



LA STRUCTURE LINGUISTIQUE DE TWEETS EN CAMPAGNE PRÉSIDENTIELLE*

1 INTRODUCTION

Le discours politique est un sujet d'analyse proéminent depuis l'antiquité parce qu'il est constamment pris entre l'ambition de représenter la voix du peuple et celle de réaliser les intérêts particuliers des politiciens et de leurs partis. Même si ces ambitions sont souvent jugées par le public critique comme égocentriques, elles sont, dans la plupart des cas, tolérées par les citoyens. Ce phénomène peut s'expliquer par la notion « *illusion* », introduite par Pierre Bourdieu (1998 : 71) : il s'agit d'un consensus auquel croient et participent tous les acteurs du jeu politique (les politiciens, les organisations et les citoyens).

Même avant l'époque des médias électroniques, les politiciens et leurs équipes de communication étaient motivés d'interagir de manière directe et continue avec les citoyens. Le Bart (1998 : 79) constate que les acteurs politiques sont en permanence « soucieux d'assurer leurs positions en affirmant le bien fondé de leurs prétentions à faire de la politique ». Ellison et al. (2006 : 416) vont encore plus loin en démontrant que les réseaux sociaux permettent un contrôle d'autant plus grand des stratégies d'auto-promotion, et Dayter (2014 : 61) parle de discours de l'éloge de soi (ang. *self-praise*).

A l'époque des nouveaux médias, l'impact économique des interactions entre les politiciens et leur public encore plus important. C'est pourquoi de nombreuses disciplines étudient le discours politique afin d'expliquer ou même prédire les résultats électoraux (O'Connor et al. 2010; Smailović et al. 2015), de faire du profilage des électeurs (Prince et al. 2007, Suignard et al. 2021) et d'analyser les émotions dans les tweets (Hamon et al. 2015). Ces recherches, basées sur des fondements linguistiques, utilisent des approches computationnelles qui permettent d'analyser de grandes quantités de textes mais permettent peu d'interprétation au niveau linguistique. Dans cette étude, nous proposons une méthodologie d'analyse systématique du discours politique sur Twitter qui permet aux linguistes sans compétences en linguistique computationnelle d'obtenir des résultats quantifiables et en même temps interprétables par des catégories linguistiques traditionnelles.

* Cette étude a été effectuée dans le cadre du programme P6-0218 *Recherches théoriques et appliquées des langues : approches contrastives, synchroniques et diachroniques* et du projet J6-2581 *Analyse computationnelle du discours médiatique à l'aide des plongements lexicaux*, financés par L'Agence nationale de la recherche scientifique (ARRS).

2 ANALYSES DU DISCOURS POLITIQUE SUR LES RESEAUX SOCIAUX

Sur le plan théorique, l'analyse du discours politique sur twitter couvre trois domaines scientifiques : l'analyse du langage des réseaux sociaux, l'analyse discursive du discours politique et la fouille de données en informatique.

Le discours sur Twitter fait partie du discours électronique médié (Panckhurst 2006 : 345). Il est doté de certaines spécificités lexicales, graphiques et syntaxiques, comme les émoticônes, abréviations et répétitions de caractères (Farzindar et al. 2013 : 7) et connaît un certain nombre d'ajustements à cause de la limitation de la quantité de caractères dans un message (280 caractères maximum, espaces comprises). Ces limitations influencent, bien évidemment, le choix de moyens stratégiques de la communication des professionnels de la politique.

Les analyses qualitatives du discours politique (Sandré 2009, Delmas 2012) examinent l'utilisation de différentes structures aux niveaux syntaxique, sémantique, énonciatif et interactionnel, ou bien elles examinent des structures avec un effet particulier sur leur public, comme c'est le cas, par exemple, de la notion du pathos dans le discours politique (Charaudeau 2008). Les résultats des analyses qualitatives permettent une compréhension approfondie du discours politique et de ses effets sur ses audiences. Roginsky et al. (2015) démontrent que le discours politique comporte des structures représentant des annonces de choses accomplies ou, plus explicitement, d'évaluations positives de soi. L'inconvénient des résultats des méthodes qualitatives réside dans le fait d'avoir été obtenus à partir de quantités de textes assez limitées.

Les méthodes du traitement automatique du langage (Pang et Lee 2008; Smailović et al. 2014) reposent dans la plupart des cas sur un ensemble d'exemples annotés manuellement. Cet ensemble d'exemples sert de modèle pour les logiciels informatiques qui identifient les caractéristiques distinctives des différents sentiments. Dans l'étape suivante, les logiciels attribuent l'un des sentiments aux exemples non-catégorisés. Ces modèles sont très efficaces sur de grands ensembles de données. Ils ont, du point de vue linguistique, deux défauts principaux : premièrement, ils ne fonctionnent pas très bien sur des ensembles de textes compilés par des linguistes parce que ces ensembles de données sont considérés comme trop petits. Deuxièmement, comme les modèles de traitement automatique du langage sont basés sur des critères statistiques (par exemple sur les N-grams, Stamatos 2013 : 421) et non pas sur l'observation des critères linguistiques (par exemple sur les pronoms possessifs, Delmas 2012 : 111), ils ne permettent pas de performer des interprétations significatives sur la corrélation entre les structures linguistiques analysées et leur effet discursif.

Il ressort des recherches mentionnées qu'au niveau de l'observation du discours politique sur Twitter, les recherches linguistiques auraient pu bénéficier de méthodologies systématisées et quantifiables et que les approches computationnelles auraient pu être complétées par des interprétations linguistiques ou humanistes. Nous proposons dans la suite un exemple de méthodologie qui pourrait permettre aux linguistes sans formation en approches computationnelles de réaliser des études systématiques de discours politique sur Twitter ou sur d'autres réseaux sociaux.

3 MÉTHODOLOGIE

Nous présentons une analyse systématique d'une sélection de tweets qui ont été publiés en pleine campagne présidentielle 2022 et reflètent l'atmosphère politique du moment. L'objectif de l'analyse était d'examiner les caractéristiques linguistiques des messages des deux candidats les plus probables pour le second tour, Emmanuel Macron et Marine Le Pen. Les messages ont été systématiquement analysés aux niveaux textuel, syntaxique, énonciatif et thématique afin de voir s'il existe des différences linguistiques pertinentes entre les deux candidats.

3.1 Compilation du corpus

La première étape consistait à compiler le corpus. Nous avons pris pour notre analyse les tweets qui ont été publiés sur les deux profils officiels des candidats entre le 28 février et le 5 mars 2022. Le corpus comporte 40 messages que nous avons segmentés en 70 phrases séparées.



Image 1 : Photo des profils officiels des candidats présidentiels Emmanuel Macron et Marine Le Pen¹

1 Sources : <https://mobile.twitter.com/macronensemble>; https://twitter.com/MLP_officiel

3.2 Annotation linguistique

Nous avons analysé chaque message aux niveaux de la longueur textuelle, de la syntaxe, de l'énonciation et de la thématique comme suite au travail présenté dans Zwitter Vitez (2020 : 157). Au niveau textuel, nous avons examiné la longueur des tweets analysés (de 1 à 4 phrases). Au niveau syntaxique, nous avons d'abord déterminé la structure syntaxique de la phrase (simple, complexe). Ensuite, nous avons examiné le type de phrase selon la méthodologie de Riegel et al. (1994 : 101) : phrase déclarative, exclamative, interrogative et injonctive. Dans l'étape suivante, nous avons identifié la présence éventuelle d'éléments énonciatifs (Benveniste 1970) explicitant les traces de l'auteur du texte (pronoms personnels, verbes modaux, jugements de valeur, spécificités orthographiques etc.). Dans la dernière étape, nous avons annoté la thématique prédominante du tweet, par exemple l'économie, l'énergie ou le travail. Le tableau (1) présente les annotations linguistiques aux niveaux textuel, syntaxique, énonciatif et thématique de deux commentaires choisis.

Tableau 1 : Les annotations linguistiques de deux commentaires choisis

| | | |
|---------------------|---|--|
| Tweet | <i>Je sortirai du marché européen libéralisé de l'électricité pour retrouver le tarif français de l'électricité, inférieur au tarif européen.</i> | <i>Face aux crises, Emmanuel Macron construit notre souveraineté agricole #avecvous.</i> |
| Auteur | Marine Le Pen | Emmanuel Macron |
| Structure du texte | 1 phrase | 1 phrase |
| Structure de phrase | Complexe | Simple |
| Type de phrase | Déclarative | Déclarative |
| Énonciation | première personne singulier | première personne pluriel |
| Thématique | Énergie | Agriculture |

Une fois que tous les tweets ont été annotés, nous avons examiné les tweets des deux candidats à tous les niveaux d'analyse linguistique. L'objectif de cette analyse était de voir s'il existe, pour chacun des deux auteurs, des différences au niveau linguistique qui pourraient aider à faire une description systématique de leurs styles.

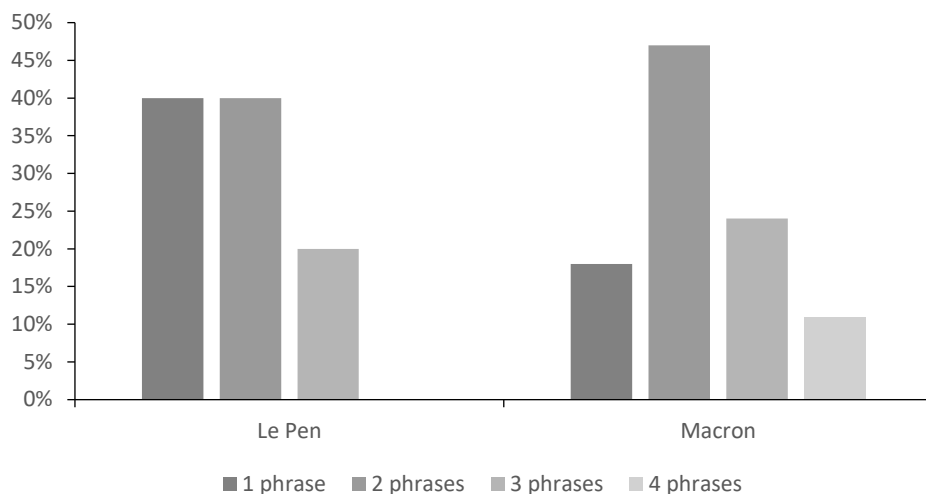
4 ANALYSE ET RÉSULTATS

Après avoir attribué à tous les commentaires une catégorie au niveau de la structure textuelle, de la structure syntaxique et des structures énonciatives, nous avons pu procéder à l'analyse des résultats.

4.1 Structure textuelle

Au niveau textuel, nous avons essayé de voir les spécificités concernant la structure et la longueur des tweets produits par Marine Le Pen et Emmanuel Macron. Le graphique (1) montre la distribution du nombre de phrases composant les tweets analysés.

Graphique 1 : La longueur des tweets produits par Marine Le Pen et Emmanuel Macron



Les résultats montrent que les tweets produits par Marine Le Pen sont le plus souvent composés d'une phrase (40 %, exemple 1) ou de deux phrases (40 %, exemple 2). Les tweets qui comportent 3 phrases ou plus sont rares (20 % des cas).

- (1) Je veux sortir du commandement intégré de l'#OTAN pour rendre à la France sa puissance et son rôle d'arbitre entre les grands blocs.
- (2) Un embargo sur le gaz et le pétrole russes aurait des conséquences cataclysmiques sur les Français. Je ne veux pas sacrifier les Français, leur pouvoir d'achat, leur emploi, l'économie française.

Les tweets produits par Emmanuel Macron, en revanche, sont composés majoritairement de deux phrases (47 %, exemple 3) ou de trois phrases (29 %, exemple 4), mais nous avons trouvé même une proportion non négligeable de tweets composés d'une phrase (18 %) et un exemple de tweet composé de quatre phrases.

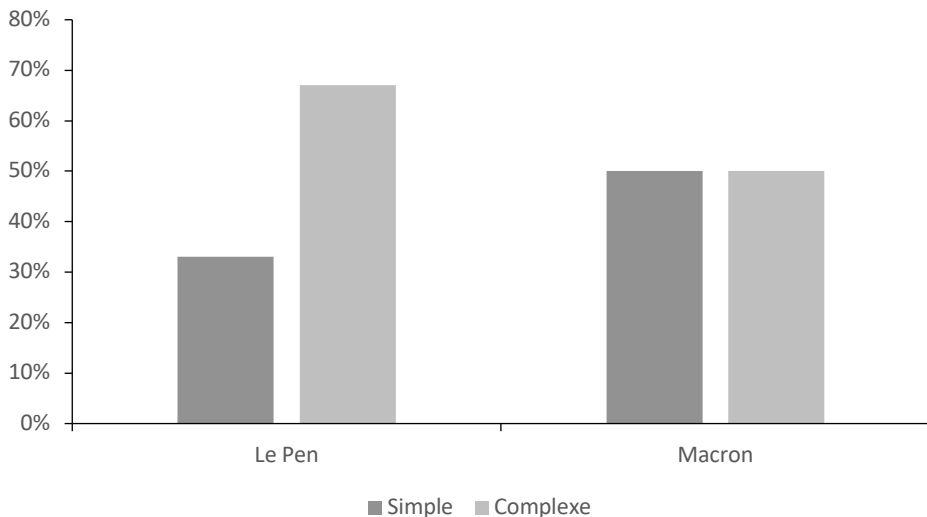
- (3) Depuis 2017, Emmanuel Macron agit sans relâche pour améliorer le quotidien de nos agriculteurs et garantir les conditions de notre souveraineté alimentaire. #avecvous pour notre agriculture et notre alimentation. #avecvous pour notre agriculture et notre alimentation.
- (4) Allez-y, dites-moi ce que vous avez sur le cœur. En 2018 dans le Puy-de-Dôme, EmmanuelMacron a rencontré Jean-Valère, fromager. 4 ans après, l'agriculteur revient sur les transformations qui améliorent son quotidien et renforcent notre souveraineté alimentaire.

Les résultats au niveau de la structure textuelle montrent de légères différences entre les deux candidats : si Macron produit des tweets typiquement composés de deux phrases, Marine Le Pen exploite de manière équilibrée la structure d'une phrase et celle de deux phrases.

4.2 Syntaxe

Au niveau de la syntaxe, nous avons d'abord déterminé la structure de phrase utilisée (Riegel et al. 1994) : simple ou complexe. Le graphique (2) montre la distribution des structures syntaxiques dans le corpus analysé.

Graphique 2 : La structure de phrase dans les tweets produits par Marine Le Pen et Emmanuel Macron

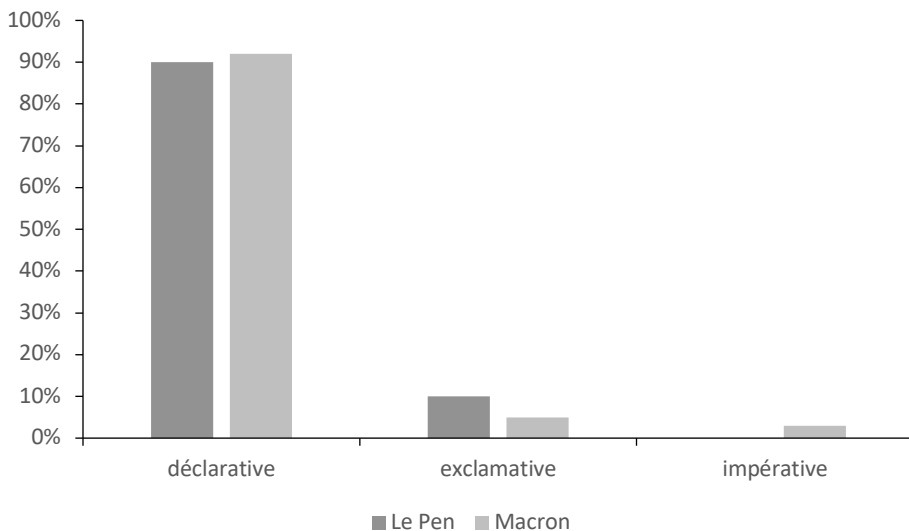


Les résultats montrent que les tweets produits par Le Pen se caractérisent par une utilisation prédominante de phrases complexes (67 %), tandis que la phrase simple n'est pas fréquemment utilisée (33 %). Ceci est démontré par les exemples (5) et (6) :

- (5) J'appelle les Français à me soutenir car je suis la mieux placée pour gagner au second tour de l'élection présidentielle contre Emmanuel Macron : tous les sondages le montrent !
- (6) Je protégerai les Français des répercussions de la crise en #Ukraine sur leur pouvoir d'achat.
Les tweets produits par Emmanuel Macron indiquent une proportion égale entre les phrases simples (50 %, exemple 7) et complexes (50 %, exemple 8).
- (7) L'école, la formation de nos jeunes est un chantier absolument essentiel.
- (8) Bientôt, le Service National Universel permettra à chaque jeune de financer totalement l'obtention du permis de conduire.

Dans la deuxième étape de l'analyse syntaxique, nous avons déterminé le type de phrase utilisée pour chaque tweet du corpus selon la classification de Riegel et al. (1994 : 101) : phrase déclarative, exclamative, interrogative et injonctive. Le graphique (4) montre la distribution des types de phrases dans le corpus analysé.

Graphique 3 : Les types de phrases dans les commentaires produits par Marine Le Pen et Emmanuel Macron



Les résultats montrent que dans les tweets produits par les deux candidats, la phrase déclarative prédomine (90 % et 91 % exemples 9 et 10) :

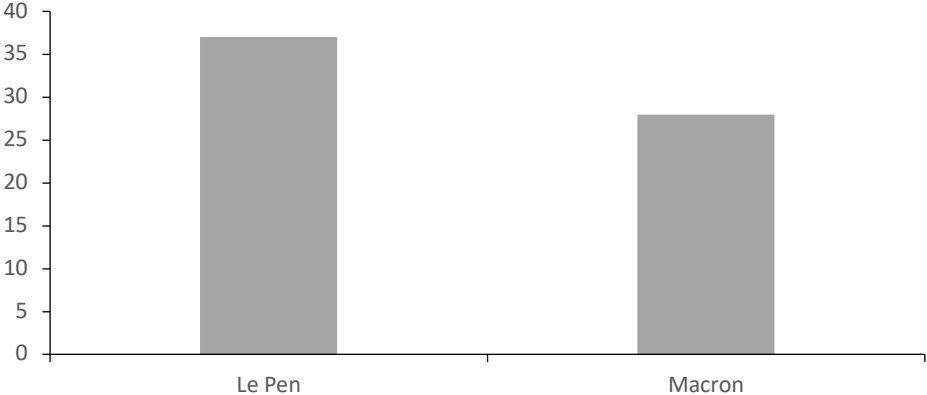
- (9) Agression d'Yvan Colonna par un détenu islamiste: il y a eu une défaillance extrêmement grave qui aurait dû entraîner la démission du ministre de la Justice.
- (10) Cette aide qui existe jusqu'à six ans de manière imparfaite, nous la mettrons jusqu'à douze ans pour les familles monoparentales.

L'analyse syntaxique des tweets analysés démontre que les différences entre les deux candidats s'esquissent au niveau de la structure de phrase et pas vraiment au niveau du type de phrase. Même si on peut trouver quelques exemples de phrase exclamative auprès des deux candidats, (10 % d'exemples analysés), cela démontre que la phrase déclarative est la structure syntaxique de choix dans le discours politique des deux candidats.

4.3 Structures énonciatives

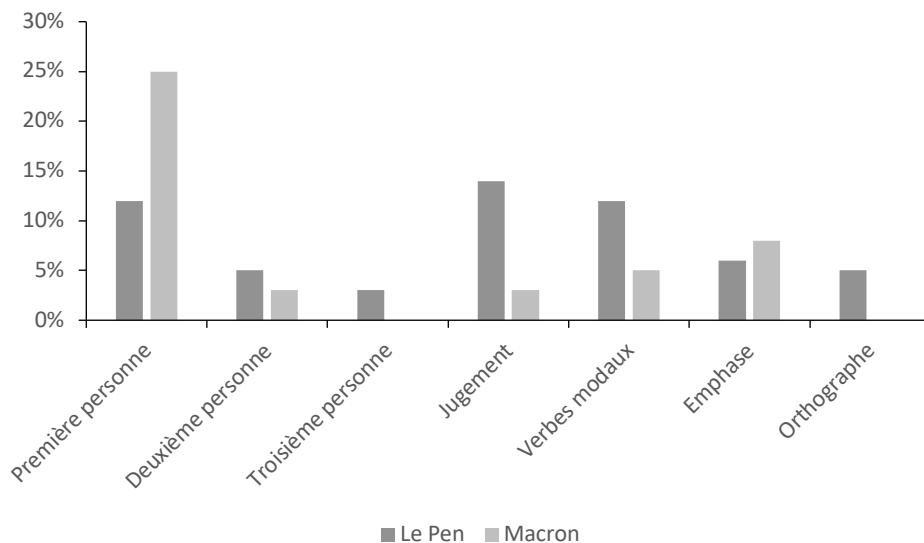
Au niveau énonciatif, nous avons essayé d'examiner si les tweets étudiés contiennent des structures révélant des traces subjectives de l'auteur (*Je veux* ou *C'est indigne*, par exemple). À partir de ce critère, nous avons établi des catégories suivantes : première personne (*J'appelle les Français à me soutenir*), deuxième personne (*Vous êtes souverain*), troisième personne (*Ce sont EUX qui sont responsables*), jugement de valeur (*il a aggravé toutes les injustices, toutes les absurdités et tous les dangers*), orthographe (*il se dit «pessimiste» et «inquiet»*), emphase (*voilà ce que je veux pour le pays*), verbes modaux (*Il faut en parler avec sang-froid*). Le graphique (4) montre la fréquence des structures énonciatives identifiées auprès de chacun des deux candidats.

Graphique 4 : La fréquence des structures énonciatives dans les commentaires produits par Marine Le Pen et Emmanuel Macron



Il ressort du graphique (4) que Marine Le Pen utilise plus de structures énonciatives (37 exemples dans le corpus analysé) qu’Emmanuel Macron (28 exemples). Ce qui semble d’autant plus intéressant que de connaître la fréquence de structures énonciatives, c’est de connaître leur distribution, ce que démontre le graphique (5) :

Graphique 5 : La distribution des structures énonciatives dans les commentaires produits par Marine Le Pen et Emmanuel Macron



À partir du graphique 5, nous pouvons constater que les tweets produits par Marine Le Pen contiennent majoritairement des structures de jugement de valeur (exemple 11), des verbes modaux (exemple 12) et des pronoms de deuxième personne (exemple 13):

- (11) L’Union européenne, **enfermée dans des dogmes absurdes** et des traités technocratiques, a interdit toutes les politiques de patriotisme économique (...)
- (12) **Je souhaite que** les demandes d’asile soient effectuées dans les ambassades et les consulats de France dans n’importe quel pays tiers (...)
- (13) On peut faire des choix différents dans **notre pays** pour peu que **VOUS le décidiez** !

Les tweets produits par Emmanuel Macron, au contraire, se caractérisent plutôt par une forte présence de pronoms de la première personne (25 %, exemple 14) et par la structure de l’emphase (exemple 15) :

(14) **Nous avons fait beaucoup de choses** mais nous devons aller beaucoup plus loin.

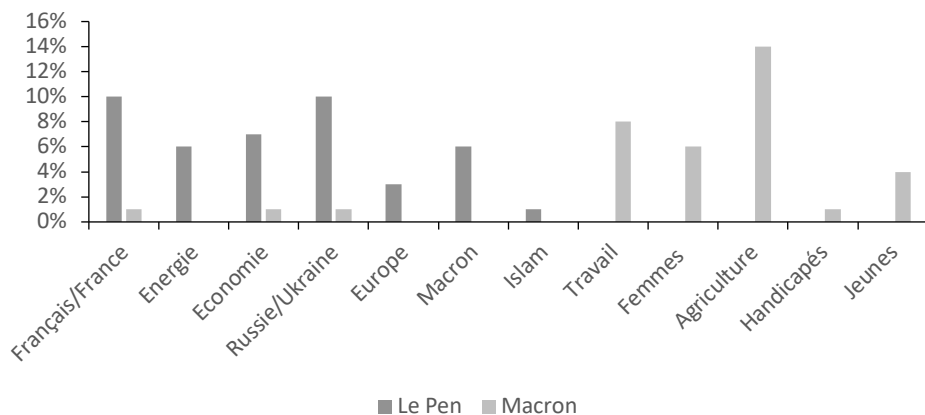
(15) **La grande cause du prochain quinquennat**, si les Françaises et les Français me font confiance, **ce sera** l'égalité femmes-hommes, encore.

L'analyse des structures énonciatives a démontré de légères différences au niveau de la fréquence de leur utilisation et des différences assez évidentes au niveau de leur distribution dans le discours des deux candidats. Vu le caractère persuasif de la communication sur Twitter, nous pouvons interpréter que Marine Le Pen utilise les jugements de valeur surtout pour culpabiliser les politiciens européens au pouvoir (*dogmes absurdes de l'Union européenne*), des verbes modaux pour visualiser les avantages de ses projets (*je souhaite que*) et les pronoms de la deuxième personne pour donner de l'importance aux électeurs (*pour peu que VOUS le décidiez*). Emmanuel Macron, de son côté, utilise les pronoms de la première personne pour souligner son investissement personnel et l'enthousiasme de son équipe (*nous avons fait beaucoup de choses*) et la structure de l'emphase pour mettre un accent particulier sur l'importance de ses projets (*La grande cause du prochain quinquennat, ce sera l'égalité femmes-hommes*).

4.4 Thématique du discours

Au niveau thématique, nous avons examiné la ou les thématiques principales des tweets étudiés (par exemple, *l'énergie* ou *les jeunes*). Les exemples liés à plusieurs thématiques ont été annotés par toutes les thématiques pertinentes (*Je protégerai les Français des répercussions de la crise en Ukraine sur leur pouvoir d'achat* ; thématiques Français, Ukraine, économie). Le graphique (6) montre la distribution de tweets catégorisés selon leur(s) thématique(s) prédominante(s).

Graphique 6 : Les thématiques dans les tweets produits par Le Pen et Macron



(exemple 16), la guerre en Ukraine et les conflits potentiels avec la Russie (exemple 17), l'économie et le pouvoir d'achat, les carburants et le contre candidat Emmanuel Macron.

(16) Vous, peuple de France, vous êtes souverain.

(17) Le président de la République se sert de la guerre en #Ukraine pour faire peur aux Français.

Les commentaires d'Emmanuel Macron sont liés aux thématiques suivantes : l'agriculture (exemple 18), le travail (exemple 19), les femmes (ce qui est probablement lié à la proximité de la journée internationale des femmes lors de la compilation du corpus) et les jeunes.

(18) Depuis 2017, Emmanuel Macron agit sans relâche pour améliorer le quotidien de nos agriculteurs et garantir les conditions de notre souveraineté alimentaire.

(19) Plus question qu'un jeune diplômé peine à trouver un travail parce qu'il n'a pas assez d'expérience professionnelle.

Nous pouvons constater que chacun des deux candidats aborde son propre champ thématique et que les thématiques abordées par les deux candidats n'apparaissent que dans quelques tweets.

5 CONCLUSION

La présente étude a démontré qu'il est possible d'identifier des différences systématiques d'ordre linguistique entre les messages produits par deux candidats politiques différents. Certaines catégories linguistiques que nous avons prises en compte (nombre de phrases, structure et type de phrase, thématique) ne représentent pas de critère d'analyse central dans des études qualitatives du discours politique, mais se sont avérées comme éléments importants de distinction systématique entre le discours des deux candidats.

Les tweets de Marine Le Pen sont formulés en une ou deux phrases de structure complexe et comportent beaucoup de marqueurs énonciatifs parmi lesquels prédominent l'utilisation des jugements de valeur, des verbes modaux et des pronoms à la deuxième personne. Parmi les thématiques qu'elle aborde prévalent la France, la guerre en Ukraine et l'économie. Dans le discours d'Emmanuel Macron, nous avons identifié une proportion égale de tweets comportant une phrase et deux phrases ainsi qu'une proportion égale de phrases simples et complexes. Au niveau énonciatif, les structures prédominantes sont l'emphase et l'utilisation des pronoms à la première personne. Macron évoque surtout les thématiques se rapportant à l'agriculture, au travail et aux jeunes.

S'il est vrai que l'échantillon analysé est limité, les résultats montrent une cohérence tout au long de l'analyse. La méthodologie proposée par la présente étude permet aux linguistes sans formation en linguistique computationnelle de réaliser des études systématiques de discours politique ou d'autres types de genres textuels présents sur Twitter. Elle est applicable aussi à des corpus plus étendus comportant d'autres profils de politiciens et de sujets abordés.

À l'époque où les méthodes de prédiction analytique et de profilage de l'auteur sont utilisées principalement par de grandes entreprises et des corporations, il semble justifié de faciliter des recherches linguistiques et humanistes plus approfondies. Ces recherches peuvent fournir des connaissances utiles pour comprendre les phénomènes politiques et pour prendre des décisions informées et responsables.

Sources primaires

Emmanuel Macron – Profil officiel pendant la campagne électorale
<https://mobile.twitter.com/macronensemble>

Marine Le Pen – Profil officiel
https://twitter.com/MLP_officiel

Références

- BENVENISTE, Émile (1970) « L'appareil formel de l'énonciation. » *Langages* 5/17, 12–18. ISSN 0458-726X.
- BOURDIEU, Pierre (1998) *Les règles de l'art, genèse et structure du champ littéraire*. Paris : Le Seuil.
- DAYTER, Daria (2014) « Self-praise in microblogging. » *Journal of Pragmatics* 61/1, 91–102.
- DELMAS, Virginie (2012) « Pour une analyse pluridimensionnelle du discours : le discours politique. » *La Linguistique* 48, 103–122.
- ELLISON, Nicole/Rebecca HEINO/Jennifer GIBBS (2006) « Managing impressions online : Self-presentation processes in the online dating environment. » *Journal of Computer-Mediated Communication* 11, 415–441.
- FARZINDAR, Atefeh/Mathieu ROCHE (2013) « Les défis du traitement automatique du langage pour l'analyse des réseaux sociaux. » *Revue TAL – Traitement Automatique des langues* 54/3, 7–16.
- HAMON, Thierry/Amel FRAISSE/Patrick PAROUBEK/Pierre ZWEIGENBAUM/Cyril GROUIN (2015) « Analyse des émotions, sentiments et opinions exprimés dans les tweets : présentation et résultats de l'édition 2015 du défi fouille de texte (DEFT). » *Actes de la 22e conférence sur le Traitement Automatique des Langues Naturelles (TALN 2015)*. Caen : HAL.
- LE BART, Christian (1998) *Le discours politique*. Paris : PUF.
- O'CONNOR, Brendan/Ramnath BALASUBRAMANYAN/Bryan ROUTLEDGE/Noah A. SMITH (2010) « From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. » *Proceedings of the International AAAI Conference on Weblogs and Social Media* 11/1, 1–2.

- PANCKHURST, Rachel (2006) « Le discours électronique médié : bilan et perspectives. » *Lire, Écrire, Communiquer et Apprendre avec Internet* 1, 345–366.
- PANG, Bo/Lillian LEE (2008) « Opinion Mining and Sentiment Analysis. » *Foundations and Trends® in Information Retrieval* 1/2, 1–135.
- PRINCE, Violaine/Yves KODRATOFF (2007) « Le Défi fouilles de textes : quels paradigmes pour la reconnaissance d’auteurs ? » *Revue des Nouvelles Technologies de l’Information RNTI E* (10), 001–014.
- RIEGEL, Martin/Jean Christophe PELLAT/René RIOUL (1994) *Grammaire méthodique du français*. Paris : PUF.
- RINN, Michael (éditeur) *Émotions et discours. L’usage des passions dans la langue*. Rennes : Presses universitaires de Rennes.
- ROGINSKY, Sandrine/Barbara DE COCK (2015) « Faire campagne sur twitter: Modalités d’énonciation et mises en récit des candidats à l’élection européenne. » *Les Cahiers du numérique* 11, 119–144.
- SANDRÉ, Marion (2009) « Analyse d’un dysfonctionnement interactionnel – l’interruption – dans le débat de l’entre-deux-tours de l’élection présidentielle de 2007. » *Mots. Les langages du politique* 89, 69–81.
- SMAILOVIĆ, Jasmina/Miha GRČAR/Nada LAVRAČ/Martin ŽNIDARŠIČ (2014) « Stream-based active learning for sentiment analysis in the financial domain. » *Information Sciences* 285, 181–203.
- SMAILOVIĆ, Jasmina/Janez KRANJC/Miha GRČAR/Martin ŽNIDARŠIČ/Igor MOZETIČ (2015) Monitoring the Twitter sentiment during the Bulgarian elections. *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics*, E. Gaussier (éd), 1–10.
- STAMATATOS, Efstathios (2013) On the robustness of authorship attribution based on character n-gram features. *Journal of Law and Policy* 21/2, 421-439.
- SUIGNARD, Philippe/Caroline ESCOFFIER/Lou CHARAUDAU/Mathieu BRUGIDOU (2021) « Que peuvent les algorithmes de plongement de mots pour l’analyse sociologique des textes ? Analyser les discours et caractériser les locuteurs des plateformes « Grand Débat National » et « Vrai Débat » » *Statistique et société* 9/1-2, 133–145.
- ZWITTER VITEZ, Ana (2020) « Le discours politique et l’expression de l’opinion sur Twitter : analyse syntaxique, lexicale et orthographique. » *Ars et humanitas* 14/1, 157–170.

Résumé

LA STRUCTURE LINGUISTIQUE DE TWEETS EN CAMPAGNE PRÉSIDENTIELLE

L’objectif de cette étude est de faire une analyse systématique de tweets publiés par Emmanuel Macron et Marine Le Pen pendant la campagne présidentielle en 2022. L’analyse est réalisée aux niveaux textuel, syntaxique, énonciatif et thématique. Les

résultats montrent de légères différences au niveau textuel et syntaxique et des différences saillantes au niveau énonciatif (jugements de valeur, verbes modaux, emphase) et thématique. La méthodologie proposée permet aux linguistes sans compétences computationnelles d'obtenir des résultats quantifiables et en même temps interprétables par des catégories linguistiques traditionnelles.

Mots-clés : discours politique, réseaux sociaux, analyse syntaxique, analyse énonciative, analyse quantitative

Abstract

THE LINGUISTIC STRUCTURE OF TWEETS IN A FRENCH PRESIDENTIAL ELECTION CAMPAIGN

The aim of this study is to carry out a systematic analysis of tweets published by Emmanuel Macron and Marine Le Pen during the 2022 presidential election campaign. We performed the analysis on textual, syntactic, discourse and thematic levels. The results show slight differences on the textual and syntactic level but major differences on the discourse level (value judgments, modal verbs, emphasis) as well as on the thematic level. The proposed methodology allows linguists without computational skills to obtain results that are quantifiable and at the same time interpretable using traditional linguistic categories.

Keywords: political discourse, social networks, syntactic analysis, discourse analysis, quantitative analysis

Povzetek

JEZIKOVNA PODOBA TVITOV V FRANCOSKI PREDSEDNIŠKI KAMPANJI

V prispevku predstavljamo sistematično analizo izbranih tvitov, ki sta jih med francosko predsedniško kampanjo leta 2022 objavila kandidata Emmanuel Macron in Marine Le Pen. Analizo smo izvedli na besedilni, skladijski, diskurzivni in tematski ravni. Rezultati kažejo precejšnjo podobnost na besedilni in skladijski ravni ter izrazite razlike na diskurzivni (vrednostne sodbe, modalni glagoli, poudarjeni zaimki) in tematski ravni. Predlagana metodologija omogoča, da jezikoslovci brez računalniškega predznanja pridobijo kvantitativno preverljive rezultate, ki jih je mogoče interpretirati s tradicionalnimi jezikovnimi kategorijami.

Ključne besede: politični diskurz, družbena omrežja, skladijska analiza, diskurzivna analiza, kvantitativna analiza