptist |
| **Education** | Bachelor's degree from Columbia University international relations<br>Juris Doctor from Harvard Law School | Degree from U.S. Naval Academy and the National War College |
| **Military Service** | – | U.S. Navy 1958-81, Vietnam (prisoner of war) |
| **Family** | spouse Michelle Obama, 2 children | spouse Cindy McCain, 7 children |
| **Political Career** | 1996-2004 Illinois Senate<br>2005-2008 Senator of Illinois | House of Representatives1982-1986<br>Senator of Arizona 1986-présent |

## 2. Présentation du corpus

La transcription par la *Commission on Presidential Debates* (CPD) des trois débats constitue notre corpus d'étude. Ces débats ont eu lieu les 26 septembre, 7 octobre et 15 octobre : deux candidas s'affrontent à tour de rôle sur une question posée [3]. Afin d'analyser ces débats dans

[2] Nous avons utilisé pour ce tableau une terminologie à laquelle sont habitués les électeurs américains. Lors de leurs enquêtes sur la population américaine, le *United States Census Bureau* utilise le terme *race* afin de déterminer l'identification de l'individu questionné. Quatre catégories sont proposées : *white*, *black*, *American Indian* ou *Alaska Native*, *Asian*, et *Pacific Islander*. Depuis 2000, les recensements officiels incluent la notion d'*ethinicity* ; http://www.census.gov/.

[3] Chaque participant bénéficie d'un temps de parole à temps égal, mais le nombre d'interventions peut différer. Les trois débats sont organisés autour de 8 segments de 10 minutes. Chaque segment introduit un nouveau sujet que chaque candidat traite à son tour. Le modérateur pose ensuite des questions subsidiaires et tente de faciliter l'échange direct entre les candidats. Au cours de l'échange direct les candidats peuvent s'interrompre, créant une situation où le tour de parole de l'un n'est pas nécessairement suivi par l'autre.

leur ensemble et de pouvoir analyser les particularités du discours des deux participants, nous avons jugé nécessaire de procéder à un alignement du corpus. Cet alignement doit assurer un nombre égal de tours de parole pour chaque candidat. Chaque intervention de l'un, est suivie d'une réponse de l'autre, éventuellement réduite à un segment vide, exemple figure ci-dessous. Les interventions des modérateurs ont été enlevées afin de ne garder que les propos des deux candidats. La casse a été forcée en minuscule. Le corpus est ainsi *aligné* et a été soumis aux logiciels spécialisés dans l'interprétation d'alignement.

Le corpus est structuré en sous parties par un ensemble de balises, comme nous pouvons le voir dans Tab. 1. Le découpage par débat est indiqué par la balise <text=debate1>. Les débats sont numérotés par ordre chronologique du premier. Les tours de parole des deux candidats sont balisés, exemple : <speaker=obama>. Nous avons également indiqué un segment par intervenant et par débat, ce qui donne : <spdb=obama1> <spdb=mccain1>. Enfin un découpage est également repérable en fonction de la question posée <question=101> et le tour de parole, indiqué dans le corpus par la balise <part=> qui contient le numéro de la question plus une lettre a-z marquant l'ordre des tours de parole.

| *text* | *debate1* | *debate2* | *debate3* |
|---|---|---|---|
| **speaker** | mccain ou obama | mccain ou obama | mccain ou obama |
| **spdb** | mccain1 ou obama1 | mccain2 ou obama2 | mccain3 ou obama3 |
| **question** | 1 + numéro de la question 01 à n (101-110) | 2 + numéro de la question 01 à n (201-210) | 3 + numéro de la question 01 à n (301-310) |
| **part** | numéro de la question + lettre du tour de parole (102a-102j) | numéro de la question + lettre du tour de parole (202a-202j) | numéro de la question + lettre du tour de parole (302a-302j) |

*Tableau 1 : Récapitulatif des noms de balises et leur contenu*

<text=debate1>

…

<speaker=obama> <spdb=obama1> <question=106> <part=106c>

§ jim, let me just make a point. I've got a bracelet, too, from sergeant – from the mother of sergeant ryan david jopeck (ph), given to me ingrain bin green bay. She asked me, can you please make sure another mother is not going through what I'm going through …

<speaker=mccain> <spdb=mccain1> <question=106> <part=106c>

§ you might think that with that kind of concern that senator obama would have gone to Afghanistan, particularly given his responsibilities as a subcommittee chairman. By the way, when I'm subcommittee chairman, we take up the issues under my subcommittee…

<speaker=obama> <spdb=obama1> <question=107> <part=107a>

§ a testimony to you, jim.

<speaker=mccain> <spdb=mccain1> <question=107> <part=107a>

§ 0000

*Figure 1 : Extrait du corpus débats Obama-McCain*

Ces balises permettent donc de suivre le débat selon plusieurs découpages : 1) les tour de parole, 2) l'ordre des débats, 3) les questions posées au cours de chaque débat, 4) les candidats.

Le corpus comporte au total 44.190 occurrences de 3.654 formes différentes. On recense 1.585 hapax, la forme la plus fréquente *the* compte 1.814 occurrences.

## 3. Le style des trois débats

### *3.1. Principales caractéristiques quantitatives des trois débats*

Notre première étude a cherché à mettre en évidence les différences entre les trois débats [4]. Le premier débat comporte un nombre supérieur d'occurrences. Les différences des décomptes de formes et d'hapax suit également cette tendance. Comme nous l'avons déjà mentionné la forme la plus fréquente du corpus est *the* ; c'est le cas pour le premier et pour le troisième débat. Quant au deuxième débat, c'est la forme *to* qui est la plus fréquente. Ce changement est sans doute à mettre au compte de la forme du débat, style *town hall meeting*, qui a sollicité plus d'interaction entre les candidats mais également avec le public. Nous regarderons de plus près cette différence au cours de l'analyse.

| | *Nb occurrences* | *Nb formes* | *Nb hapax* | *Fréquence Max* | *Forme Max* |
|---|---|---|---|---|---|
| **Debate 1** | 15377 | 2149 | 1077 | 668 | the |
| **Debate 2** | 14238 | 1945 | 917 | 557 | to |
| **Debate 3** | 14575 | 2019 | 995 | 587 | the |

*Tableau 2 : Principales caractéristiques des débats*

L'AFC est le résultat d'un ensemble de calculs, permettant une visualisation de la proximité des différentes parties sélectionnées en fonction des formes dont la fréquence est supérieure ou égal à 10. Dans Fig. 2, nous observons que les productions correspondant à chacun des deux candidats sont bien distinguées sur le plan du vocabulaire employé. Par ailleurs, nous constatons une certaine homogénéité pour chaque débat dans la mesure où ils apparaissent ici sur le même axe.

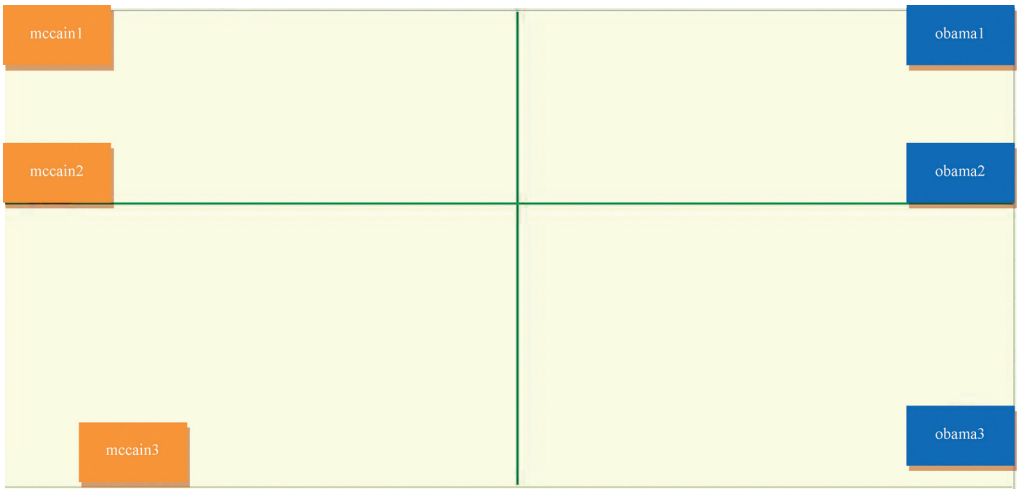


*Figure 2 : Analyse factorielle des correspondances sur la partie <spdb=>*

### *3.2. Principales caractéristiques quantitatives pour chacun des candidats*

D'abord, regardons le nombre d'occurrences par candidat et par débat, selon la partition <spdb>. Obama est largement en tête avec 7.948 occurrences pour le premier débat (spdb=obama1), alors que McCain compte 7.430 occurrences pour celui-ci (spdb=mccain1), au dessus du troisième débat où il compte 7.016 occurrences (spdb=mccain3). Quant à Obama, le troisième débat voit

[4] Le logiciel *Lexico 3* a été utilisé pour effectuer ces analyses.

moins d'occurrences, un total de 7.558 (spdb=obama3). Enfin, c'est le deuxième débat qui voit le moins d'occurrences pour les deux candidats, McCain, à 6.772 (spdb=mccain2) correspondant à une baisse de 8% par rapport à la moyenne et Obama (spdb=obama2) compte 7.452, largement en dessous du nombre d'occurrences lors du premier débat, soit une baisse de 6%. Cette diminution chez McCain est très apparente dans Fig. 3 plus bas. Cette constatation est aussi valable pour les formes et les hapax. La nature spécifique du deuxième débat pour les deux candidats apparaît clairement dans la figure ci-dessous montrant sa décroissance par rapport au premier débat. Bien que le temps de parole des deux candidats soit identique le discours de McCain est moins dense que celui de son adversaire.

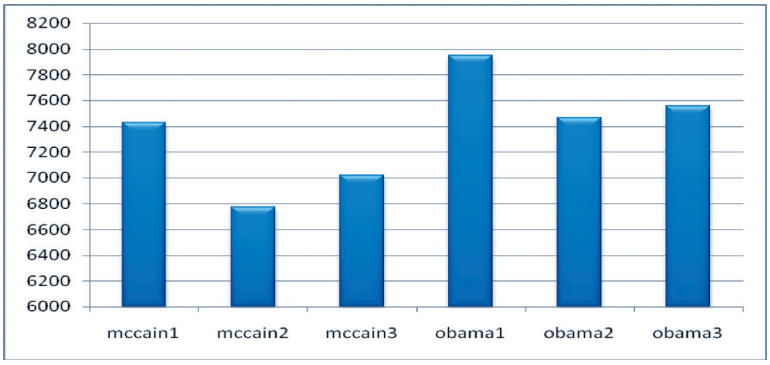


*Figure 3 : Le nombre d'occurrences par candidat et par débat*

Le deuxième débat se distingue également par la forme la plus fréquente, *and* chez McCain et *to* chez Obama. Le retour au contexte montre que ce phénomène est dû à la nature particulière de ce débat qui voit une forte fréquence de formes verbales.

Pour chaque débat, Obama emploie plus d'occurrences que McCain. Cependant, Fig. 4, l'accroissement du vocabulaire, montre que McCain a tendance à utiliser une plus grande variété de formes qu'Obama. Au cours du débat, McCain renouvelle plus son vocabulaire. A la lumière de ce résultat il apparaît que le vocabulaire d'Obama est plus pauvre que celui de McCain.

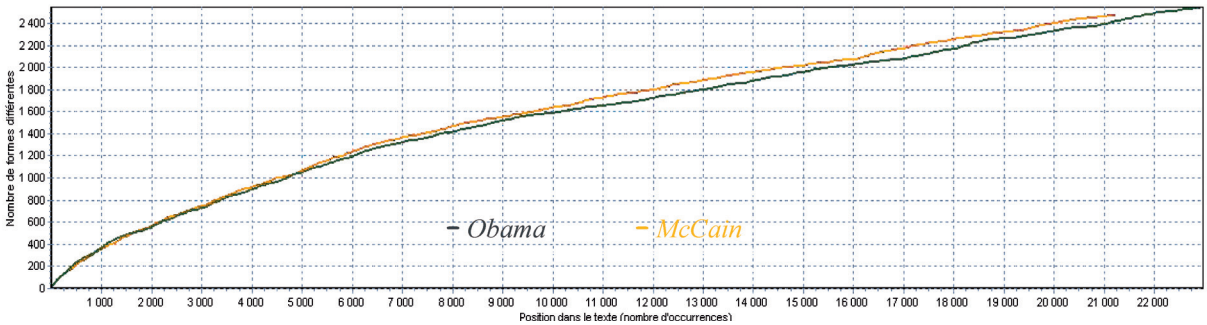


*Figure 4 : Accroissement de vocabulaire pour chacun des candidats*

## 4. Spécificités des candidats

Notre deuxième étude a cherché à relever les particularités de chacun des candidats par le calcul des spécificités. Le tableau ci-dessous présente le résultat de ce calcul [5]. Les termes du Tab 3

[5] Voir le modèle hypergéométrique appliqué aux formes du corpus, tel qu'il est présenté par Lafon (1980).

sont les spécificités majeures du corpus. Les formes et segments répétés montrés ici sont fournis par ordre décroissant de spécificité du plus spécifique au moins spécifique.

| *Spécificités 5+de Obama* | | | | | | *Spécificités 5+de McCain* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Forme* | *FTot* | *FP* | *Forme* | *FTot* | *FP* | *Forme* | *FTot* | *FP* | *Forme* | *FTot* | *FP* |
| senator mccain | 117 | 115 | agree | 15 | 15 | senator obama | 132 | 132 | m not | 13 | 13 |
| mccain | 121 | 117 | eight years | 26 | 23 | obama | 135 | 134 | again | 48 | 37 |
| make | 155 | 117 | here | 59 | 44 | i know | 42 | 40 | united states | 69 | 48 |
| sure that | 63 | 55 | senator mccain and i | 13 | 13 | and | 1568 | 878 | taxes | 45 | 35 |
| is | 557 | 356 | end up | 13 | 13 | america | 98 | 74 | it be | 13 | 13 |
| i think | 153 | 111 | for example | 13 | 13 | my friends | 22 | 22 | to tell | 11 | 11 |
| going | 264 | 178 | talk about | 13 | 13 | the point is | 19 | 19 | record | 37 | 29 |
| john | 37 | 34 | to deal with | 24 | 22 | i have | 67 | 51 | americans | 46 | 35 |
| making | 30 | 29 | what i | 41 | 34 | friends | 31 | 28 | but the | 21 | 19 |
| going to | 233 | 161 | true | 28 | 24 | my | 136 | 92 | and i | 183 | 113 |
| to make | 93 | 73 | getting | 26 | 23 | that senator obama | 19 | 19 | against | 48 | 36 |
| think | 178 | 128 | important for | 14 | 14 | know | 162 | 109 | barrel | 13 | 13 |
| important | 69 | 56 | the last | 46 | 36 | will | 162 | 107 | by the way | 18 | 17 |
| we are | 71 | 56 | the last eight years | 21 | 19 | by the | 27 | 25 | know that | 20 | 18 |
| make sure | 52 | 42 | policies | 28 | 25 | ll | 64 | 49 | know how | 21 | 19 |
| just | 130 | 92 | example | 31 | 26 | i know how to | 15 | 15 | business | 33 | 27 |
| sure | 83 | 63 | are going to | 28 | 25 | point is | 20 | 19 | whether it be | 12 | 12 |
| last | 63 | 49 | deal | 35 | 29 | we need to | 41 | 34 | the united states senate | 14 | 14 |
| are going | 36 | 31 | provide | 31 | 27 | I know how | 17 | 17 | i | 970 | 519 |
| going to be | 58 | 46 | eight | 28 | 24 | fought | 19 | 18 | of | 1005 | 540 |
| that we | 203 | 135 | got to make | 14 | 14 | in america | 37 | 31 | strategy | 31 | 25 |
| over the last | 24 | 23 | actually | 19 | 18 | a long | 19 | 18 | the united | 71 | 49 |
| of you | 16 | 16 | point i | 13 | 13 | the point | 24 | 22 | pork | 11 | 11 |
| making sure that | 45 | 38 | s important | 15 | 15 | sit down | 15 | 15 | senator obama has | 14 | 14 |
| absolutely | 21 | 20 | some | 118 | 81 | control | 16 | 15 | others | 11 | 11 |
| senator mccain and | 17 | 17 | us to | 20 | 19 | united | 75 | 53 | of course | 19 | 17 |
| we are going to | 15 | 15 | right | 86 | 61 | need to | 50 | 37 | | | |
| we are going | 17 | 16 | over the last eight years | 14 | 14 | | | | | | |
| that | 1375 | 786 | over the | 37 | 31 | | | | | | |
| we | 1113 | 639 | invest | 14 | 14 | | | | | | |
| for us to | 15 | 15 | ve got to make | 13 | 13 | | | | | | |
| are | 318 | 201 | because | 148 | 98 | | | | | | |
| making sure that | 15 | 15 | over | 67 | 50 | | | | | | |
| this is | 56 | 43 | | | | | | | | | |

*Tableau 3 : Les spécificités coef 5 et plus chez Obama et McCain*

Ce calcul permet de mesurer la probabilité pour qu'une forme d'une fréquence *f* apparaisse *k* fois dans une partie *i* sélectionnée du corpus. Nous avons choisi de nous concentrer uniquement sur les formes/segments répétés ayant un coefficient de 5 ou plus. Les fréquences FP données

ici correspondent au nombre d'occurrences de la forme pour la partie <speaker=> de chaque candidat. Le coefficient est calculé à partir de la fréquence totale (FTot ci-dessous) de la forme dans le corpus.

D'abord, nous remarquons l'importance du nombre de formes coef 5 et plus du côté du candidat démocrate. Il n'est pas étonnant de voir que les segments les plus spécifiques de chaque candidat sont les noms de leur adversaire (*senator mccain*, *senator obama*). Ainsi, Obama s'adresse fréquemment à McCain et inversement. De manière plus globale, nous observons la tactique d'Obama qui consiste à rappeler constamment la présidence de Bush au cours de son discours avec les segments tels *over the last eight years*, *over the last*, *the last eight years*... De son côté, le candidat républicain rappelle son service et son expérience dans le gouvernement américain par les formes de type : *america*, *united states*, *the united states senate, record*... Sa tendance à vouloir se rapprocher de l' « américain moyen » apparaît clairement ici avec le segment *my friends,* qu'il est le seul à utiliser. Les deux candidats s'opposent également sur certains sujets observés par les formes *invest* du côté démocrate et *pork barrel* [6] (deux formes séparées dans le tableau) du côté républicain.

Nous avons choisi d'élaborer certaines observations du Tab. 3 dans les parties qui suivent.

### *4.1. Les verbes : un débat marqué par l'action*

Chez les deux candidats la moitié des segments répétés sont des expressions verbales. Chaque participant utilise des expressions verbales différentes. Obama emploie souvent les formes du futur avec l'expression *going to*, en revanche, chez McCain, nous observons plutôt la forme *will* ou xx'*ll* (contraction de *will*).

De part et d'autre, les formes désignant le futur sont utilisées de manière plus fréquente au cours du deuxième débat. McCain continue à employer *will* dans le troisième débat, alors que la forme *going to* est moins utilisée par Obama.

Au-delà de l'expression *going to* qui dénote le futur en anglais, les spécificités d'Obama montrent également la présence de formes verbales au présent progressif -*ing*. D'ailleurs, les deux expressions *make sure that* et *making sure that* sont constamment martelées. Le présent progressif -*ing* et les formes en *going to* semblent être sur représentées dans le deuxième débat.

Les expressions *I think* chez Obama et *i know* chez McCain sont parmi les segments les plus spécifiques à chaque candidat avec des spécificités respectives de +8 et +11 respectivement. En étendant la requête aux verbes *think* vs. *know,* nous observons, de manière générale, que *think* est le plus souvent employé au cours du troisième débat pour Obama. Chez McCain, c'est dans le premier et de manière très fréquente dans le deuxième débat que *know* est utilisé. McCain tente de se distinguer par rapport à son expérience politique en employant la forme *know* (*I know that*). Comme observé dans Tab. 3 des spécificités McCain évoque souvent sa carrière (forme *record*), présente de manière significative lors du deuxième débat. McCain abandonne ces formes dans le dernier débat face à un changement de tactique d'Obama par rapport aux deux débats précédents, ce que nous verrons plus loin.

[6] *Pork barrel* désigne les dépenses gouvernementales pour des projets très localisés dans le but de rapporter de l'argent à la région d'un représentant du gouvernement.

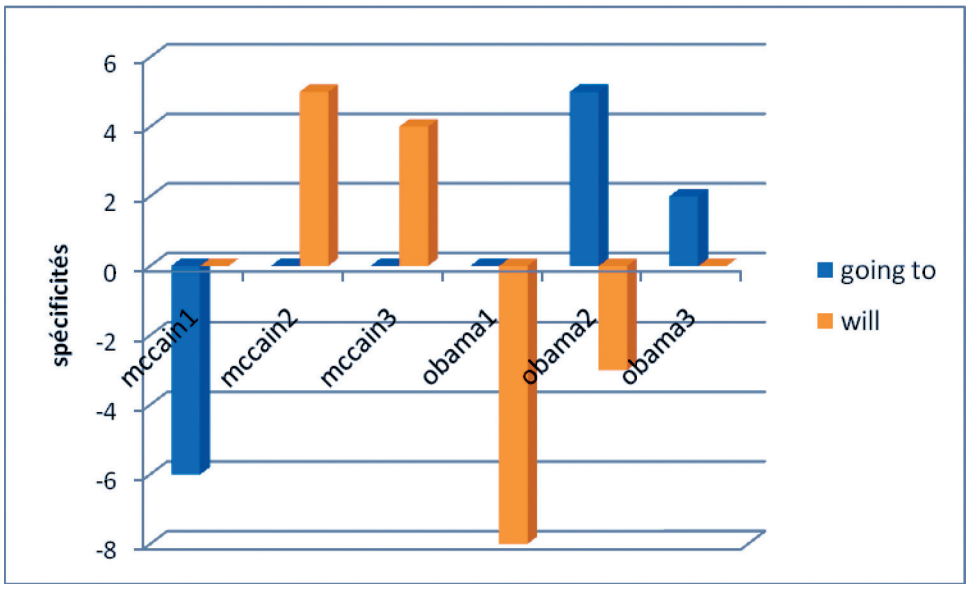


*Figure 5 : Spécificités de will et going dans le discours de McCain et Obama*

### *4.2. Les adresses : l'irrégularité de leur représentation*

Le tableau 3 sur les spécificités montre clairement que chaque candidat désigne son adversaire de manière fréquente. En effet, Obama nomme le *senator mccain* directement avec 115 occurrences de ce segment. McCain, quant à lui, désigne le *senator obama* 132 fois. Chaque candidat tente de se distinguer de son adversaire en l'évoquant dans son discours.

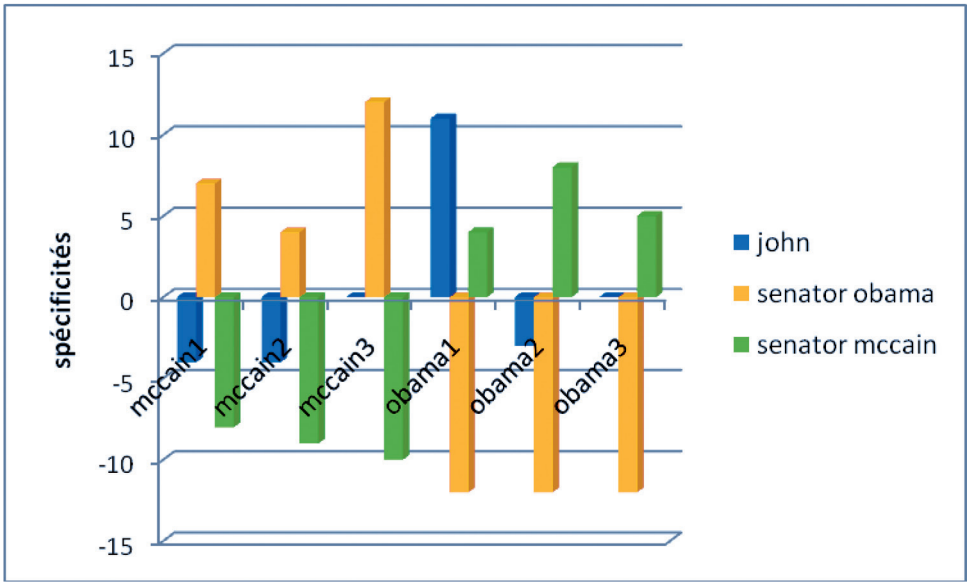


*Figure 6 : Spécificités de john, senator mccain, et senator obama dans le discours d'Obama et McCain*

Il est intéressant de remarquer qu'Obama emploie également le prénom de McCain, *john*, dans des contextes très similaires à ceux où il choisit le segment *senator mccain*. En revanche McCain n'emploie jamais le prénom d'Obama et utilise toujours le segment *senator obama*. Comparons, dans les exemples qui suivent, le contexte de *senator mccain* et *john* dans le discours du candidat démocrate.

(i) f the most conservative republicans who **john** already mentioned to set up what
ve to change the culture . tom - - or **john** mentioned me being wildly libera
ime ' s worth of tax increase . § now, **john** mentioned the fact that business
ng , because , you know , 10 days ago, **john** said that the fundamentals of th
(ii) agree on the importance of energy, but **senator mccain** mentioned earlier the
of us engaging in direct diplomacy . **senator mccain** mentioned henry kissin
powerful because of petro - dollars, as **senator mccain** mentioned . § that mea
(iii) let ' s go back to the original point. **john** , nobody is denying that $ 18 bi
ey are still sending out videotapes and **senator mccain** , nobody is talking ab
(iv) gment of this war . § and so john likes **john** , you like to pretend like the
against it ; senator obama voted for it. **john** , you want to give oil companies

Dans les exemples (i) et (ii) Obama ne s'adresse pas directement à McCain, au contraire, il se réfère à des discours passés de son adversaire avec les verbes *mentioned*, *said*. Dans ce cas l'usage de *john* ou de *senator mccain* ne semble pas changer le sens de son discours. L'exemple (iii) montre où Obama semble s'adresser, cette fois, personnellement à McCain. Les segments *john* ou *senator mccain* ponctue le discours, marqué par une pause transcrite avec une virgule. Ainsi, il distingue le discours de McCain de celui « partagé » par Obama et le public, ce qui peut être observé par l'usage de la forme *nobody* suivant la virgule. Les mêmes phrases ne changeraient pas de sens si McCain n'était pas impliqué directement. L'effet produit aurait été certainement différent si Obama avait parlé en général. Enfin, il est possible d'observer dans l'exemple (iv), des contextes dans lesquels le segment *senator mccain* n'apparaît jamais. Lorsqu'Obama s'adresse clairement à McCain, avec la forme *you*, il emploie son prénom.

### *4.3. Les pronoms : changement de stratégie dans le dernier débat*

Dans les spécificités de chaque candidat (Tab. 3), nous remarquons de nombreux segments qui contiennent pronom *we* pour Obama et *I* chez McCain. Cette recherche a été étendue au groupe de pronoms de la première personne du pluriel *we*, *us*, *our, ours* (we + dans la figure ci-dessous), et au groupe de pronoms de la première personne du singulier *I*, *me*, *my*, *mine* (I+ dans la figure ci-dessous). La spécificité des pronoms est très apparente lors du premier débat et également lors du deuxième débat avec une fréquence moindre. Seul le troisième débat marque une différence avec les deux autres pour Obama. Il semble moins utiliser les pronoms de la première personne du pluriel, indiqué ici par la spécificité négative de we+. Il s'agit sans doute d'une rupture avec la stratégie utilisée précédemment.

En résumé, nous observons au fil des débats une décroissance des pronoms personnels spécifiques à chaque candidat (observés également dans Tab. 3, première personne du pluriel chez Obama et première personne du singulier chez McCain). Le troisième débat semble marquer un changement de stratégie dans la façon dont les deux candidats s'adressent au public.

## 5. Les segments répétés par les candidats

### *5.1. La répétition : spécificité d'Obama*

Nous avons vu que le discours d'Obama avait un plus grand nombre de formes (Fig. 3), mais d'après l'accroissement de vocabulaire (Fig. 4) il y a moins de formes par nombre d'occurrences que dans le discours du candidat républicain. Serait-ce parce que le candidat démocrate se répète souvent ? Il est possible de regarder du côté des segments répètes afin d'approfondir cette question.

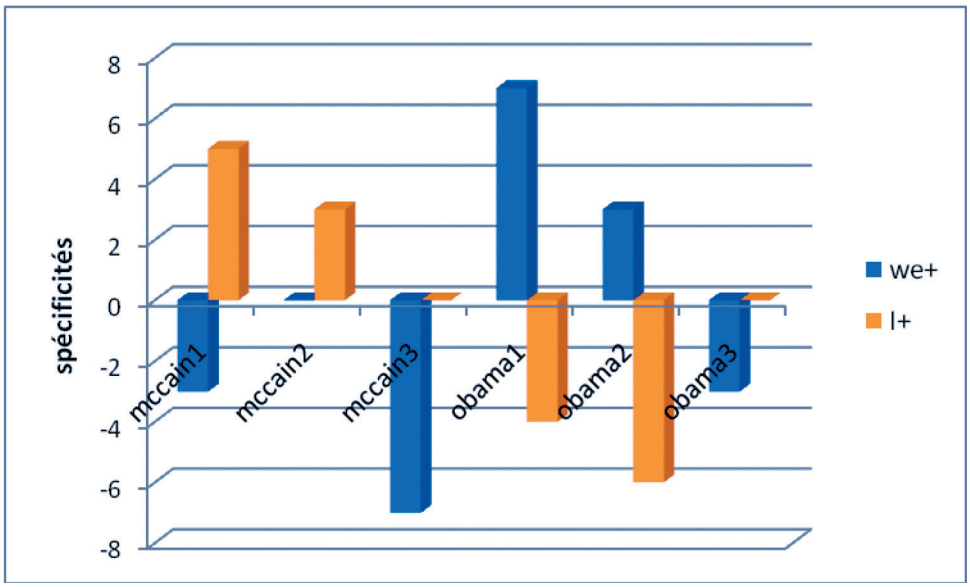


*Figure 7 : Spécificités de we et de I dans les discours d'Obama et de McCain*

Les segments répétés sont des séquences comprenant 2 formes ou plus qui ont deux occurrences ou plus dans le corpus. Nous avons choisi d'observer ici les segments répétés en fonction de leur taille. D'abord les segments de 7 formes ou plus, ensuite les segments de 5 à 6 formes, 3 à 4 formes et enfin les segments constitués de seulement deux formes. En analysant la ventilation de ces tranches sur les parties <spdb=> dans le corpus, il est frappant de noter que les segments répétés de toute longueur sont une spécificité d'Obama et par conséquent sont sous-représentés chez McCain.

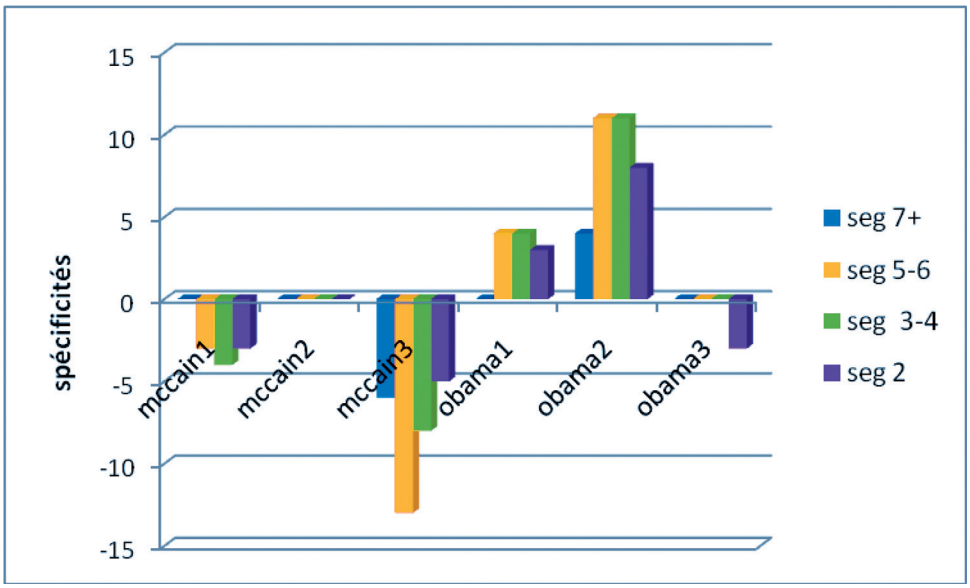


*Figure 8 : Spécificités des différents segments dans le discours d'Obama et de McCain*

### *5.2. Une répétition partagée*

Le nombre de segments répétés communs de part et d'autre dépend fortement de leur taille. Il est intéressant de remarquer que dans la tranche des segments de 7 formes ou plus, seulement deux sont employés par les deux candidats, *we still have a long way to go* (v) et *a lot of work*

*to do in* (vi). Cette observation montre que les segments de cette taille sont très spécifiques au candidat qui les a produits, c'est-à-dire ce qui est dit par l'un n'est en général pas répété par l'autre.

(v) *Partie : mccain1, Nombre de contextes : 1*
ncies are doing a great job. **but we still have a long way to go** before we
*Partie : obama1, Nombre de contextes : 1*
ms of securing potential targets, **but we still have a long way to go** . § we ' ve got

(vi) *Partie : mccain1, Nombre de contextes : 1*
jihadist elements. so we ' ve got **a lot of work to do in** afghanistan . but I ' m con
*Partie : obama1, Nombre de contextes : 1*
we , i think , are going to have **a lot of work to do in** the next administration to

Lorsque la taille des segments diminue, les segments communs aux deux candidats deviennent de plus en plus fréquents. Nous avons opéré une distinction entre les segments qui sont partagés, c'est-à-dire communs aux deux candidats et ceux qui spécifiques à chaque candidat, ne sont jamais employés par l'autre. Globalement, il semblerait que presque la moitié des segments répétés sont employés par les deux candidats avec au moins une occurrence d'un côté ou de l'autre (Fig. 9a).

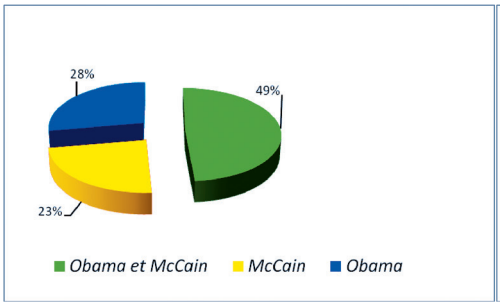


*Figure 9a :*
*Répartition des segments répétés*

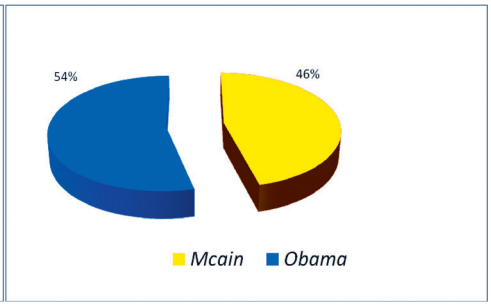


*Figure 9b :*
*Répétitions totales pour chaque candidat*

La tendance à la répétition reste néanmoins plus forte chez Obama, observée dans Fig. 9b qui compte la fréquence totale des segments répétés pour chaque candidat (17.201 pour Obama et 14.696 pour McCain). Il est alors possible de remarquer des écarts de type une occurrence d'un segment (vii) du côté républicain et, pour le même segment, 9 occurrences (viii) du côté démocrate. Cette remarque est également conforme aux résultats de l'analyse de la spécificité des segments répétés (section 5.1).

(vii) *Partie : mccain, Nombre de contextes : 1*
ke sure we ' re not handing the **health care system** over to the federal

(viii) *Partie : obama, Nombre de contextes : 9*
americans to go down , the - a **health care system** that is broken , ener
es . i make sure that we have a **health care system** that allows for every
th korea . § we have to fix our **health care system** , which is putting an
- term we ' ve got to fix our **health care system** , we ' ve got to fix
umber two , because that broken **health care system** is bad not only for
aveling of the employer - based **health care system** . § that , i don ' t
th away. we ' ve got to fix or **health care system** and we ' ve got to
tter . it ' s not improving our **health care system** . it ' s just a giveaw
raveling of the employer -based **health care system** . § all i want to do ,

## 6. Conclusion

L'analyse des spécificités et des segments répétés de chaque candidat a révélé certaines particularités langagières de chacun des candidats. Les constructions répétées observées dans la partie 5 nous amènent à des questions sur la façon dont chaque candidat se prépare au débat. Il semblerait qu'Obama construise son discours de manière plus pondérée que son rival. En effet, Obama utilise souvent de segments de 7 formes ou plus (Fig. 8). Cette façon d'employer des « blocs » de mots probablement mémorisés pourrait aider à économiser un temps précieux et nécessaire à la formulation du contenu de ses réponses. Le candidat démocrate a tendance à employer la première personne du pluriel *we* alors que son adversaire tend à utiliser des constructions en *I*, première personne du singulier (Fig. 7). Ainsi, nous voyons une différence dans l'implication de la proximité avec la population américaine. Lorsqu'Obama choisit les constructions en *I*, il s'agit le plus souvent d'expressions telles *i think*.

McCain essaie de créer l'image d'un homme politique confiant. Il se repose sur son expérience avec les expressions telles *i know* ce qui montre un personnage sûr de ses positions. Dans un jeu du chat et de la souris, le candidat républicain tente de mettre de la distance entre lui-même et son rival. McCain emploie toujours le segment *senator obama* pour désigner son opposant. Le candidat démocrate réduit cette distance en utilisant la forme *john*, le prénom de son adversaire (Fig. 6). Contrairement à Obama, McCain emploie un nombre inférieur de répétitions. Il donne l'impression que son discours n'aurait pas été rédigé en amont et serait donc plus spontané. Ce nombre de répétitions moindre explique certainement la légère supériorité de nombre de formes par nombre d'occurrences (voir figure 4 accroissement de vocabulaire) de McCain, donnant également un sentiment de spontanéité dans son discours. Les segments tels *my friends* lors du deuxième débat sont peut-être des pauses marquant un temps de réflexion dans le discours du candidat républicain, tout comme les constructions répétées chez Obama.

Cette spontanéité du discours de McCain pourrait expliquer son usage de constructions en *will*, car à cette forme est souvent attribuée des caractéristiques sémantiques différentes de *going to*, forme souvent utilisée par Obama (Fig. 5). Il est clair que la distinction très nette entre les deux candidats sur l'usage du futur mériterait une analyse spécifique sur la distribution et la contextualisation des deux constructions.

Sur l'ensemble des débats, le candidat démocrate a été plus présent en termes de volume de mots. Toutefois, le deuxième débat se distingue de deux autres. Ce débat compte le moins d'occurrences pour McCain et Obama. Le deuxième débat semble donc être le point tournant. Cette observation pourrait correspondre au sentiment, de la part de la presse, qu'aucun des deux candidats s'est distingué lors du premier débat. Au cours du deuxième débat tant au nombre d'occurrences que de formes, Obama occupe clairement l'espace. D'un côté il emploie beaucoup de formes verbales, ce qui est à l'origine de la fréquence de la forme *to*. McCain, quant à lui, emploie *and* de manière fréquente, signe d'un discours probablement moins répété au préalable.

Les procédures de textométrie que nous avons utilisé ici ont aidé au déchiffrement de ces débats. Les outils textometriques (afc, spécificités, segments répétés, etc.) nous ont permis de suivre les variations au cours des débats tant sur le plan quantitative mais également sur le plan du vocabulaire utilisé. Ces méthodes ont révélé quelques particularités linguistiques dans le discours de chaque candidat qu'il nous faudra encore étudier, pour comprendre les ressorts qui ont conduit au succès démocrate.

## Références

Bonnafous S. and Tournier M. (1995). Analyse du discours, lexicométrie, communication et politique. *Langages*, vol. 29, 117 : 67-81.

Connolly K. (2008). *US rivals spar in first TV debate*, 27, septembre, BBC News http://news.bbc.co.uk/2/hi/americas/7639102.stm.

Labbé C. and Labbé D. (1997). Que mesure la spécificité du vocabulaire. In *Lexicometrica*, vol. 3.

Lafon P. (1980). Sur la variabilité de la fréquence des formes dans un corpus. *Mots*, 1 : 127-165.

Lebart L. and Salem A. (1994). *Statistique Textuelle*. Paris : Dunod.

Née E. (2008). Insécurité et élections présidentielles dans le journal Le Monde. In *Lexicométrica*, numéro thématique « Explorations Textuelles ».

Salem A. (1987). *Pratique des segments répétés. Essai de statistique textuelle*. Paris : Klincksieck.

**Corpus**

*Commission on Presidential Debates*: http://www.debates.org/.

*Commission on Presidential Debates Announces Sites, Dates, Formats, and Candidate Selection Criteria for 2008 General Election* http://www.gwu.edu/~action/2008/chrndeb08/cpd111907pr.html.