Classification d'un corpus hétérogène : la page *Facebook* de soutien au « bijoutier de Nice » (septembre 2013)

Alexia Ducos[1], Valérie Bonnet[2], Pascal Marchand[3], Pierre Ratinaud[4]

[1] Université de Toulouse – alexia.ducos@iut-tlse3.fr
[2] Université de Toulouse – valerie.bonnet@iut-tlse3.fr
[3] Université de Toulouse – pascal.marchand@iut-tlse3.fr
[4] Université de Toulouse – ratinaud@univ-tlse2.fr

Abstract

The lexicometric analysis provided by the *Iramuteq* software (Ratinaud) is applied to the online content of a Facebook page dedicated to the support of the "jeweler from Nice" media event. The corpus, composed of 3,958,482 occurrences, corresponds to the content published between September 11 and September 28, 2013. To optimize the textual analysis, the corpus underwent cleaning operations. In total, the classification operation relates to 31,745 active forms revealed by a first statistic report (nine classes were also revealed at that time). The similarity analysis (ADS) of the textual matrix then allowed us to modelize, via graphs, similar and disparate elements of the Internet users' discourses. Therefore, the main challenge of the lexicometric analysis is indeed to put the discursive and linguistic use of enunciation forward in a specific socio-political context.

Résumé

L'analyse lexicométrique proposée par le logiciel *Iramuteq* (P. Ratinaud) est appliquée ici au contenu en ligne publié sur la page Facebook en soutien au bijoutier de Nice. Cette page a été créée le jour même du fait divers, largement relayé par les médias. L'extraction du contenu de ce corpus a été réalisée sur la période du 11 au 28 septembre 2013, ce dernier était alors composé de 3 958 482 occurrences. Des opérations de nettoyage du corpus ont été nécessaires afin d'optimiser l'analyse textuelle du contenu. Au total, c'est à partir des 31 745 formes actives révélées par un premier rapport statistique que l'opération de classification s'est effectuée (neuf classes sont définies). L'analyse de similitude (ADS) de la matrice textuelle nous a, par la suite, permis de modéliser par graphes les éléments analogues et disparates produits par le discours des internautes. Dès lors, l'enjeu de l'analyse lexicométrique est bien de révéler l'usage discursif et linguistique des commentaires produits, dans un contexte socio-politique spécifique.

Mots-clés : Classification descendante hiérarchique, analyse de similitude, corpus hétérogène, discours polémique, expression citoyenne, *Iramuteq*

1. Introduction

1.1. Contexte

Ce 11 septembre 2013, est largement relaté le récit médiatique de Stephan Türk, bijoutier dans la ville de Nice, qui suite au vol à main armée de son commerce, a blessé mortellement un de ses deux agresseurs par tirs d'arme à feu dans le dos. Ce fait divers s'inscrit dans le paysage médiatique français déjà touché par une série d'attaques similaires ébranlant l'opinion publique. Dans cette affaire, une réaction citoyenne a massivement investi une page sur le réseau social Facebook, créée le jour-même du fait divers, et dénommée « *Soutien au bijoutier de Nice* ». Le propos de la page ne contient qu'une phrase descriptive rédigée par son administrateur « *Soutenons ce bijoutier qui ne faisait que son travail* » ; pour le reste, sa popularité ne tient qu'à deux variables présentes, caractéristiques de la mobilisation

citoyenne : nombre/temps. En effet, c'est la portée « *organique* » ou « *virale* » (soit « *l'effet de buzz* ») permise par la puissance du réseau social qui a permis à la page d'atteindre en seulement trois jours le million de « *likes* ». Cette affaire a connu un développement médiatique particulier à la suite des accusations de « bidonnage » de la page Facebook par les partis d'extrême droite. En effet, ceux-ci auraient, selon certaines rumeurs, acheté des faux « *likes* » à l'étranger. L'intérêt de la constitution du corpus et de son analyse par *Iramuteq* (Ratinaud) nous permettra ainsi d'appréhender la totalité d'un discours particulièrement polémique, dans le cadre de recherches en Sciences de l'Information et de la Communication. Par « *polémique* », nous retiendrons la définition proposée par Ruth Amossy (2011), comme « *un mode de gestion des conflits par la polarisation extrême et la confrontation radicale des positions antagonistes* », complétée par la déclaration de Dominique Garand (1998), pour qui « *toute parole polémique est issue du conflictuel* ».

1.2. Enjeux de l'analyse lexicométrique

Si l'on considère les commentaires publiés sur la page Facebook comme l'apanage d'une expression citoyenne dans laquelle sommeille l'utopie ultime d'une « *e-démocratie* » ou d'une « *démocratie 2.0* » (Flichy, 2011), les données statistiques lexicométriques et leur interprétation sont un premier moyen d'étude vers la recherche sémantique. Il sera pertinent de comprendre les enjeux et caractéristiques du discours polémique. Aussi, la classification lexicale permettra de faire remonter les différents foyers de controverses à plusieurs niveaux, tant sur les plans de la légalité, de la morale ou bien encore de l'émotion. Les questions préalables à l'étude lexicométrique sont donc les suivantes :

- Quelles thématiques discursives sont à l'origine des différents foyers de polémiques ?
- Y a-t-il réellement un débat d'idées ou seulement la révélation d'une agrégation d'opinions partagées ?

2. Le corpus

2.1. Extraction de la page Facebook

Comme nous l'avons auparavant évoqué en introduction, le corpus a été constitué à partir des données publiées sur la page *Facebook* « *Soutien au bijoutier de Nice* » du 11 au 28 septembre 2013. Cette extraction est ainsi majoritairement composée des commentaires d'internautes. Pour extraire le corpus, nous avons utilisé l'API[1] de *Facebook* qui a permis de nous fournir le contenu des commentaires associés à chacun des *posts* d'une page, ainsi que des informations comme la date, le nom de l'auteur du commentaire et le nombre de "*likes*" qu'il a reçu[2]. La première analyse statistique globale sans nettoyage nous révèle les caractéristiques suivantes :

nombre de textes :	132279
nombre d'occurrences :	3948188
nombre de formes :	49032
moyenne d'occurrences par forme :	80.52
nombre d'hapax :	25863 (0.66% des occurrences - 52.75% des formes)
moyenne d'occurrences par texte :	29.85

Tableau 1. Statistiques générales du corpus extrait

[1] *Application Programming Interface*

[2] http://graph.facebook.com

2.2. Opérations de nettoyage

Le corpus en l'état souffre d'une variabilité telle dans la graphie qu'une analyse automatique serait considérée comme peu stable : erreurs orthographiques, approximations syntaxiques, termes orduriers, etc… Une première étape de travail devait alors en rétablir l'homogénéité (Ratinaud et Marchand, 2011). Pour ce faire, la procédure de correction orthographique était indispensable, bien que laborieuse ; nous l'avons menée sur toutes les lexies non reconnues par le dictionnaire du correcteur hunspell[3], possédant au moins 5 occurrences. Nous avons conscience que ce travail de correction orthographique repose sur un principe de *« destruction créatrice »* de fautes, mais de par la taille du corpus, ces effets négatifs sont minimes au regard de l'analyse. Une suppression des URL a également été exécutée afin d'éviter la présence de termes indésirables (tels « *http* », « *youtube* » ou « *watch* »), voire de *spams*. Ce défi correctionnel a constitué une charge de travail importante en termes de travail humain.

2.3. Corpus corrigé

Les variables du corpus sont les suivantes : *like (le nombre de « *like(s)* » associé au commentaire), *id (le post sur la page), *date (la date à laquelle a été publié le commentaire). Pour le propos de cette communication, nous ne les considérerons pas et en resterons aux distributions lexicales. L'analyse statistique globale après nettoyage nous révèle les caractéristiques suivantes :

nombre de textes :	132279
nombre d'occurrences :	3958482
nombre de formes :	45174
moyenne d'occurrences par forme :	87.63
nombre d'hapax :	25259 (0.64% des occurrences – 55.91% des formes)
moyenne d'occurrence par texte :	29.93

Tableau 2. Statistiques générales (corpus lemmatisé)

La réalisation d'un nuage de mots nous permet au préalable d'observer les formes lexicales pleines les plus utilisées du corpus :

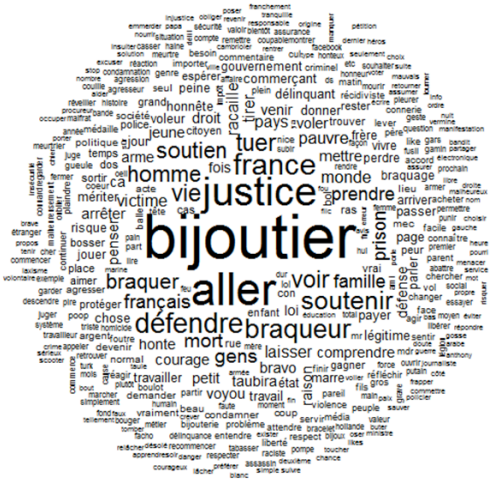
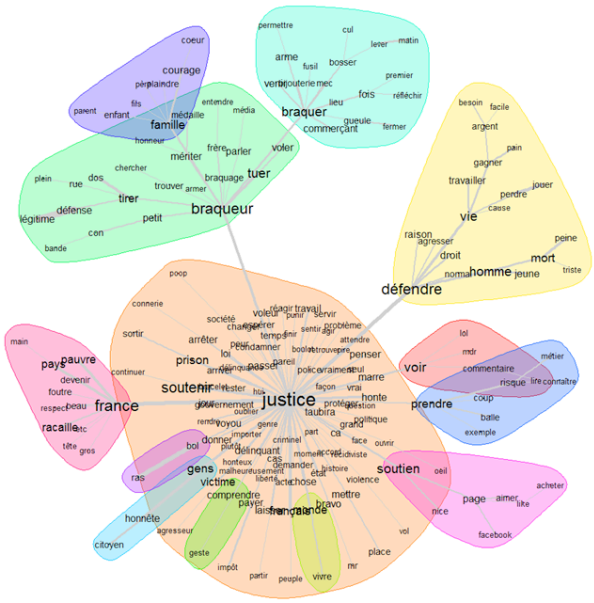


Figure 1. Nuage de mots des 400 formes pleines principales

[3] http://hunspell.sourceforge.net

Lexique 3, Université Paris V. *La documentation officielle de Lexique 3*, New B., Pallier C., Ferrand L., 2005

L'analyse de similitude d'une matrice textuelle offre, par rapport au nuage de mots, une perception globale des rapports des formes entres-elles[1] : proximités et distances des items sont mises en évidence.


famille
braquer
braqueur
tuer
tirer
voler
parler
défendre
vie
homme
mort
voir
prendre
justice
prison
soutenir
pauvre
pays
france
racaille
gens
victime
soutien
penser
marre
honte
mettre
comprendre
payer
bravo
place
page
facebook
commentaire
honnête
citoyen
agresseur
geste
impôt
partir
peuple
vivre
rnr
vol
nice
aimer
like
achreter
oeil
ras
donner
delinquant
criminel
cas
demander
histoire
violence
chose
état
taubira
grand
politique
face
outrir
part
ca
importer
oublier
genie
voyou
rendre
respect
beau
tête
gros
devenir
foutre
continuer
sortir
main
connerie
poop
société
voleur
punir
servir
sentir
agir
attendre
trouver
pareil
police
aimer
seul
façon
vrai
coup
batile
exemple
métier
risque
lire
contraitre
lol
mdr
besoin
facile
argent
pain
gagner
travailler
perdre
jouer
cause
raison
agresser
droit
normal
jeune
triste
peine
arrêter
loi
condamner
boulot
arriver
passer
rester
hus
gouvernement
oublier
imperter
plaid
bol
honteux
malheureusement
liberté
acte
mérite
frère
braquage
chercher
trouver
armer
plein
rue
dos
défense
légitime
petit
con
bande
coeur
courage
périplaindre
fils
parent
enfant
médaille
entendre
média
honneur
arme
fusil
bosser
premier
verbe
fouterie
mec
fois
réfléchir
lieu
commerçant
queule
fermer
permettre
cul
lever
matin
travail
servir
problème
attendre
penser
marre
honte
mettre
bravo
place
violence
soutien
page
facebook
commentaire
voir
prendre
justice
victime
gens
prison
soutenir
pauvre
pays
france
racaille
honnête
ras
comprendre
payer

3.1. Profil des classes

3.1.1. Profil de la classe 9

La classe 9 (7,2%) concerne la polémique autour du concept de légitime défense. Le thème de la légalité est donc ici fortement utilisé et les internautes usent d'un lexique factuel pour ré-interpréter leur définition personnelle de la légitime défense (et ce, majoritairement au prisme du légal). La classe 9 répond fortement aux soutiens de la classe 8 (la morale s'oppose aux réactions émotionnelles).

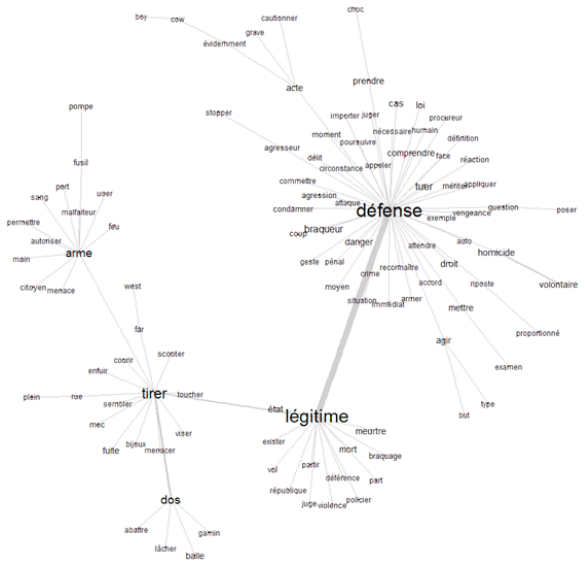


Figure 4. Graphe de la classe 9

Voici quelques segments de texte caractéristiques[6] de la classe 9 : « *N'**importe**-quoi, si on commence à s'entre-**tuer** même en **cas** d'**agression** que va devenir ce monde ? **Tirer** dans le **dos** d'un homme qui **prend** la **fuite** ce n'est ni du courage ni de la **légitime défense**.* », « *Tout le monde oublie qu'il n'y a aucune **légitime défense** puisque sa vie n'était plus en **danger** il a **tiré** dans le **dos** d'un **fuyard** par **vengeance,** on n'est pas au **far west*** », « *Dans ce **cas** sa responsabilité **pénale** n'est pas **retenue** pour les **atteintes** qu'elle a pu elle-même **causer**. En **état** de **légitime défense**, dans le **cas** présent, la **riposte** est **disproportionnée** à l'**attaque** et non simultanée ; de plus, le justicier **tire** dans le **dos** en prime* ».

3.1.2. Profil de la classe 8

Le lexique de cette classe (16,6%) correspond aux soutiens envers le bijoutier. Ici, Stephan Türk et son geste (tirer sur un de ses agresseurs) sont félicités par les internautes. Il s'agit d'un discours dont le champ lexical consiste en la congratulation : *courage, soutien, soutenir, bravo, chapeau, approuver, amitiés, félicitation, admirer, raison, saluer*, etc… Le discours associe au bijoutier des valeurs héroïques, il est qualifié de « *modèle* », un « *symbole* ». Tout comme dans la littérature, cette héroïsation (au sens inchoatif du terme) suppose la possession/détention de certaines caractéristiques et l'adhésion à certaines valeurs perceptibles dans le lexique : *courageux, héros, respect, médaille, admiration, solidarité,*

[6] Tous les exemples de segments de textes caractéristiques cités dans cet article ont été préalablement corrigés afin de faciliter leur lecture. Les formes en gras sont les plus caractéristiques des classes.

sympathie, bravoure, etc…cette classe est homogène ; des caractéristiques sont des adjectifs, des valeurs, des substantifs. Mais ce processus de personnification n'est pas sans complexité, il est à la fois emblématique pour ceux qui l'approuvent, tragique, car coupable aux yeux de la loi (transgression), et problématique (le « héros » pour certains est un « anti-héros » pour d'autres dans la classe 9, incompatible avec la morale collective) : l'acte du bijoutier est ainsi débattu par les énonciateurs des classes, approuvé ou critiqué.

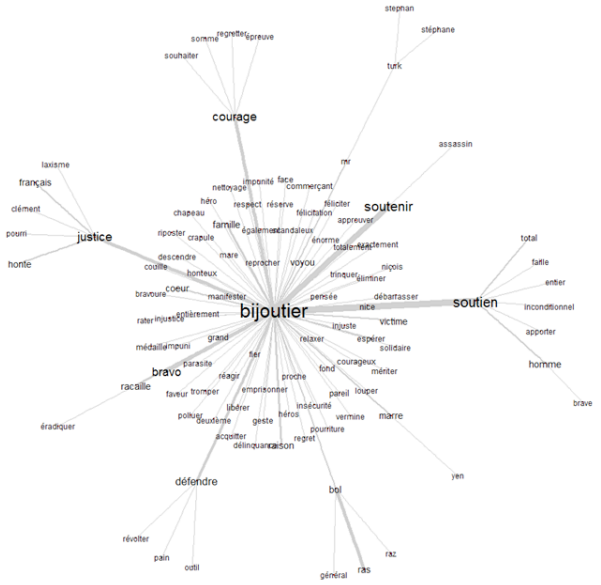


Figure 5. Graphe de la classe 8

Voici quelques segments de texte caractéristiques de la classe 8 : « *Ayant été* ***commerçante*** *et agressée et cambriolée à* ***raison*** *de 7 actes par année sans aucun* ***soutien*** *de la police, je* ***soutiens*** *ce* ***bijoutier****. L'exaspération a des limites,* ***bravo*** *!* », « *Je suis de tout* ***cœur*** *avec le* ***bijoutier*** *de* ***Nice*** *; autour de moi beaucoup de gens le* ***soutiennent****,* ***bravo*** *pour son* ***courage*** *», « Ce* ***bijoutier*** *mérite une* ***médaille*** *pour avoir* ***défendu*** *sa* ***vie*** *et son* ***gagne-pain****.* ***Bravo*** *Monsieur Türk, vous êtes un héros des temps modernes (…) ».*

3.1.3. Profil de la classe 6

Cette classe (6,3%) est formée à partir d'un lexique relatif au problème sociétal de réinsertion des délinquants après une peine de prison. Ce profil est donc à mettre en miroir avec le projet de réforme pénale de la garde des Sceaux Christiane Taubira, en matière de lutte contre la récidive (le but est d'entraîner une diminution des incarcérations, en supprimant les peines plancher et en instaurant des aménagements de peines. La loi prévoierait l'étude, à chaque deux-tiers des peines de prison, d'une « sortie sous contraintes »).

La ministre et son projet de réforme sont vivement critiqués dans cette classe. Le lexique dominant est le suivant : *peine, prison, Taubira, récidiviste, aménagement, violeur, Cayenne, sursis, bagne, relâcher, remise, pédophile, purger, sortir, recommencer*, etc… A partir de l'affaire du bijoutier de Nice, le débat sur la réforme des peines de prison est ainsi relancé en ligne par les citoyens. Pour certains internautes, c'est même la peine de mort qui devrait être réinstaurée en France. Et alors que le sujet est celui du vol à main armée, les viols, la pédophilie ou encore le terrorisme sont de nouvelles thématiques évoquées.

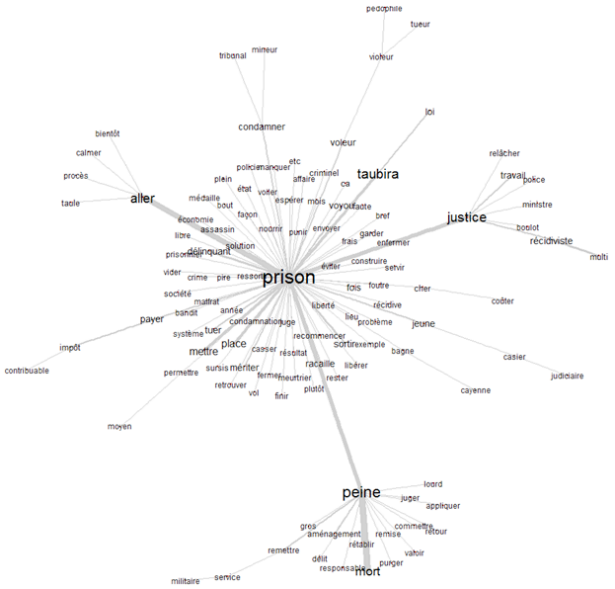


Figure 6. Graphe de la classe 6

Voici certains de ces segments caractéristiques : « *Je vote pour la* ***peine*** *de* ***mort,*** *il y aura plus de* ***place*** *dans les* ***prisons*** *et surtout moins de connards dans les rues. C'est le contraire que veut faire Madame* ***Taubira****, la France sera contaminée par ces vermines, faut les éradiquer* », « *C'est grâce à la* ***justice*** *française que nous sommes obligés de nous faire* ***justice*** *nous-mêmes, si on* ***rétablissait*** *la* ***peine*** *de* ***mort*** *aux* ***violeurs*** *et* ***assassins*** *d'enfants on ferait du ménage en* ***prison*** *et il y aurait à* ***nouveau*** *de la* ***place*** ».

3.1.4. Profil de la classe 1

Cette classe (17,4%) fait état de vives critiques contre l'Etat et la gérance de ses fonctions régaliennes, notamment concernant la justice et la sécurité intérieure. Le gouvernement est attaqué, le Président, le Premier Ministre et les deux ministres de la justice et de l'intérieur sont visés au premier chef. L'image de la France est considérée comme salie, elle est « *honteuse* ».

Figure 7. Graphe de la classe 1

Voici les lemmes les plus significatifs de ces attaques : *France, pays, justice, protéger, gouvernement, citoyen, peur, voyou, force, laxisme, loi, liberté, impôt, égalité, protection, fraternité, punir, gendarmerie, autorité*, etc… Nous nous rendons bien compte que le thème de l'insécurité est largement présent, où le sentiment de peur est exacerbé. Les systèmes sécuritaires et pénaux de la France sont condamnés : ils sont définis comme défaillants, car ils ne protègeraient pas les honnêtes citoyens français et n'effraieraient plus les délinquants. Quelques segments de textes caractéristiques de la classe 1 : « *La **France** est devenue un **pays** où les **honnêtes** gens ont des devoirs, et où les **délinquants** ont des **droits*** », « *Ah oui elle est belle la France ! La **justice** aussi, elle **préserve** les **délinquants** et **condamne** les **gens** qui ne font que se **défendre** alors moi je dis **honte** à la **justice**, **honte** à tous ces **gens** qui **dirigent** ce **pays*** ».

3.1.5. Profil de la classe 7

La classe 7 (représentée pour 10,7%) est fondée à partir d'un lexique où les verbes « *jouer* » et « *perdre* » sont associés (lien de cause à effet) pour expliquer le caractère fataliste de la situation. Le fait de braquer est perçu comme un « *jeu* », un « *métier* » à « *risques* ». Voici certaines formes pleines du profil : *perdre, jouer, risquer, métier, mort, assumer, triste, conséquence, cowboy, compassion*, etc… Alors que la responsabilité du braqueur dans sa mort est clairement présentée (« *bien fait pour sa gueule* »), certains des internautes adressent également, dans cette classe, leurs condoléances à la famille de la victime.

Figure 8. Graphe de la classe 7

Voici quelques segments de texte caractéristiques de la classe 7 : « *Quand on veut **jouer** les **braqueur**s on prend des **risques** et on les assume. Il a **joué**, il a **perdu**. Tant mieux. », «La **mort** d'un jeune homme de 19 ans est **triste**, **condoléances** à ses parents, mais malheureusement quand on **braque** une **bijouterie** avec une **arme** il faut s'attendre au **danger** ».*

3.1.6. Profil de la classe 2

La classe 2 (6,6%) est celle de la France « *travailleuse* », « *qui se lève tôt* » (*confer* un héritage des discours de Nicolas Sarkozy), et contre toute forme d'assistanat. La mise en scène du récit de vie du bijoutier moyen, son aspect « routinier » crée une forte identification des internautes.

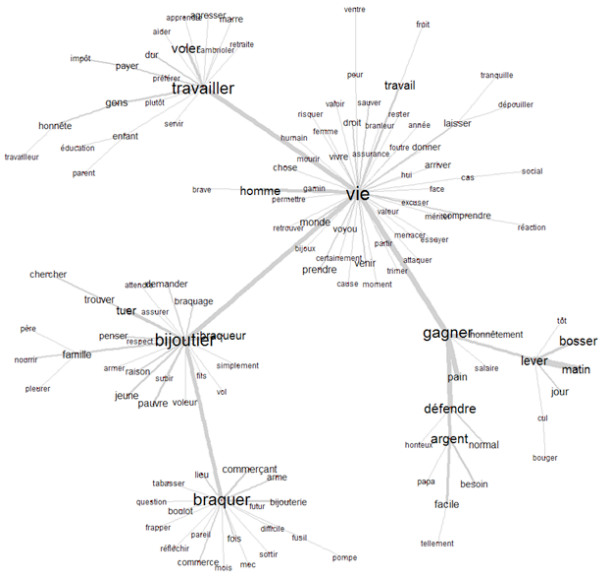


Figure 9. Graphe de la classe 2

Voici son lexique : *lever, matin, travailler, dur, gagner, vie, bosser, argent, pain, nourrir, impôt, payer, labeur, tôt, subvenir, salaire, retraite, pognon, charge, honnêtement, smic*, etc… Quelques segments de texte caractéristiques de la classe : « *Et oui il a* ***ôté*** *la* ***vie*** *à de pauvres* ***jeunes*** *cas* ***sociaux*** *qui* ***venaient*** *le* ***braquer*** *car lui* ***travaille*** *honnêtement et c'est pas* ***normal*** *de se* ***lever*** *le* ***matin*** *pour* ***gagner*** *sa* ***vie,*** *les* ***pauvres*** *gamins qui ne* ***demandaient*** *qu'à* ***vivre*** *sur le dos des honnêtes* ***gens*** », « *Je soutiens ce* ***père*** *de* ***famille*** *qui a voulu protéger son* ***gagne****-pain je suis moi-même* ***commerçante*** *chaque* ***jour*** *je me* ***lève*** *pour aller soulever mon rideau je* ***gagne*** *ma* ***vie*** *en* ***travaillant*** *durement* ».

3.1.7. Profil de la classe 3

Cette classe (12,8%) concerne des critiques à l'encontre de la famille de la victime. Les internautes mettent en avant le problème de l'éducation des jeunes, et des valeurs transmises. La responsabilité du voleur est associée au fait qu'il allait devenir père et donc avoir des responsabilités de parents. Les témoignages du frère de la victime dans les médias sont aussi condamnés. Voici quelques segments caractéristiques de la classe *: « Le* ***frère*** *de la* ***victime*** *qui dit qu'il ne comprenait pas pourquoi le* ***bijoutier*** *a* ***tiré*** *sur un* ***jeune père*** *de* ***famille****, nous ce qu'on ne comprend pas c'est pourquoi ton* ***frère*** *qui a 20 ans et est* ***père*** *de* ***famille*** *va* ***braquer*** *une* ***bijouterie*** », « ***Témoignage*** *étonnant du* ***frère*** *du* ***braqueur****, lui-même* ***commerçant*** *: cela dit en* ***passant,*** *il savait qu'il* ***allait*** *être* ***papa*** *il avait* ***besoin*** *d'****argent*** *alors il a* ***braqué*** *une* ***bijouterie****…bah oui c'est* ***normal*** », « *Et le* ***frère*** *de la* ***victime****,* ***braqueur****,* ***voleur****, qui ose dire qu'il y a des* ***assurances*** *pour les* ***bijoux*** *! Non mais quelle* ***mentalité*** *! Allez, allons tous* ***braquer*** *des* ***magasins****, ils sont* ***assurés*** *après tout.* »

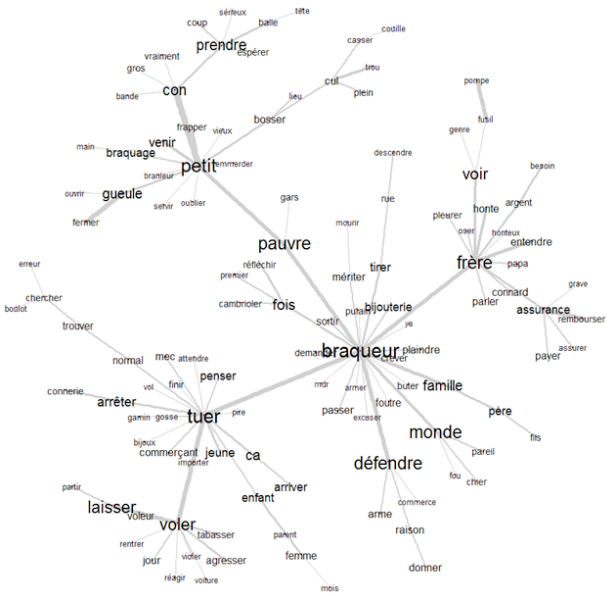


Figure 10. Graphe de la classe 3

3.1.8. Profil de la classe 4

Le lexique de la classe 4 (12,3%) révèle la présence d'un débat interne autour de l'authenticité et de la validité des « *likes* » sur la page Facebook, et qui fait écho à la polémique largement médiatisée (*cf.* nuage de mots figure 6). Le lexique dominant comprend donc les lemmes : *page, Facebook, like, aimer, acheter, front national, signer, médias, faux, manipulation ami, cliquer, statistique, robot, truquer, arnaque, fake,* etc…

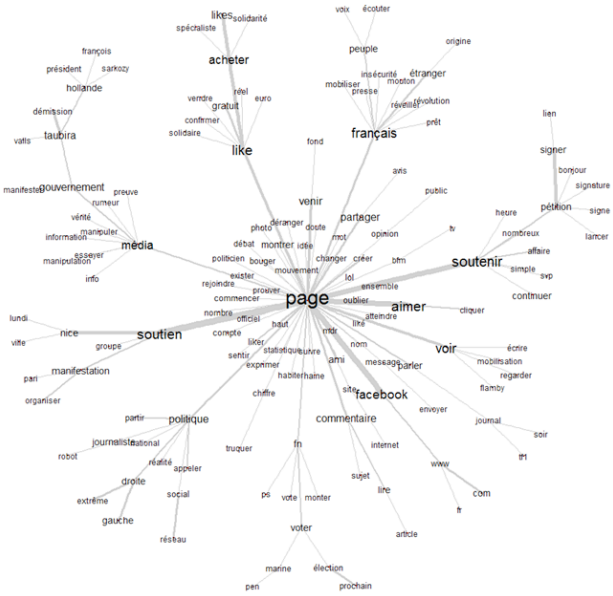


Figure 11. Graphe de la classe 4

Les segments de texte caractéristiques de la classe nous permettent de mettre au jour la controverse : « *J'ai 122* ***amis*** *sur* ***Facebook*** *et 19 ont* ***liké*** *la* ***page****, ils n'ont pas été* ***payés*** *et leur* ***vote*** *est bien réel. Il faut accepter la* ***vérité****.* », « *Le problème c'est que la suspicion sur le* ***nombre*** *de* ***likes*** *sur votre* page *pose* ***problème*** *; en effet, vous ne seriez que 250 000, le reste viendrait de* ***pays étrangers*** *ou de* ***likes achetés*** », *« Notre enquêtrice est allée fouiner sur un site fort intéressant :* ***Socialbakers****, qui détaille la provenance des* ***likes*** *sur des* ***sites****.* »

3.1.9. Profil de la classe 5

Les formes pleines possédant le plus d'occurrences de la classe 5 (10%) mettent au jour une polémique fondée autour du racisme : *facho, raciste, arabe, blanc, musulman, noir, bougnoule, libanais, nazi, heiss, juif, maghrébin, couleur, gitan, racaille,* etc…

Figure 12. Graphe de la classe 5

Les internautes se répondent en s'interpellant via leur pseudo Facebook : *Samy, Axel, Laura, Luigi, Pascal, Hakim, Anthony*, etc… Ici, nous avons affaire à la classe où l'interaction verbale entre les internautes est la plus manifeste. Aussi, nous pourrions avancer l'idée selon laquelle il y a hybridation dans l'utilisation du réseau social exploité comme lieu de forum de discussion, hybridation admise notamment par les nouvelles fonctionnalités de Facebook qui permettent de répondre à des commentaires. S'il existe une configuration discursive propre aux forums, voire une intertextualité[7], il serait pertinent d'étudier une telle présence au sein de cette classe : interpellation, présence de smileys, réponses par liens, jugements appréciatifs ou dépréciatifs, etc… Autre paramètre propre aux forums, les jugements par l'orthographe, où les internautes faisant des fautes de Français sont stigmatisés. Quelque phrases d'exemples : « *Pour certains , il est grand temps de* ***retourner*** *à l'****école*** *! Car niveau* ***orthographe****, y'a du travail ! Et niveau* ***éducation*** *aussi ! Ah oui c'est vrai, vous défendez ce pauvre braqueur ! Eh bien, on va avancer avec des gens comme vous !* », « *Toutes ces* ***fautes*** *d'****orthographe*** *me donnent mal à la tête. Pensez à prendre des* ***cours*** *plutôt que de vous occuper des autres,* ***bande*** *de* ***cas sociaux*** *!* », « *Tu noteras* ***fanny marie****, leur* ***orthographe****. Ca en dit long sur l'atrophie de leur* ***cerveau*** *!* » Comme l'estime (Breton, 2001), la fracture numérique peut s'avérer sociale (et non seulement géographique), puisque les internautes qui s'expriment mal ont tendance à être exclus de certains forums de discussion en ligne. Des modérateurs ferment des « *topics* » pour cause d'« *incompréhensibilité* ». Voici un autre segment de texte qui dénonce l'orthographe défaillante d'internautes : « *Message pour les* ***racistes*** *et nationalistes extrémistes : si vous aimez votre pays, apprenez déjà à* ***écrire français****, une fois l'****orthographe*** *maîtrisée vous serez un poil plus crédibles* ***bande*** *de* ***fachos*** *de* ***merde*** *que vous*

[7] Un « système de règles implicites qui sous-tendent cet intertexte, le mode de citation qui est jugé légitime dans la formation, le type, ou le genre de discours dont relève ce corpus » (Maingueneau, 1984:83).

êtes ». Une autre tendance est à repérer, puisque le lexique injurieux est ici particulièrement prégnant, et ce, autant du côté des défenseurs du bijoutier, que des opposants à la page : *facho, pute, salope, baiser, insulter, sale, dégager, bobo, niquer, bouffon, conasse, bâtard*, etc… Ces insultes sont majoritairement scabreuses, le champ lexical du sexe est remarquable. Patrice Flichy[8] (2008) perçoit dans cette utilisation de l'invective une tendance qu'il dénomme « *flame wars* » qu'il traduit en « *guerre d'injures* » (ou phénomène des « *haters* »). Flichy s'appuie sur les travaux de Mark Poster[9] (1997) pour avancer la thèse selon laquelle « *le débat* [en ligne] *ne tend pas vers l'élaboration d'une position commune, mais plutôt vers une multiplication de points de vue contradictoires* ». Pour Amossy (2011), les interactions polémiques dans les forums de discussion[10] constituent *« une* agora *imaginaire […] d'un genre très particulier, puisque fondés sur l'antagonisme et la violence verbale »,* où *« l'exacerbation de la confrontation agonique entre internautes joue un rôle dans la construction d'un groupe interne au sein duquel les membres peuvent se conforter mutuellement dans leur combat commun »*. Voici quelques segments de texte caractéristiques de la classe 5 : « *Je suis choquée par les* ***propos*** *injurieux,* ***racistes****,* ***fachos*** *et j'en passe que je lis sur cette* ***page****, et des deux côtés en plus, c'est lamentable* », « ***sale arabe*** *de* ***merde****, tous des* ***fils*** *de* ***pute****, que les* ***arabes crèvent*** *tous, il faut jeter des bombes nucléaires en Afrique comme ça ils ne viendront plus nous pourrir la* ***vie*** !», « *Sarah Maria Francesca pour ta gouverne, le voleur était* ***français*** *et le bijoutier* ***arabe*** *donc informe-toi avant de parler au lieu d'écrire des* ***propos racistes*** », « *Ta* ***gueule*** *Roger va* ***sucer*** *des* ***bites****, tu t'es pris pour qui là ? Un magistrat pour juger les gens, ou un politicien ? Va te laver le* ***cul****.* »

3.2. Analyse factorielle de correspondances

L'analyse factorielle de correspondances positionne ici les formes lexicales pleines à partir de deux facteurs.

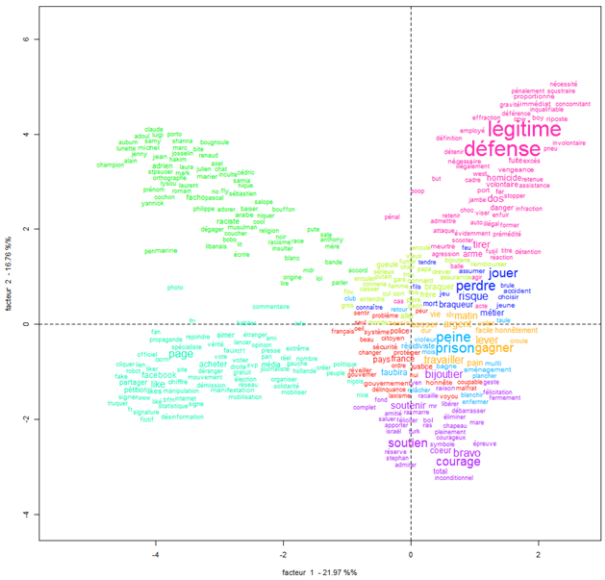


Figure 13. AFC du corpus

[8] Flichy, 2008.

[9] Poster, 1997.

[10] *Ibid*, 2011.

Nous pouvons dès lors constater que si la page Facebook s'intitule bien « soutien au bijoutier de Nice », les thématiques produites par les énonciateurs ne sont pas exclusivement propres au soutien. La page, transformée en véritable *agora* politique, est le théâtre de polémiques diverses dont l'affaire du bijoutier n'est que l'élément déclencheur. L'analyse classificatoire permet de pointer cette convergence des interdiscours, difficilement mesurable dans le cadre d'un très gros corpus comme le nôtre.

Figure 14. Interprétation des facteurs des axes diagonaux

Nous pourrions interpréter cette AFC à partir de ses axes diagonaux, afin de mieux cerner les caractéristiques et singularités des lexiques de classes. Les émotions positives en soutien au bijoutier s'opposent aux émotions négatives où une « *guerre d'injures* » fait rage. Et, dans un autre axe diagonal, l'affaire est débattue sur les faits (controverse de fond), autant que sur la forme (anti-controverse). Médias et réseau social s'auto-alimentent dans la production des informations. Nous constatons grâce à la classification de ce gros corpus hétérogène, que l'expression citoyenne en ligne tient plus à un échange d'opinions qu'à un débat d'idées.

A partir de cette étude, nous pouvons nous interroger autour du degrès d'engagement politique des internautes, alors même que la manifestation physique dans les rues de Nice en soutien au bijoutier, n'a mobilisé que 800 à 1 000 personnes. Alors, à quel type de pratique avons-nous affaire ? Quels sont ses aspects communautaires, sont-il propres aux réseaux sociaux ? Et, si communauté il y a, est-elle réellement consciente ?

Références

Amossy R. (2011). *La coexistence dans le dissensus*, *Semen*, n°31.

Amossy R. et Marcel B. (2011). *Introduction : la polémique médiatisée*, *Semen* n°31.

Breton P. (2000). *Le culte d'Internet, une menace pour le lien social* ? Paris, La Découverte.

Flichy P. (2008). *Internet et le débat démocratique*, *Réseaux* 4/2008 (n° 150), p. 159-185.

Garand D. (1998). *Propositions méthodologiques pour l'étude du polémique*, Hayward Hayward A. and Garand D. (1998). *États du polémique*, Montréal, Nota Bene, 211-268.

Maingueneau D. (1984). *Genèses du discours*, éditions Mardaga.

Poster M. (1997). A chapter *Cyberdemocracy : The Internet and the Public Sphere,* in a book Holmes D. (ed), *Virtual Politics, Identity and Community in Cyberspace*, Sage, Londres, 1997, pp. 212-228.

Ratinaud P. et Marchand P. (2012). *Recherche improbable d'une homogène diversité : le débat sur l'identité nationale*. Langages, 187, 93-107.

Ratinaud P. et Marchand, P. (2012*). L'identité nationale : l'alibi du débat. Le Discours et la langue*. Vol 3. n1, 175 et sq.

Reinert M. (1983). *Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte*, *Les cahiers de l'analyse des données,* Vol VIII, n° 2, p 187-198.