



**HAL**  
open science

# COMMENT LE DISCOURS POLITIQUE FRANÇAIS THEMATISE LA QUESTION DE L'EUROPE : QUELQUES ELEMENTS DE REPONSE A PARTIR DU CORPUS POLITITWEETS

Julien Longhi

► **To cite this version:**

Julien Longhi. COMMENT LE DISCOURS POLITIQUE FRANÇAIS THEMATISE LA QUESTION DE L'EUROPE : QUELQUES ELEMENTS DE REPONSE A PARTIR DU CORPUS POLITITWEETS. L'Europe dans les médias en ligne, Jun 2016, Gennevilliers, France. halshs-01508988

**HAL Id: halshs-01508988**

**<https://shs.hal.science/halshs-01508988>**

Submitted on 15 Apr 2017

**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.

**COMMENT LE DISCOURS POLITIQUE  
FRANÇAIS THEMATISE LA QUESTION DE  
L'EUROPE : QUELQUES ELEMENTS DE  
REPONSE A PARTIR DU CORPUS  
*POLITITWEETS***

Julien Longhi, Université de Cergy-Pontoise, AGORA

Les réseaux sociaux numériques constituent un terrain qui renouvelle l'analyse du discours politique, notamment en proposant de nouveaux observables au chercheur. Cet article illustre et valorise le recours à ces nouvelles données, par le recours au corpus *Polittweets* (Julien Longhi *et al.*, 2014). Constitué dans le cadre d'une recherche centrée sur le lexique politique, à partir d'analyses issues des nouveaux moyens de communication, il permet de faire des analyses spécifiques sur différents thèmes ou phénomènes linguistiques. Dans cet ouvrage consacré à l'Europe et aux discours de / sur l'Europe, nous proposons de décrire la manière dont ce sujet est thématisé dans ce corpus, afin de saisir la façon dont le discours politique français investit cette question.

Cette étude sera menée en trois temps. Dans un premier temps, nous présenterons brièvement le contexte de l'étude (l'analyse du discours numérique et le tweet politique comme genre du discours politique, le corpus *Polittweets*), puis nous procéderons au traitement du corpus, en lien avec l'Europe. Nous détaillerons tout d'abord l'analyse des tweets qui contiennent la forme « Europe », afin de voir comment ces messages qui lexicalisent ce thème de l'Europe se caractérisent. Puis nous terminerons cette étude en observant, dans un sous-corpus de tweets idéologiques identifiés dans une précédente recherche (Sarah Djemili, Julien Longhi *et al.*, 2014, 16-25), une thématique liée à l'Europe (identifiée grâce au logiciel *Iramuteq*). Nous montrerons ainsi que le thème de l'Europe est abordé dans le discours politique numérique français de manière négative, problématique ou anxiogène, ce qui éclaire peut-être, comme cause ou conséquence, la représentation qu'en ont les citoyens. Nous verrons aussi que cette vision ne diffère pas forcément de celle des médias en ligne, ce qui renforce l'intérêt de l'étude des médias sociaux comme sources d'information.

## **L'ANALYSE DU DISCOURS POLITIQUE NUMERIQUE**

Nous considérons que le tweet politique est un genre spécifique du discours politique (Julien Longhi, 2013) et avons donc cherché, par la constitution de *Polittweets*, à obtenir un corpus qui permette de travailler sur ce genre de discours, de le caractériser, et de l'appréhender sous différentes formes.

### **1.1. L'ANALYSE DU DISCOURS NUMERIQUE ET LE TWEET POLITIQUE COMME GENRE DU DISCOURS POLITIQUE**

La constitution d'un corpus de tweets correspond à un double objectif, pensé dès le départ comme des éléments solidaires d'une même recherche :

1. Se doter d'un corpus pour réaliser une recherche centrée sur le lexique politique, à partir d'analyses d'observables issus des nouveaux moyens de communication.
2. Participer dans le cadre du projet CoMeRe<sup>1</sup> à la constitution d'un ensemble de corpus de communications médiatisées par les réseaux.

Avec l'intégration de cette recherche dans le projet CoMeRe, se pose la question des discours médiés par les réseaux, et concernant les tweets également, la question du discours numérique. En effet<sup>2</sup>, Twitter est un service de microblogage, permettant aux utilisateurs de publier des messages courts (140 caractères maximum). Le slogan d'origine de Twitter, « What are you doing? » le définissait comme un service permettant de raconter ce qu'on fait au moment où on le fait. Prenant acte de l'utilisation du service pour s'échanger des informations et des liens, Twitter le remplace par « What's happening? » (« Quoi de neuf ? » ou encore « Que se passe-t-il ? »

---

<sup>1</sup> <https://corpuscomere.wordpress.com/apropos/> : corpus de communication médiée par les réseaux (consulté le 11/04/2017).

<sup>2</sup> Nous utilisons des aspects de la page Wikipedia consacrée à Twitter, mais l'ajustons / l'enrichissons en fonction d'évolutions ou spécificités.

dans la version française). Twitter s'enrichit régulièrement de nouvelles fonctionnalités : insertion d'images, de vidéos, possibilité de commenter des tweets, etc. L'usage fait par les politiciens se fonde donc sur les possibilités offertes, notamment la production de messages courts, éventuellement agrémentés d'une image et prenant en compte des éléments technodiscursifs (Marie-Anne Paveau, 2013) propres au réseau. Paveau (2013, 2) montre que « les discours sur Twitter, comme sur les RSN et le web 2.0 en général, sont difficilement analysables hors de leur écosystème : il s'agit en effet d'écritures natives en ligne, de nature composite, car intégrant une dimension technologique », considérant notamment « le tweet comme forme technodiscursive spécifique » et « le hashtag comme technoforme fournissant au discours de Twitter l'une de ses caractéristiques fondamentales, son investigabilité ». Le tweet politique est donc une forme particulière du discours politique, qui en reprend certaines caractéristiques, mais les rend spécifiques au regard des conditions technologies propres au réseau.

## **1.2 LE CORPUS *POLITITWEETS***

Le choix des comptes Twitter a été guidé par le critère de l'influence sur ce réseau, afin de tenir compte des personnalités ont le plus d'influence dans le domaine de la politique sur Twitter. Pour cela, les étapes suivantes ont été appliquées :

1. Partir de 7 personnalités de 6 groupes politiques ;
2. Récupérer toutes les listes créées par des utilisateurs où ils étaient cités : cela a conduit à 7087 listes ;
3. Sélectionner parmi ces listes celles qui avaient au moins 6 twittos (utilisateurs de Twitter) et qui contenaient la chaîne de caractères \*politic\* dans le nom ou descriptif de la liste : 120 listes ;
4. Sur ces 120 listes récupérer les 2934 twittos ;
5. Travail par seuil, en ne retenant que les comptes présents dans plus de 12 listes : nous arrivons ainsi à 205 twittos politiques ;
6. Sur ces 205 comptes nous avons récupéré les 200 derniers tweets de chacun au 27 mars 2014, soit 34273 tweets.

Cette procédure nous permet d'avoir un ensemble de messages provenant de comptes « politiques », en garantissant que cette

étiquette n'est pas projetée sur les données par le chercheur, mais qu'elle est choisie par les usagers (à l'étape 3) par la dénomination de leur liste avec la chaîne \*politic\*.

Les tweets, ainsi que les métadonnées (extraites grâce à l'API Twitter), sont stockés dans une base de données, qui permet de recueillir le texte des tweets ainsi que l'ensemble des informants. Voici un exemple avec quelques champs sélectionnés :

tweet_text	screen_name	favoritecount	retweetcount
Toute ma gratitude va à l'extraordinaire personnel ...	valtrier	2286	6337
Merci du fond du cœur à tous ceux qui ont envoyé d...	valtrier	1830	3429
"Soyons dignes, soyons patriotes, soyons Français ...	NicolasSarkozy	1632	5711
Avant d'entrer à la présidence de la République, je...	fhollande	1401	6806
Que chacun d'entre vous sache combien je suis reco...	NicolasSarkozy	1350	2617
"Vous pourrez compter sur moi à chaque fois qu'il y...	NicolasSarkozy	614	1832
"Jamais je ne pourrai vous rendre tout ce que vous...	NicolasSarkozy	499	2499
@nicodomenach je pourrais aussi entrer en arrière ...	ramayade	464	2979
Joie d'accueillir le père Georges, rayonnant de gén...	fhollande	446	1129
#FF @Elysee pour suivre toutes les informations su...	fhollande	427	993

## 1 Capture d'écran de la base de données développée au laboratoire ETIS

L'enjeu pour une approche linguistique est donc d'utiliser ce matériau pour élaborer le corpus *Politiitweets*. Ces données ont donc été mises en forme selon le standard TEI (le lecteur pourra abondamment se documenter sur le format TEI en ligne<sup>3</sup>, ou voir Longhi et Wigham, 2015 pour une synthèse) : notamment, les éléments précédés de # ou @ (les technomots, cf. Paveau) sont « codés » grâce au format TEI, afin de baliser les différents aspects sémiotiques et technologiques spécifiques au tweet. Ainsi, voici un tweet tel qu'il apparaît dans le corpus :

<sup>3</sup> <http://www.tei-c.org/index.xml> (consulté le 11/04/2017)

```

<post xml:id="cmr-politweets-a448491380948885504" who="#cmr-politweets-p80820758"
when="2014-03-25T17:07:32" xml:lang="fra">
  <p><distinct type="twitter-retweet"><ident>RT</ident>
  <addressingTerm><addressMarker>@</addressMarker><addressee type="twitter-account"
  ref="https://twitter.com/LePG_35022597"
  >LePG</addressee></addressingTerm>:</distinct> A <distinct type="twitter-hashtag"
  ><ident>#</ident><rs ref="https://twitter.com/search?q=%23Bruxelles&src=hash"
  >Bruxelles</rs></distinct>,
  <addressingTerm><addressMarker>@</addressMarker><addressee type="twitter-account"
  ref="#cmr-politweets-p80820758">JLMelenchon</addressee></addressingTerm> conclue
  la rencontre-débat sur le <distinct type="twitter-hashtag"><ident>#</ident><rs
  ref="https://twitter.com/search?q=%23GMT&src=hash">GMT</rs></distinct>. Nous
  live-tweetons - <distinct type="twitter-hashtag"><ident>#</ident><rs
  ref="https://twitter.com/search?q=%23UE&src=hash">UE</rs></distinct>
  <distinct type="twitter-hashtag"><ident>#</ident><rs
  ref="https://twitter.com/search?q=%23USA&src=hash">USA</rs></distinct>
  <distinct type="twitter-hashtag"><ident>#</ident><rs
  ref="https://twitter.com/search?q=%23Europe&src=hash">Europe</rs></distinct> -
  <ref target="http://t.co/hNGmnAagIM">http://t.co/hNGmnAagIM</ref></p>
</trailer>
<fs>
  <f name="medium">
    <string>Twitter for iPhone</string>
  </f>
  <f name="retweetcount">
    <numeric value="22"/>
  </f>
  <f name="isRetweet">
    <binary value="true"/>
  </f>
  <f name="retweetedstatus_id">
    <numeric value="448490939338588160"/>
  </f>
</fs>
</trailer>
</post>

```

## 2 Mise en forme du tweet dans le corpus

Ce format (issu du XML, « langage de balisage extensible ») est particulièrement bien adapté à la constitution de documents numériques sur lesquels des analyses pourront être menées, du fait du balisage qui permet l'introduction de métadonnées.

C'est donc à partir de ce corpus que l'analyse est menée, notamment à partir de la recherche de la forme « Europe ».

## L'ANALYSE DES TWEETS QUI CONTIENNENT LA FORME « EUROPE »

Dans un premier temps, nous avons donc cherché dans les tweets du corpus ceux qui contiennent la forme « Europe ». Nous obtenons ainsi 343 tweets, ce qui est assez peu mais cohérent puisque le 27 mars 2014, date d'extraction du corpus, est une date de l'entre-deux tours de l'élection municipale 2014 : les messages sont donc

davantage centrés sur le contexte local ou national.

## 2.1. LES FORMES EMPLOYEES DANS LES TWEETS CONTENANT « EUROPE »

Le relevé du nombre d'occurrences des formes (lemmes) les plus employées, avec le logiciel *Iramuteq*, est intéressant :

europe	342
européen	30
france	23
ukraine	21
politique	20
débat	18
ue	17
notreeurope	17
grand	16
besoin	15
projet	13
russie	12
proposition	11
nation	11
bruxelles	10

### 3 Formes les plus employées dans les tweets contenant la forme « Europe »

Ce tableau nous indique que le recours à la forme « Europe » se fait notamment dans des tweets qui mettent en discours l'Europe sous une autre forme (« européen », « ue » ou « notreeurope »), ou d'autres entités (« france », « ukraine », « russie », « bruxelles », « nation »). On trouve sinon des formes en lien avec le domaine politique (« politique », « débat », « projet »).

Si on s'intéresse aux co-occurrences, c'est-à-dire aux formes employées avec « Europe », et plus largement celles qui gravitent dans son environnement, on obtient le schéma suivant :



#### 4 Analyse des similitudes des tweets contenant « Europe »

On constate que deux champs sémantiques principaux sont présents : la guerre (« défendre », « désarmer », « guerre ») et le domaine politique, soit par les formes d'expression (« tribune », « débat », « blog », « colloque »), soit par des thématiques (« écologie », « valeur », « démocratie », etc.).

## 2.2. LES THEMES DEVELOPPES A PROPOS DE L'EUROPE

Cela se retrouve dans la classification thématique que propose *Iramuteq*, à partir du modèle Alceste : dans ce petit corpus, le



logiciel détecte deux classes. Les énoncés caractéristiques de la première sont sur le thème du renouveau :

\*\*\*\* \*Europe \*Michel\_Barnier\_10

score : 18.63

rteu\_marktl entreprenariat social offre de nouveaux modèles économiques dans une europe qui change découvrez en

\*\*\*\* \*Europe \*Hervé\_Morin\_7

score : 17.73

rtlinternautefodirectpolitique on a besoin d uneeurope intégrée puissante capable de porter le modèle européen dans le monde

\*\*\*\* \*Europe \*Harlem\_Désir\_6

score : 17.56

rt ps\_ep2014 harlemdesirl europe a besoin d un nouveau souffle il faut rompre avec les mesures imposées par les conservateur

Nous interprétons cette classe comme thématissant le besoin de renouveau lorsqu'il est question de l'Europe. En effet, les termes « nouveau », « change », « besoin », indiquent que les tweets consistent à demander ou proposer des changements pour l'Europe telle qu'elle est conçue.

L'autre classe est beaucoup plus contextuelle, et évoque l'Europe dans le contexte de l'invasion de la Crimée par la Russie, contestée par l'Ukraine :

\*\*\*\* \*Europe \*J\_P.\_Chevènement\_21

score : 112.58

lukraine devrait être un pont entre la russie et l europe plutôt qu une pomme de discorde cc gaucheavenir

\*\*\*\* \*Europe \*Yannick\_Jadot\_1

score : 95.58

ukraine nous soutenons la perspective de sanctions mais l europe est ambiguë et divisée sur la russie ee2014

\*\*\*\* \*Europe \*J\_P.\_Chevènement\_35

score : 95.58

dépêcheafpukrainechevènement plus de raison pour une guerre  
froide en europeukrainerussieeurope

\*\*\*\* \*Europe \*J\_P.\_Chevènement\_3

score : 94.54

c est une erreur pour l europe de vouloir choisir entre l ukraine  
et la russie nous devons nous associer les uns et les autres  
28min

\*\*\*\* \*Europe \*J\_P.\_Chevènement\_15

score : 88.36

je pense que l europe devrait être créatrice de paix et de ne pas  
ranimer des remugles de guerre froide 28min ukraine

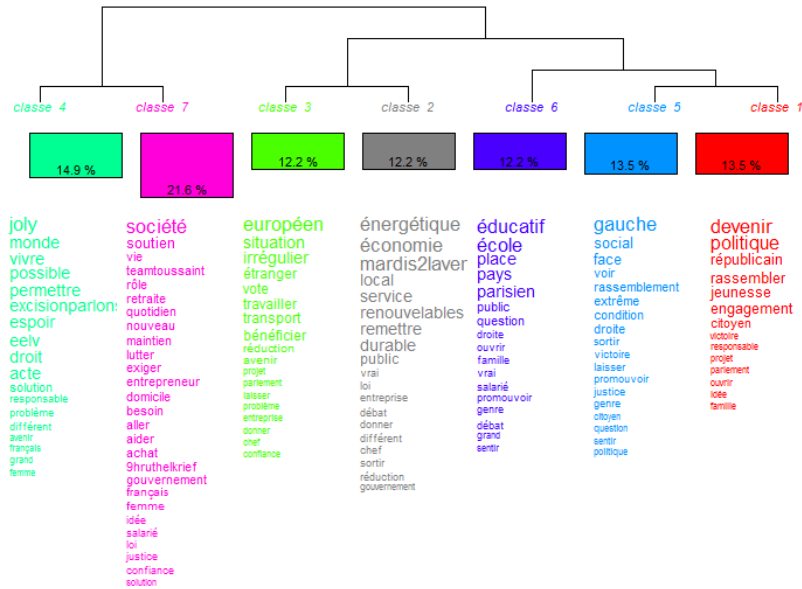
Il est donc question de l'Europe en ce qu'elle doit prendre des  
responsabilités politiques dans cette crise, ou un territoire qui est le  
lieu d'affrontements qui la divisent.

La vision de l'Europe telle qu'elle apparaît dans les tweets qui la  
mentionnent est donc relativement négative, puisqu'elle est associée  
à un renouveau ou à un changement qui serait nécessaire, et à un  
conflit dans lequel elle ne prend pas ses responsabilités. Cette  
méthode focalise sur une forme particulière, ce qui peut introduire un  
certain biais dans le recueil. Aussi, nous proposons de compléter ces  
résultats par une autre analyse : celle-ci s'appuie sur un travail  
précédent (S. Djemili, J. Longhi *et al.*, 2014) dans lequel nous avons  
détecté des tweets idéologiques (nous ne pouvons détailler ici  
l'analyse). Nous allons donc chercher, dans cet échantillon, à voir  
comment l'Europe est présentée.

## **LES TWEETS IDEOLOGIQUES QUI TRAITENT DE L'EUROPE**

En effet, lors de l'analyse Alceste du corpus de tweets  
idéologiques par *Iramuteq*, nous obtenons 7 classes qui permettent de

dégager différentes thématiques :



## 5 Thématiques issues du sous-corpus idéologique

Nous pouvons observer que la classe 3, qui représente environ 12,2% des tweets idéologiques, concerne l'Europe, avec l'adjectif « européen » comme première forme identifiée. Les termes associés à « Europe » sont « irrégulier », « étranger », etc., tel que l'on peut le distinguer dans le nuage de mots spécifique à cette classe :



## 6 Nuage de mots de la classe relative à l'Europe

Pour mieux appréhender la nature des messages produits, voici trois segments de texte les plus caractéristiques de cette classe :

\*\*\*\* \*idéologique \*tweet170

score : 234.52

comme le dit gillescarrez aucune capitale européenne ne donne 75% de réduction dans les transports aux étrangers en situation irrégulière

\*\*\*\* \*idéologique \*tweet169

score : 196.78

les étrangers en situation irrégulière bénéficient de 75% de réduction dans les transports les travailleurs pauvres et les chômeurs payent

\*\*\*\* \*idéologique \*tweet116

score : 120,02

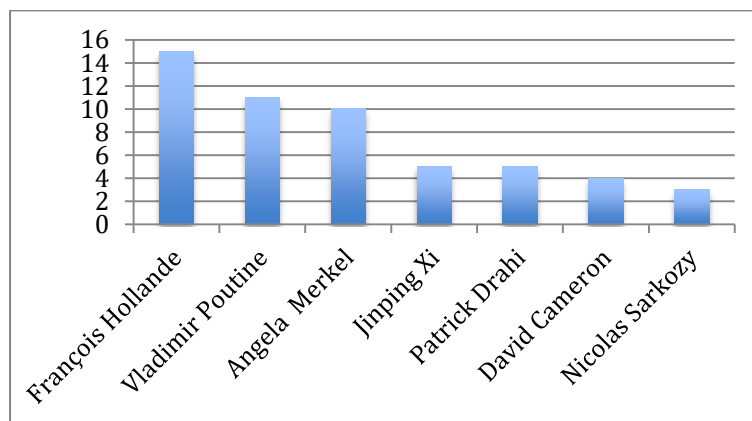
la crise n est pas encore 1 souvenir ms continuons à travailler pr l avenir le cadre d'1 projet européen qui mérite confiance retrouvée

Ces exemples indiquent que l'Europe (qu'elle soit présente formellement ou non) est thématisée en relation avec les aides accordées aux étrangers (présentées comme excessives en France), la nécessité de renouveler le projet européen, ou encore les étrangers en situation irrégulière. On voit donc que selon cette méthode, le recueil des tweets qui s'intègrent à la thématique européenne donnent aussi une image négative à l'Europe.

## MEDIAS SOCIAUX OU MEDIAS EN LIGNE, QUELLES DIFFERENCES ?

Pour terminer cette étude et afin de contribuer au questionnement plus spécifique au projet LEMEL à propos de l'Europe dans les médias en ligne, nous avons recherché dans la base de journaux *Factiva* les articles contenant le terme « Europe » publiés entre le 1<sup>er</sup> et le 27 mars 2014 par l'AFP, afin de voir si la représentation médiatique « en ligne » est cohérente avec celle observée par les comptes politiques influents.

*Factiva* permet d'extraire plusieurs types d'informations, par exemple les dirigeants cités dans les articles. Pour ce petit corpus de l'AFP contemporain de notre corpus de tweets, voici les dirigeants cités :



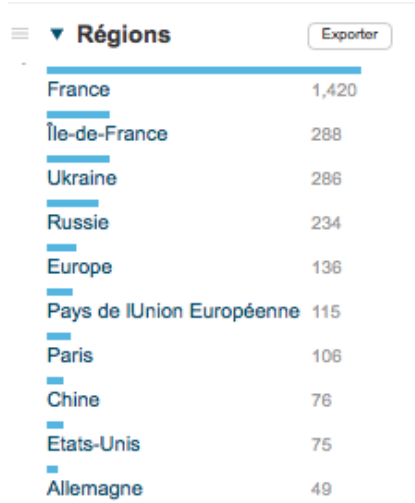
7 Dirigeants cités dans la presse en ligne

Si nous mettons cela en rapport avec une autre information, celle liée aux mots-clés, l'interprétation devient plus claire :

second tour premier tour  
clôture négative  
charcuterie française  
crise ukrainienne imperial tobacco  
des effectifs villes intelligentes  
lockdown mode élections municipales

## 8 Mots-clés dans la presse en ligne

En effet, avec « François Hollande », « Nicolas Sarkozy », « second tour », « premier tour », nous percevons que le premier sujet de la presse est la campagne municipale, même dans les articles qui mentionnent l'Europe. Mais la question russe est bien présente, avec « Vladimir Poutine » et « crise ukrainienne ». Ceci est confirmé par les « régions » convoquées dans ces articles :



## 9 Régions relatives aux articles

La France est en tête, mais l'Ukraine et la Russie représentent également une part importante des régions convoquées.

On repère également des préoccupations économiques avec « clôture négative » ou « charcuterie française », qui montrent le lien fait dans la presse en ligne entre l'Europe et les questions économiques ou financières.

Ces quelques points étayent à notre sens l'intérêt qu'il y a à coupler une analyse des médias en ligne à une analyse des médias sociaux, puisqu'on peut observer certaines convergences thématiques, et donc objectiver ce qui relève spécifiquement du type de média, ou du sujet. Dans le cas de l'Europe, on peut considérer que ce sujet, tel qu'il est véhiculé dans les médias (en ligne ou sociaux), concerne les enjeux politiques et électoraux, les crises politiques en Europe, ou les questions et difficultés économiques.

## CONCLUSION

Dans cet article, nous avons essayé de contribuer aux questionnements du projet LEMEL, notamment celui-ci : quelles représentations de l'Europe se forment dans, se formulent par, proposent les médias en ligne lors du déroulement d'événements médiatiques, pendant des périodes normales (hors-événement européen majeur) ?<sup>4</sup>. Pour cela, nous avons analysé les mises en discours de la thématique européenne dans un corpus de tweets politiques. Ce traitement, à travers l'analyse des productions issues d'un réseau social, a permis d'éclairer d'une certaine manière la défiance, parfois pointée par les médias, des citoyens envers l'Europe. En effet, la manière dont la classe politique française investit cette thématique se fait largement de manière négative : besoin d'un changement, lien avec des crises, des dérives, des dénonciations, etc. Si les analyses médiatiques du projet LEMEL des résultats conséquents sur les discours des médias européens, notre contribution permet, à travers un autre genre de discours, de comprendre aussi ce qui conditionne le discours sur l'Europe. Les pistes évoquées à la fin de l'article, avec quelques éléments à propos des médias en ligne, incitent à poursuivre le couplage entre l'analyse

---

<sup>4</sup><http://lemel.hypotheses.org/13> (consulté le 11/04/2017)

des médias en ligne et des médias sociaux, car les convergences permettent de faire ressortir les spécificités d'un thème, ce qui peut renforcer son intérêt, alors que les différences permettent de saisir ce qui est relatif à une forme d'information.

## REFERENCES BIBLIOGRAPHIQUES

DJEMILI, Sarah, LONGHI, Julien, *et al.*, « What does Twitter have to say about ideology? ». **In** : FAAß, Gertrud, RUPPENHOFER, Josef, *Workshop Proceedings of the 12<sup>th</sup> Edition of the Konvens Conference* [en ligne], Hildesheim, Universitätsverlag Hildesheim, 2014. Disponible sur <<http://www.uni-hildesheim.de/konvens2014/data/konvens2014-workshop-proceedings.pdf>> (consulté le 11/04/2017), p.16-25

LONGHI, Julien, « Essai de caractérisation du tweet politique », *L'Information grammaticale*, n° 136, 2013.

LONGHI, Julien, MARINICA, Claudia, BORZIC, Boris, ALKHOULI, Abdulhafiz, « Polititweets : corpus de tweets provenant de comptes politiques influents ». **In** : CHANIER, Thierry, *Banque de corpus CoMeRe* [en ligne], Nancy, Ortolang.fr, 2014. Disponible sur < <https://repository.ortolang.fr/api/content/comere/v3.3/cmr-polititweets.html> > (consulté le 11/04/2017)

LONGHI, Julien, WIGHAM, Ciara R., « Structuring a CMC corpus of political tweets in TEI: corpus features, ethics and workflow » [en ligne]. *Corpus Linguistics 2015*, Lancaster, 2015. Disponible sur <<https://halshs.archives-ouvertes.fr/halshs-01176061>> (consulté le 11/04/2017)

PAVEAU, Marie-Anne, « Analyse discursive des réseaux sociaux numériques ». **In** : *Dictionnaire d'analyse du discours numérique, Technologies discursives* [carnet de recherche] [en ligne], 2013. Disponible sur <<http://technodiscours.hypotheses.org/431>> (consulté le 11/04/2017)