SM+S
social media + society

Article

Social Media + Society
January-March 2017: 1–13

DOI: 10.1177/2056305117693647
journals.sagepub.com/home/sms
SAGE

The Charlie Hebdo Attacks on Twitter: A Comparative Analysis of a Political Controversy in English and French

Nikos Smyrnaios[1] and Pierre Ratinaud[2]

Abstract

In this article, we propose an original method combining large-scale network and lexicometric analysis to link identifiable communities of Twitter users with the main discursive themes they used in the aftermath of the Charlie Hebdo attacks in Paris, France in 2015. We used this method to compare tweets and user networks in French and in English. We observed that the majority of the users who tweeted about Charlie Hebdo were people without any particular affiliation, who were shocked by the attacks and immediately expressed themselves through emotionally charged messages. But rather quickly their proportion decreased and they participated less in politically polarizing discussions. On the other hand, we found that smaller, highly politicized, and polarized groups had similar attitudes toward the events: they were less engaged immediately after the attacks in emotional expression of sympathy and shock, but they participated vividly in the following days in polemical discussions or engaged themes. Other findings include the central position of mainstream media and the existence of groups of users that aggregated on the basis of nationality. More generally, our results show clearly that even the most dramatic events such as a terrorist attack with innocent victims do not produce homogeneous reactions online. Rather, political engagement and cultural dispositions are keys to understand different attitudes on Twitter.

Keywords

Twitter, Charlie Hebdo, network analysis, discourse analysis, web sphere, controversy

Introduction

For several years now, the Internet has become a privileged field for controversies, that is, the process during which players involved develop "arguments and conflicting views that lead them to offer different versions of the social and the natural world" (Callon, 1986, p. 175). Controversies are particularly numerous in the political field, where constructed and discussed *public issues* are defined as "problems that are subject to treatment in any form whatsoever on the part of public authorities and thus involve decision-making" (Garraud, 1990, p. 27). In such a context, a multitude of players can take part in the controversy: authorities, media, and various organizations (associations, lobbies, political parties) and also ordinary citizens with no specific affiliation. While there is a large literature on online controversies in the study of science, technology, and society (STS; Marres, 2015), little empirical research has been conducted on political controversies that are triggered online by unpredictable events such as terrorist attacks.[1] This is probably because such dramatic events are considered to produce largely consensual reactions. But is this always the case?

Regarding the attacks perpetrated against the Charlie Hebdo newspaper and a kosher grocery in Paris, in January 2015, this is only partly true. On 7 January 2015, around 11.30 a.m., two individuals murdered a part of the Charlie Hebdo satirical newspaper team. The next day, another individual murdered a policeman near Paris. On 9 January, the Kouachi brothers, responsible for the Charlie Hebdo attacks, took hostages and locked themselves in a company building in Dammartin-en-Goële, a town 50 km northeast of Paris. The same day, Amedy Coulibaly, responsible for the policeman's murder the previous day, hid in the Hyper Kasher, a kosher grocery located in eastern Paris, and shot

[1]University of Toulouse 3, France
[2]University of Toulouse 2, France

Corresponding Author:
Nikos Smyrnaios, Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales (LERASS), University Toulouse 3, 115D route de Narbonne, BP 67701, F-31077 Toulouse Cedex 4, France.
Email: smyrnaios@free.fr

four customers. Both hostage takings finished almost simultaneously around 5.00 p.m., when the police stormed the premises and killed the terrorists. These attacks, known as the "Charlie Hebdo attacks," made a total of 17 victims and had a very strong impact in France and throughout the world. On 11 January, the French government organized a "Republican march," attended by 44 heads of foreign states. It led to the largest public gathering ever known in France. The bloody attack on Charlie Hebdo and the dramatic events that followed sparked a huge wave of reactions in France and in the world. As is the rule now, these reactions quickly flooded social networking sites and Twitter in particular. Starting from 7 January, millions of tweets were published using hashtags such as #CharlieHebdo and #JesuisCharlie.

The majority of journalists, politicians, and pundits in the French and international media, as expected, condemned the attacks, expressed solidarity with the victims and pressured authorities to find and punish those responsible. But when it came to defining the causes, controversy arose. Some on the right of the political spectrum pointed out that Islam is by "nature" a violent or rearward religion, while others insisted on the "provocations" of Charlie Hebdo against Muslims and denounced the rising Islamophobia in France. At the same time, a debate took place concerning the appropriate response by authorities opposing those that backed the "national union" strategy of the French government and the presence of worldwide political leaders in Paris for a march against terrorism, while others criticized this stance as hypocritical and inefficient. Finally, some evoked the unstable geopolitical context of the Middle East and the tumultuous history of the region as being the deeper roots of the problem. Thus, the central question of this research is the following: "Were these controversies present in online reactions to the Charlie Hebdo attacks, and if yes, in what manner were they expressed and by whom?" To answer that question, we propose an original method to link identifiable communities of Twitter users with the main discursive themes they used. First, we operate a large-scale network analysis upon a sample of tweets related to the attack that allows us to detect highly linked communities of users. Then, we operate an automated discourse analysis based on a lexicometric method. Finally, we combine the two to link the communities of users to the lexical clusters that they used the most in French and in English.

Theoretical Framework

The peculiarity of debates that emerge out of events, such as the Charlie Hebdo attacks, when they take place online, as opposed to other institutionalized spaces, is the difficulty in defining its temporal, spatial, and sociopolitical frontiers and oppositions. The dynamics and extent of public reactions are unpredictable. The course of the discussions, the kind of discourse that develops, the interactions that take place between groups and individuals, and the ideas and representations that are expressed depend on a multitude of factors that are difficult to predict a priori and a fortiori in different cultural and linguistic environments. Political controversies that occur online and worldwide are far from presenting the characteristics of a normative public space. That is why, in order to grasp their complexity, we use a particular theoretical and methodological framework that we propose to apply on the case of the Charlie Hebdo attacks.

In this work, we consider online political controversies as *web spheres*. This concept is both a theoretical proposal and a methodological tool. A web sphere is defined by Schneider and Foot (2006) as a set of freely accessible digital resources spread over different web pages or servers interconnected by links, which refer to a specific event or theme. It is therefore a micro-public sphere circumscribed by both a thematic focus and time limits. This thematic orientation can be diverse in nature and more or less well defined (a news story, a social or political problem, a scientific or religious controversy etc.), but it always involves a public issue as it was defined earlier. A web sphere is limited in time but its lifespan is not known in advance. It depends on the ad hoc engagement of participants, which in turn is variable and can be influenced by many factors (e.g., an unforeseen event can revive a controversy that seemed off). One of these factors in the case of the Charlie Hebdo attacks is the emergence of a *nexus*. Nexuses are "affective prelogical knots shared by a large number of individuals" (Rouquette & Moscovici, 1994, p. 68) which, at a specific time, are unquestionable and command profound mobilization of crowds and clear-cut stands. They become visible in conflict, threat, or crisis situations and produce homogeneous and uniform public discourses. Following this definition, the #JesuisCharlie hashtag and the "Republican march" in the streets of Paris following the attacks seem to be part of such a nexus. Nevertheless, as recent research shows, public reactions on Twitter were much more diverse and controversial (An, Kwak, Mejova, Alonso Saenz De Oger, & Gomez Fortes, 2016; Badouard, 2016; Giglietto & Lee, 2015). Other hashtags that emerged at the time, such as #JesuisAhmed and #JenesuispasCharlie, produced critical and even deviant discourses that aimed specifically in countering unanimous solidarity to Charlie Hebdo and were strongly influenced by particular social and cultural contexts.

Indeed, a web sphere can be entered by spontaneous or organized participants, diverted from its original objective of framework (e.g., producing clear-cut and unanimous stands in the case of a nexus), used like a "refraction chamber" (Rieder, 2012) to talk about something else, but nevertheless still maintain a certain temporary and thematic coherence. The objective of this work is precisely to map the different controversial public reactions on Twitter on the Charlie Hebdo attacks and locate their main actors and discursive frameworks based on a more "neutral" hashtag (#CharlieHebdo), a priori less emotionally charged than #JesuisCharlie and not outright polemical like #JenesuispasCharlie.

This can be done based on another particular characteristic of web spheres which is the fact that they leave digital traces that can be collected and processed asynchronously. Such a method allows for a long view on online social interactions and discourses that is difficult to achieve with traditional methods of social sciences (ethnography, sociological survey etc.). Therefore, in this research, we carry out an "empiricist implementation of controversy analysis" that uses digital methods to tell "what the issues of contestation are, who the actors are, and where they are based" (Marres, 2015, p. 663). Of course, the implementation of digital methods to analyze a web sphere resulting from a political controversy also raises a series of important epistemological and methodological problems such as the lack of transparency resulting from technological black-boxing (Rieder & Röhle, 2012).[2] But one of the great advantages of the web spheres approach is that it takes into account digital resources and links, that is, *texts* and *interactions*, and this in a dynamic time frame. Through these three dimensions (texts, interactions, temporality), it is possible to monitor and analyze controversies and political debates that take place online (in this case on Twitter) with enough perspective to capture the overall picture but finely enough to avoid over-interpretation.

Method

In this article, we propose an analysis of a sample of tweets upon which we apply an innovative method (Smyrnaios & Ratinaud, 2014). We implement a protocol for identifying the three components of a web sphere: (1) communities of users within Twitter based on retweets and mentions (links) about Charlie Hebdo, (2) discursive themes mobilized by each of these communities (texts), and (3) their evolution over time. The first step to identify communities of users is to apply a network analysis on our sample. Indeed, a group of Twitter users can be analyzed as a network where profiles are connected by interactions such as retweets and mentions. Retweets are equivalent to citations of other users' messages and mentions of users' aim to talk *to* or *about* them. Research has shown that the "influence" of Twitter users, that is to say the scale of their tweets' impact upon other users, depends less on the number of their followers than on the number of retweets and mentions they receive (Cha, Haddadi, Benevenuto, & Gummadi, 2010) but that the network structures for retweets and mentions can be different (Conover et al. 2011). Indeed, users may retweet and mention other users for different reasons (basically retweets are a form of confirmation whereas mentions are a form of discussion). In our study, we decided to include both retweets and mentions in the network analysis because our objective is to understand precisely how both of these forms of interaction that occur simultaneously structure online debate. Indeed, in the study of an online political controversy, it is interesting to observe the structure of the network created by users citing each other but also conversing with one another as both of these forms of interaction are deeply interrelated. Thus, the topology of this kind of network results from the intensity of interactions between Twitter profiles including both mentions and retweets. In that way, groups of users that interact often are close to one another and create clusters that represent communities inside which the most mentioned and retweeted profiles are central nodes.

The second step of our method is to operate an automated discourse analysis based on the Reinert lexicometric method. The method proposed by Reinert (1990) was first implemented in the ALCESTE software and has been recently integrated in an open-source software called IRaMuTeQ that allows for new developments (Ratinaud, 2014). IRaMuTeQ breaks down texts into smaller parts called *segments* based on punctuation and on size criteria[3] and then classifies them into clusters based on lexical similarity. This second phase of the analysis uses the top-down hierarchical classification algorithm described by Reinert (1983), operating a series of bi-partitions of the corpus based on a correspondence analysis (CA). The first step of this phase is to lemmatize vocabulary and to build a presence/absence matrix (1/0) that crosses tweets (in rows) and full forms (in columns). This matrix undergoes a series of bi-partitions based on a CA. The algorithm goes by always taking as input the biggest of the generated clusters, until the number of clusters set by the user is reached. As in the case of Topic Models (Zhao et al., 2011), the determination of this parameter remains difficult and it results from a back-and-forth process between analysis and interpretation. In a last phase, small clusters are eliminated. The size of the eliminated clusters is also a parameter that depends on the user. In the analysis presented, we have only kept clusters with at least 4500 tweets that are significant given the size of the sample. In summary, our analysis proceeds to a grouping of textual units on a lexical co-occurrence criterion. The clusters generated by the lexical classification are sets of tweets that tend to contain the same words. In this manner, we can obtain the main discursive frames and themes that emerged on Twitter around Charlie Hebdo. Finally, the third step of our method is to link communities of users to the lexical clusters that they used the most. To do that, we use the Chi Square statistic that compares the tallies or counts of categorical responses between two (or more) independent groups, in this case communities of users and lexical clusters.

Dataset

Several difficulties arise in sampling Twitter on a subject such as the Charlie Hebdo attacks: first, the extraordinary volume of messages produced on the subject; second, the linguistic diversity of messages due to the global impact of the event; and finally, the very high density of interactions between users, especially around the worldwide popular hashtag #JesuisCharlie. Therefore, rather than working on #JesuisCharlie, characterized by an emotionally charged and fairly homogeneous discourse centered on condolences and expressions of solidarity or polemical hashtags such as #JesuisAhmed and #JenesuispasCharlie, we decided to analyze a more diverse sample of tweets to shed light on the plurality of debates and interactions that took place. Thus, we collected all messages containing the string *charliehebdo*, which was the first and most commonly used on the subject, between 7 January 1.00 p.m. and 12 January midnight, after the march that took place in Paris, with the DMI-TCAT software.[4] This sampling process produced a set of 3.66 million tweets in different languages. Given the method used to retrieve the tweets, it seems likely that the sample is not totally comprehensive. While this is an important limit of the study, we estimate that, given the volume of the sample, our results are still at least partially valid.[5] Chart 1 shows the proportion of the different languages represented in this sample.

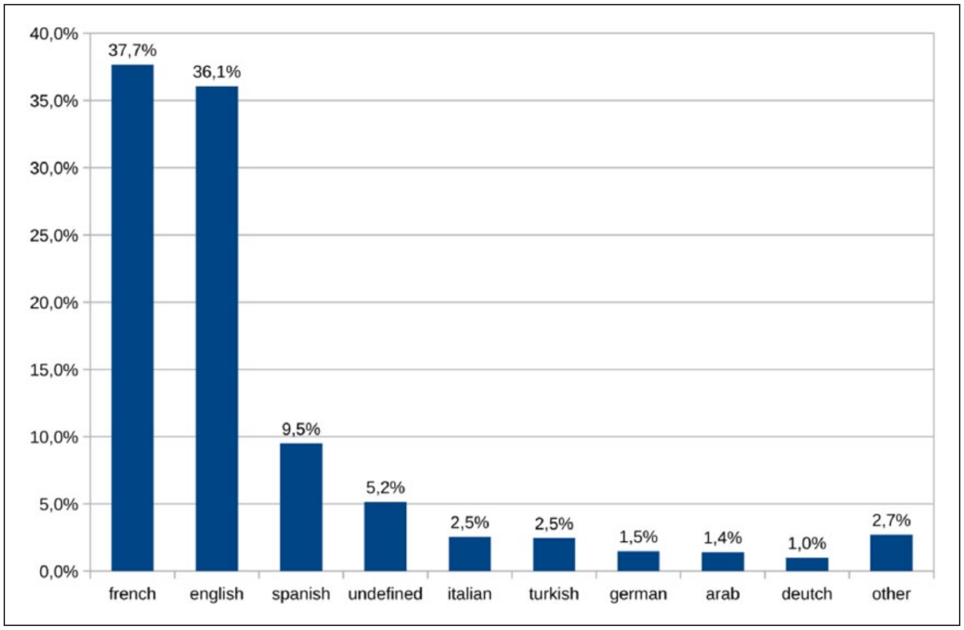


Chart 1. Proportion of main languages in the sample.

Not surprisingly, we found that French is overrepresented in this sample (37.7% of the tweets). In normal circumstances, French represents only 2% of messages.[6] Furthermore, a very important part of the messages in each language consists of retweets, as shown in Chart 2, indicating highly structured exchanges around few accounts that concentrate a large number of citations.

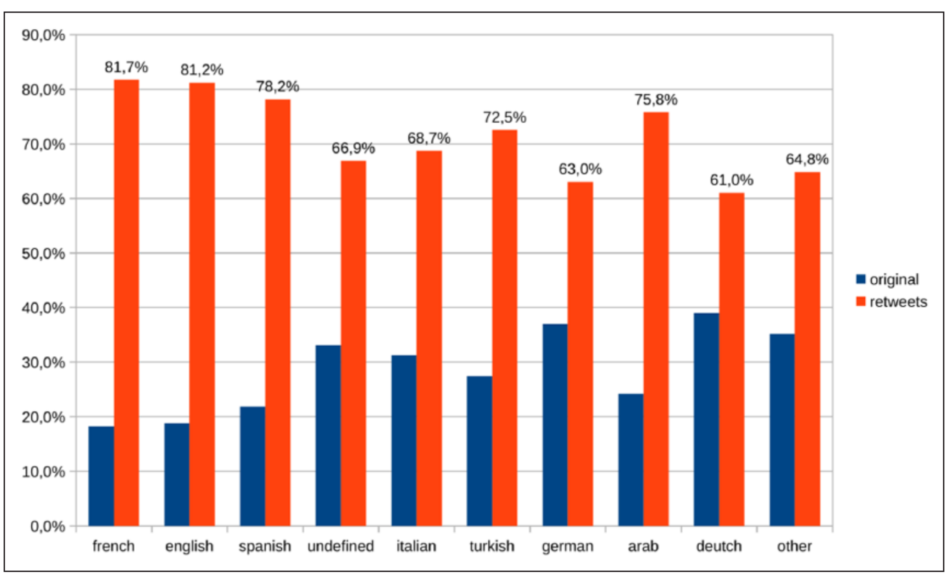
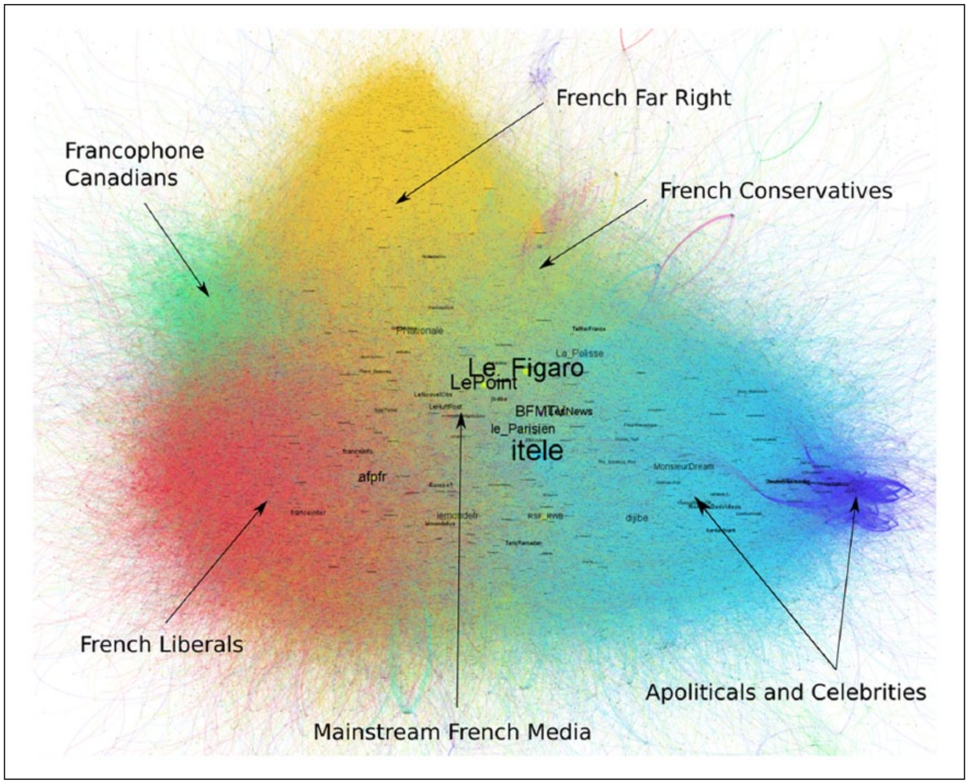


Chart 2. Proportion of tweets and retweets in the main languages of the sample.

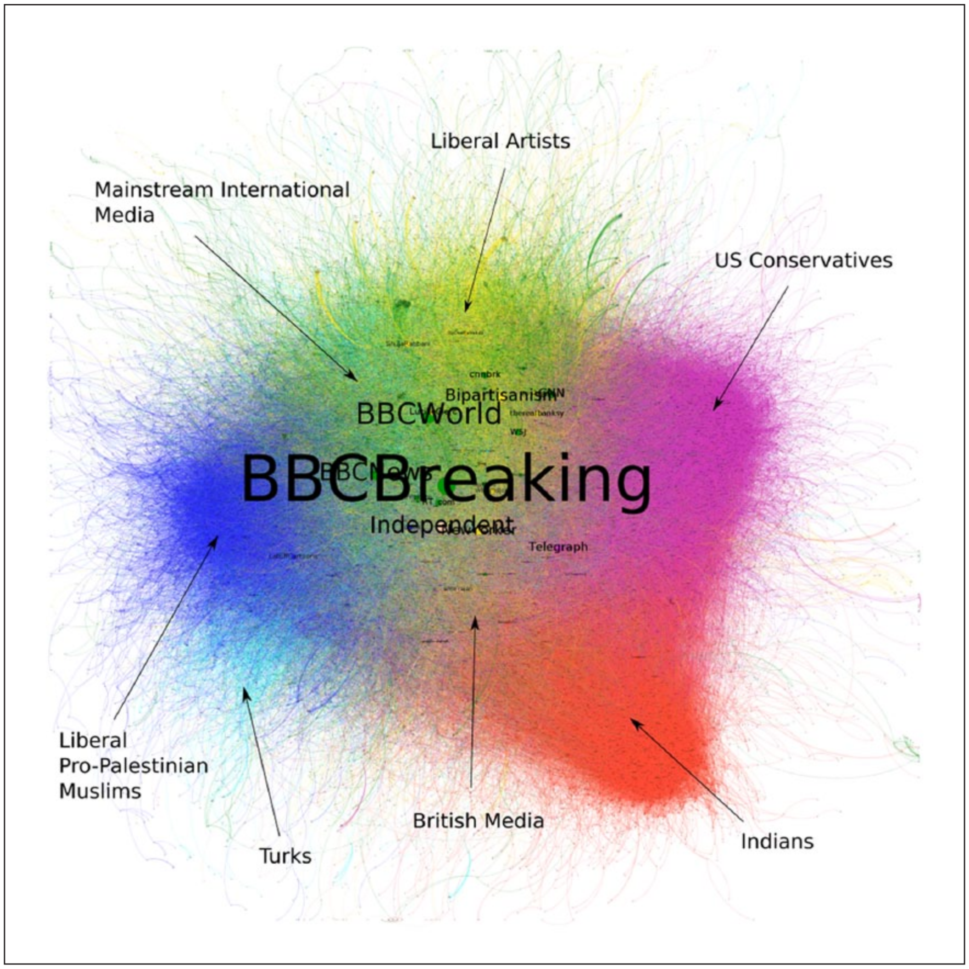


Graph 1. Network based on mentions and retweets of users having produced at least 10 tweets in French including the string *charliehebdo*.

The French sample consists of 1.38 million tweets, while the English one consists of 1.3 million tweets. To avoid the bias of repetition through retweeting, we conducted the lexical analysis only upon the corpus of original tweets, thus, excluding retweets but maintaining replies which can be an important factor for the lexical diversity of the sample (252,194 original tweets in French and 248,500 original tweets in English). Similarly, in order to concentrate on users with significant interest for Charlie Hebdo for the network analysis, we selected only those with at least 10 tweets and retweets containing the string *charliehebdo*. Using a minimum of 10 tweets as a starting point is a way to limit the size of graphs for practical reasons. Nevertheless, the final graph contains not only these users but also all those the former retweeted or mentioned throughout the sampling period. Thus, the total number of users included in the graphs grows exponentially when increasing the starting sample to users with less than 10 tweets, without necessarily substantially altering the result of the analysis. Indeed, small contributors are often assigned to very small communities that are not central to our analysis. The final graph of the French network obtained on the basis of these criteria is composed of 53,196 distinct users and 375,800 links, while the English one of 40,441 users connected by 239,474 links. The French network is nearly two times as dense as the English one ($3.76{\cdot}10^{-7}$ vs $7.05{\cdot}10^{-7}$). This is probably due to the fact that the English network composed users from all over the world and the French network mainly composed French.

Communities of Twitter Users Around Charlie Hebdo

After defining the sample, the first step we took was to identify communities of users by studying the graph of retweets and mentions in this corpus. Specifically, we examined the tweets of users with at least 10 tweets and we looked for user names mentioned in the message. The tweet issuer and the mentioned names became the vertices of a graph, the relationship between users symbolized by the links between these peaks. In the end, we obtained a directed graph of interactions between users. We then used the OpenOrd algorithm proposed by Blondel, Guillaume, Lambiotte, and Lefebvre (2008) and available in Gephi[7] that looks for the nodes that are more densely connected together than the rest of the network to determine communities inside both networks. These communities are composed of sets of users who tend to mention and retweet each other often. Research has shown that communities identified in Twitter networks correspond to opinion or affinity groups (Mousavi & Gu, 2015). Graph 1 and 2 give an account of the main communities detected that represent 64% of users of the English graph and 73% of the French one.



Graph 2. Network based on mentions and retweets of users having produced at least 10 tweets in English including the string *charliehebdo*.

The dots of the graph represent Twitter accounts that produced a message containing the string *charliehebdo* or accounts mentioned in such messages. The lines between the dots represent the interactions between these accounts (retweets (RTs) and mentions). The topology of the graph is the result of the intensity of interactions between accounts. The more two accounts are characterized by an intense two-way communication between them, the closer they are on the map. Dot size (smaller to bigger) depends on the number of mentions and RTs received. The color depends on the community to which they belong. On the two graphs, we have "appointed" names to communities based on the most mentioned accounts through an inductive method. Basically, we identified the most mentioned accounts in each community and then we manually examined their Twitter profiles to characterize them on the basis of two criteria that emerged as pertinent: political position and nationality. Finally, we named the whole community based on the characterization of these accounts. This step is purely qualitative, but it is the only way to qualify communities of thousands of linked users. Its validity has been tested on different empirical grounds (Smyrnaios & Ratinaud, 2014).

The Structure of the Networks

The network structure in both samples is quite different. In the French network, the main clustering factor is political homophily. In the English network, the most prominent clustering factor is nationality, followed by political homophily. On the other hand, a common feature is the position of mainstream media that we find in clusters at the heart of both networks: Parisian media such as Le Figaro, Le Point, Le Parisien, Le Nouvel Observateur, and Le Monde in the French sample (bright yellow cluster) and international media such as the BBC, CNN, Russia Today, AFP, and The Wall Street Journal in the English sample (green cluster). This means that they receive many mentions by number of common user accounts belonging to all the other communities. One particularity about French media is that the two news channels iTélé and BFMTV are associated with different cluster (sky blue and purple, respectively) in the French sample, a feature probably related to their coverage of live events that generated a lot of live tweeting. Near the center of the French graph, there is also a group of accounts related to official bodies (Ministry of the Interior, Police, Gendarmerie, Police Prefecture). As in the case of the aforementioned media accounts, these were disseminating raw information and instructions to the public in relation to the terrorist attacks.

However, a small group of media stands out of the central media cluster mentioned above and is part of a much larger set (red, bottom left of the graph). These are public service media (France Inter, France Culture, RFI, Radio France, France 24, FranceTVinfo) and others positioned on the left of the political spectrum (Libération, L'Humanité, Mediapart, Politis, Reporterre). In the same set, there are many accounts of progressive journalists and leftist politicians (Michel Mompontet, Sylvain Lapoix, Jean-Luc Mélenchon, or Martine Billard) as well as other users clearly stating their commitment to the left. On the other hand, the far right and part of the conservatives form another cluster (bright orange on the top) structured around a few accounts with a lot of mentions: Fdesouche a blog engaged in the far right, the populist National Front's official account, that of Fabrice Robert, president of the Identity Bloc, a radical group, the official account of National Front's leader Marine Le Pen but also right-wing media such as Valeurs Actuelles and Atlantico. Nicolas Sarkozy and the accounts of the conservative party occupy a boundary position between the group in question and the rest of the graph by creating their own (small) group. Finally, two different clusters at the right of the graph (clear and dark blue) are composed of users with no particular or visible political affiliation that aggregate around show business personalities (we called these two communities Apoliticals and Celebrities).

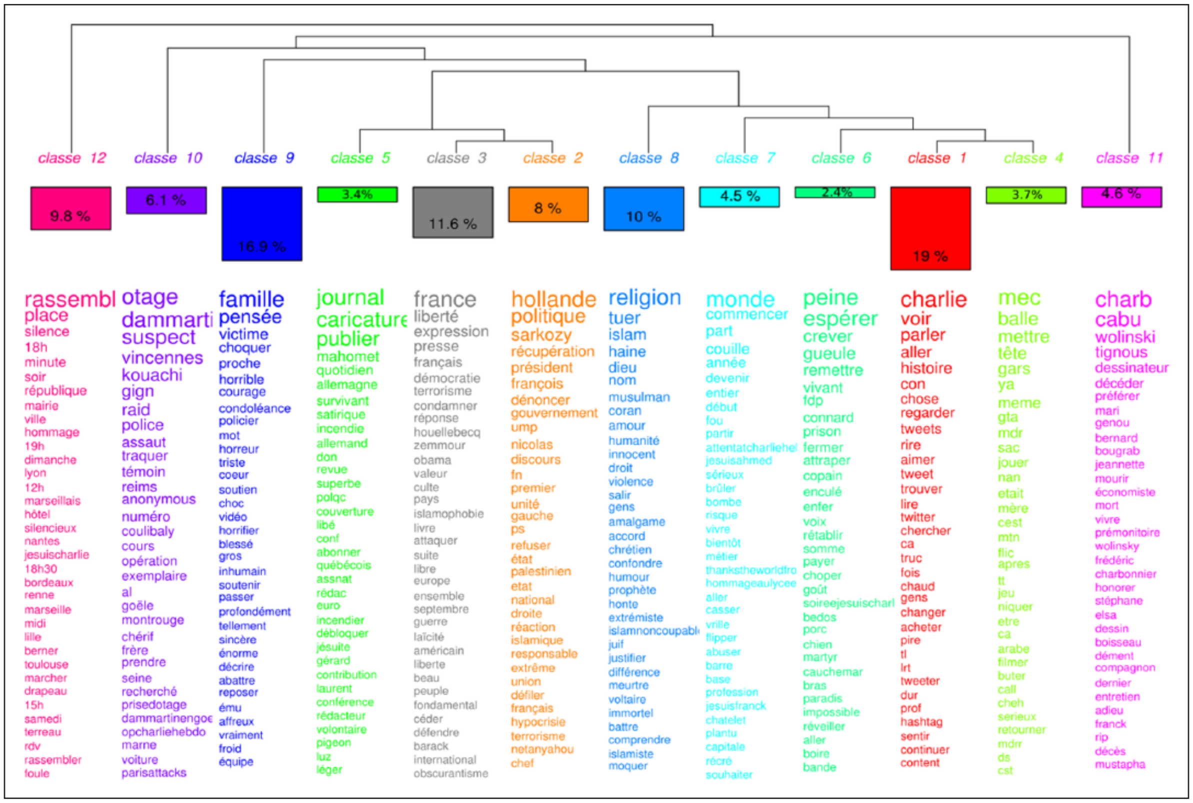
A comparable phenomenon can be observed in the English sample with a national twist; some progressive and urban US media such as New York Magazine, Slate, The New Yorker, Buzzfeed, The Huffington Post, or The Daily Beast are mostly linked to artists such as Banksy, Shuja Rabbani, Rob Tornoe, or Lucille Clerc, forming the main nodes of a community we called Liberal Artists (yellow). On the other hand, conservative media such as The Daily Telegraph and Fox News are linked to Republican pundits such as Greta Van Susteren and Megyn Kelly to form a right-wing US community comprising numerous partisans of the Republican Party (fuchsia). Some important nodes of this network such as Arsen Ostrovsky and Yair Rosenberg are openly pro-Israeli. On the exact opposite of the graph, we find a Liberal Pro-Palestinian network composed mainly of Muslims (blue). What is interesting about this community is that only one media, the Middle East Monitor, is central. The two most important nodes are a Turkish journalist, Borzou Taragahi, and Latuff, a Brazilian cartoonist known for his pro-Palestinian positions. Finally, another common feature between the two graphs is the presence of minority national or cultural communities that aggregate in distinct clusters. This is the case with Indians and Turks in the English sample and francophone Canadians (Québec) in the French sample.

The Discourses Around Charlie Hebdo

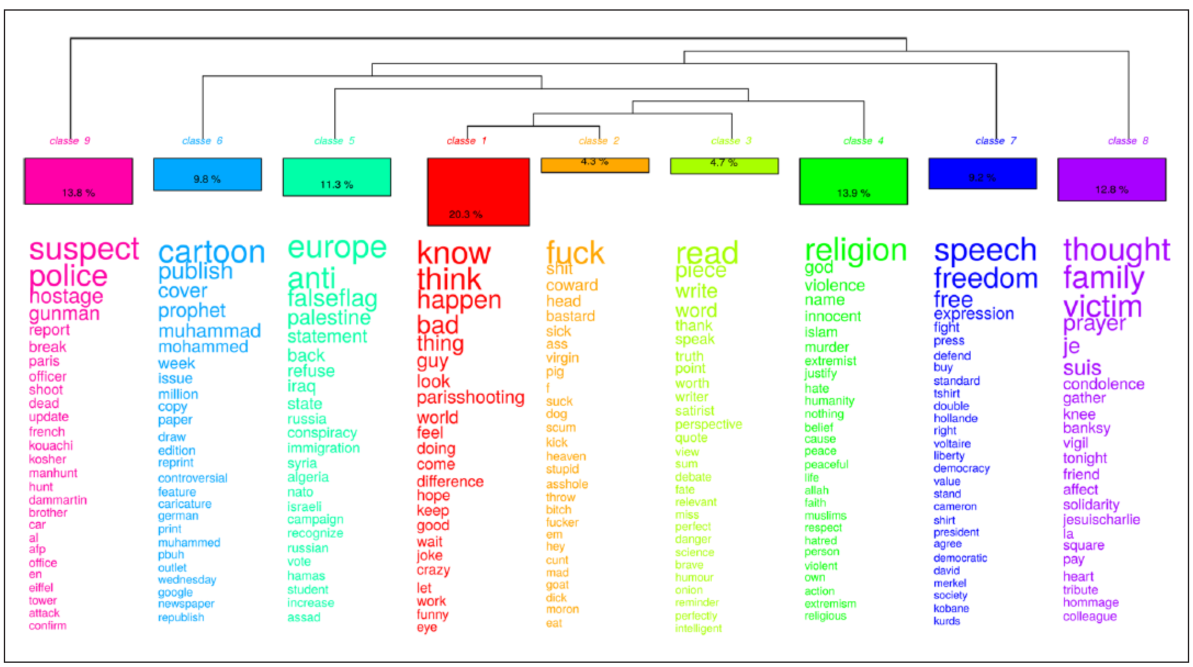
After having identified the main communities of users around Charlie Hebdo, we wanted to link them to the different discourses that were produced on Twitter. To do that, we therefore submitted the corpus of original tweets containing the string *charliehebdo* to a downward hierarchical classification with IRaMuTeQ. In other words, we used the software to group in separate categories (lexical clusters) the tweets that tend to contain the same words. Dendrogram 1 and 2 show the classification trees, the size of each cluster, and the overrepresented lexicon in each cluster, respectively, in the French and the English corpus.

Through the classification of the French and the English corpus of tweets, we obtained, respectively, 12 and 9 lexical clusters. Some of them refer to common themes and frames, while others are different. The common themes include factual descriptions of the hostage takings that occurred on 9 January (Cluster 10 in French and 9 in English), emotional messages expressing sympathy and condolences to the families of the victims (Cluster 9 in French and 8 in English), messages expressing resentment and horror about what happened (Cluster 7 in French and 3 in English), declarations about the importance of freedom of expression that was targeted by the attacks (Cluster 3 in French and 7 in English), descriptions of Charlie Hebdo with reminders that it published cartoons of the prophet Mohammed in 2011 (Cluster 5 in French and 6 in English), denunciations of religious fanaticism and killing in God's name (Cluster 8 in French and 4 in English), and insults against the terrorists with numerous occurrences of the term "pig" (Cluster 6 in French and 2 in English). The clusters that are exclusive to the French corpus include calls to the marches that took place in France on 10 and 11 January (Cluster 12), comments on policy responses and critiques against politicians such as François Hollande and Nicolas Sarkozy for exploiting the terrorist attacks (Cluster 2), comments on the video showing a terrorist shooting down a police officer during the attack (Cluster 4), and tweets that list the names of the victims of this attack (Cluster 11). The clusters that are exclusive to the English corpus are composed of messages referring to the international implications of the attacks notably in the Middle East as well as denunciations of the attacks as false flag or conspiracy (Cluster 5) and a discussion about the true nature of Charlie Hebdo (Cluster 6).

What we notice here is that the most common themes in the two samples are emotional reactions, insults, factual descriptions and information, tributes to freedom of expression, and denunciations of religious fanaticism. In other words, the common themes are also mostly consensual. These themes are also the most prominent in size. On the other hand, those that are exclusive in each sample are those that refer to



Dendrogram 1. Classification tree, sizes of clusters (classe) as percentage of the corpus and over-represented words in each cluster in the French corpus of original tweets.



Dendrogram 2. Classification tree, size of clusters (classe) as percentage of the corpus and over-represented words in each cluster in the English corpus of original tweets.

particular political debates and events in each linguistic space (protests and criticizing politicians for exploiting the terrorist attacks in French, the Middle East conflict in English) as well as the most polemical (conspiracy theories and discussions about Charlie Hebdo's true nature in English).

Another interesting variable to understand the structure of the Charlie Hebdo web sphere is temporality. Charts 3 and 4 project links between the date of publication of tweets and lexical clusters. Links are estimated through a chi2 reflecting the trend to find in clusters a statistical overrepresentation (a higher proportion) or a statistical underrepresentation (a lower proportion) of tweets produced at each dates. The bars going to the right signal an overrepresentation of tweets from this date in the cluster: the bars going to the left signal an underrepresentation.

Globally, the production of tweets is coherent with the events. In both samples, emotional reactions and condolences come first (Clusters 9 in French and 8 in English are overrepresented on 7 and 8 January). Factual description and information coincides with particular events (for instance, the hostage takings described in Cluster 10 in French and 9 in English that is overrepresented on 9 January or the protest against terrorism that took place in Paris described in Cluster 12 in French that is overrepresented on 12 January). Commentary and more polemical debates are overrepresented toward the end of the period.

Which Discourses for Which Community?

The last phase of our analysis is based on examining the presence of lexical clusters among the different communities of users. In other words, this phase shows which were the main types of discourse that were used by each community (and those that were not). Charts 5 and 6 graphically render overrepresentation and underrepresentation of clusters of tweets for each group of users.

When we compare the two charts, the first obvious common feature is the overrepresentation of factual descriptions of the hostage takings (Cluster 10 in French and 9 in English) among both French and English speaking mainstream media. Both media groups also show underrepresentation of denunciations of religious fanaticism and killing in God's name (Cluster 8 in French and 4 in English). Another common feature is the particularity of highly politicized groups compared to others. In the French sample, calls in favor of freedom of expression and against terrorism (Cluster 3) but

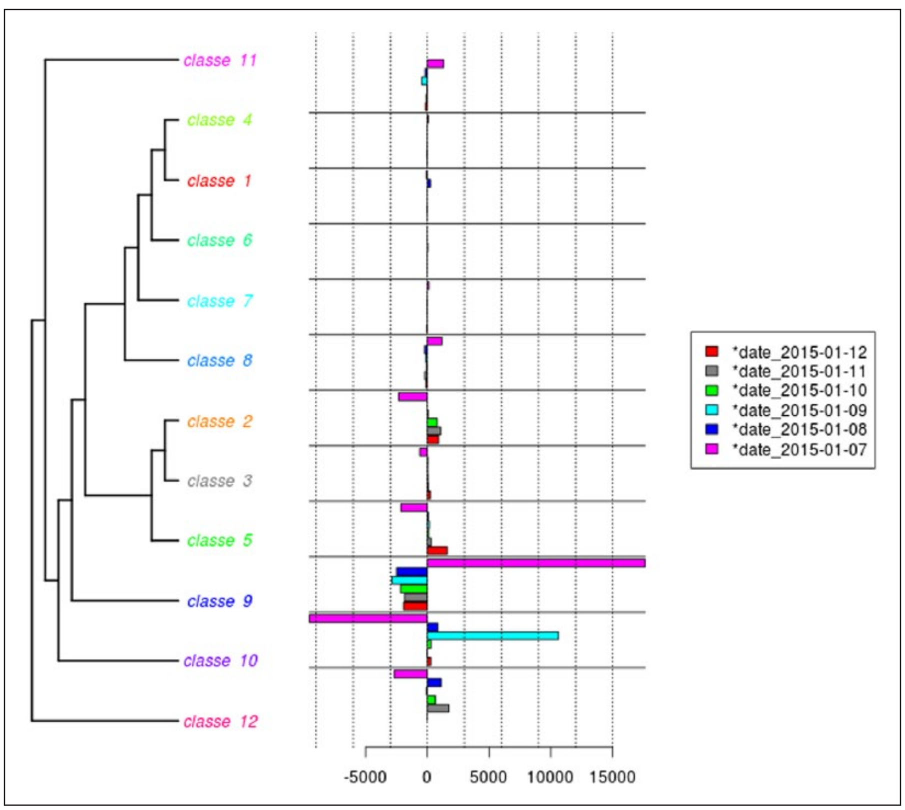


Chart 3. Links between lexical clusters and date in the French sample (Chi2).

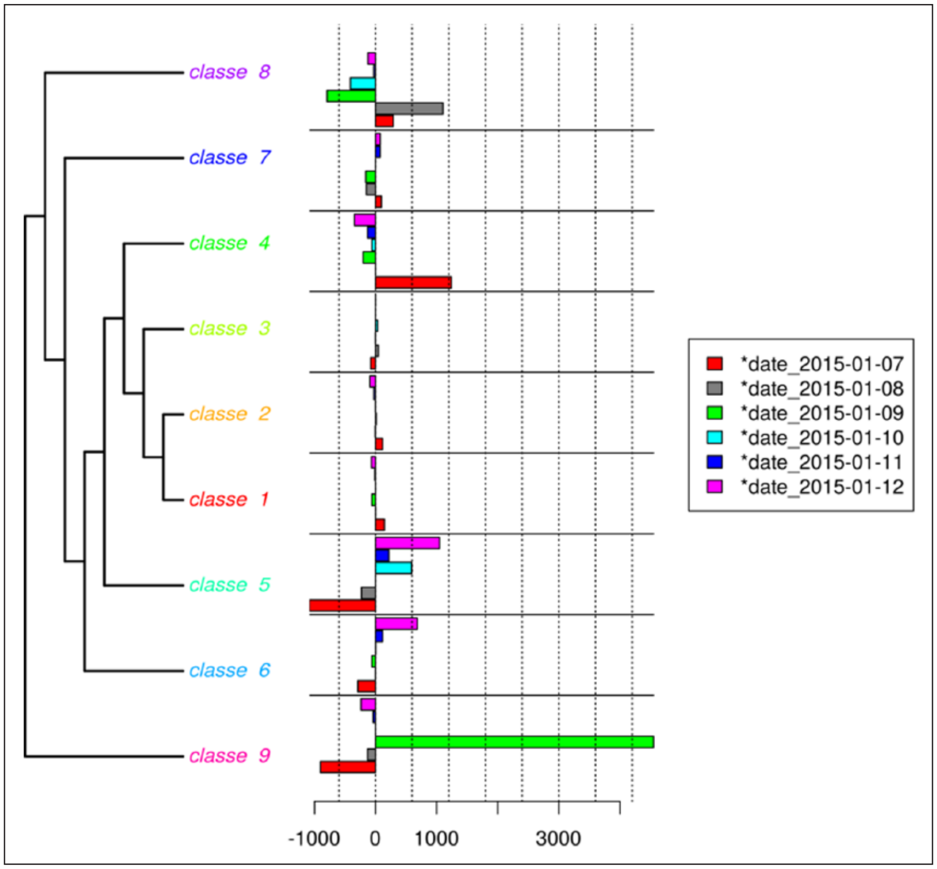


Chart 4. Links between lexical clusters and date in the English sample (chi2).

also critiques against policy responses and politicians, especially, François Hollande (Cluster 2), are overrepresented in the Far right. The latter is also prone to factually describe the hostage takings (Cluster 10). Emotional messages expressing sympathy to the families of the victims are underrepresented in both communities of the Left and, more importantly, the Far right (Cluster 9). In the English sample, factual descriptions of Charlie Hebdo as well as insults are slightly overrepresented among Conservatives (Clusters 6 and 2). But both Conservatives and Liberal Pro-Palestinian Muslims are very strongly prone to mention the Middle East and conspiracy theories (Cluster 5) and not to express sympathy for the family victims (Cluster 8).

Finally, when introducing temporality, we observe that the most politicized communities in both samples tend to be underrepresented on 7 January, the day of the attack on Charlie Hebdo, and overrepresented between 10 and 12 January.

Discussion and Conclusion

Our study has produced an overview of the web sphere that developed on Twitter around Charlie Hebdo, using an innovative method to articulate communities of users and discursive themes through a period of 6 days after the attacks. Of course, our method as well as the results it produced has several limitations. First of all a web sphere has no existence per se. There is no central body that declares its birth and defines its scope and duration. A web sphere is the result of a necessarily partial observation. Therefore, there can be no absolutely comprehensive corpus for the analysis, despite the illusion that can be created by very big samples. In our case, two key choices that were made in the sampling process have limited the scope of the study: the time frame of 6 days and the keyword we used to gather tweets. These decisions were rationally taken to produce a representative sample: the period starts with the attack on Charlie Hebdo, covers all the other important events of the sequence (attack on Hyper Kasher, hostage takings, manhunt, and death of the perpetrators) and finishes on a Monday, the day after the Paris march; *charliehebdo* was chosen because it was the most popular and most neutral keyword used to tweet about the events. But, nevertheless, it is possible that our results may have been different if we had extended the period of observation and included in the sample tweets containing other keywords. Another methodological caveat is the qualitative characterization of lexical clusters and Twitter users' communities. In both cases, we chose to characterize them on the basis of the most important entities they contained (the most often used

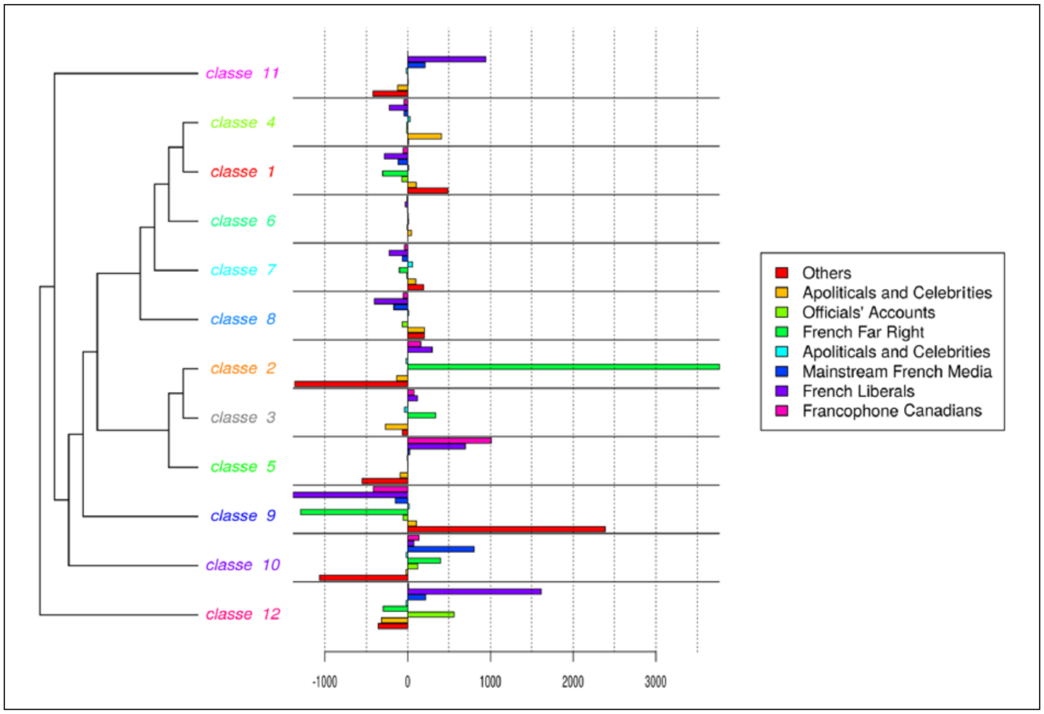


Chart 5. Links between lexical clusters and communities in the French sample (chi2).

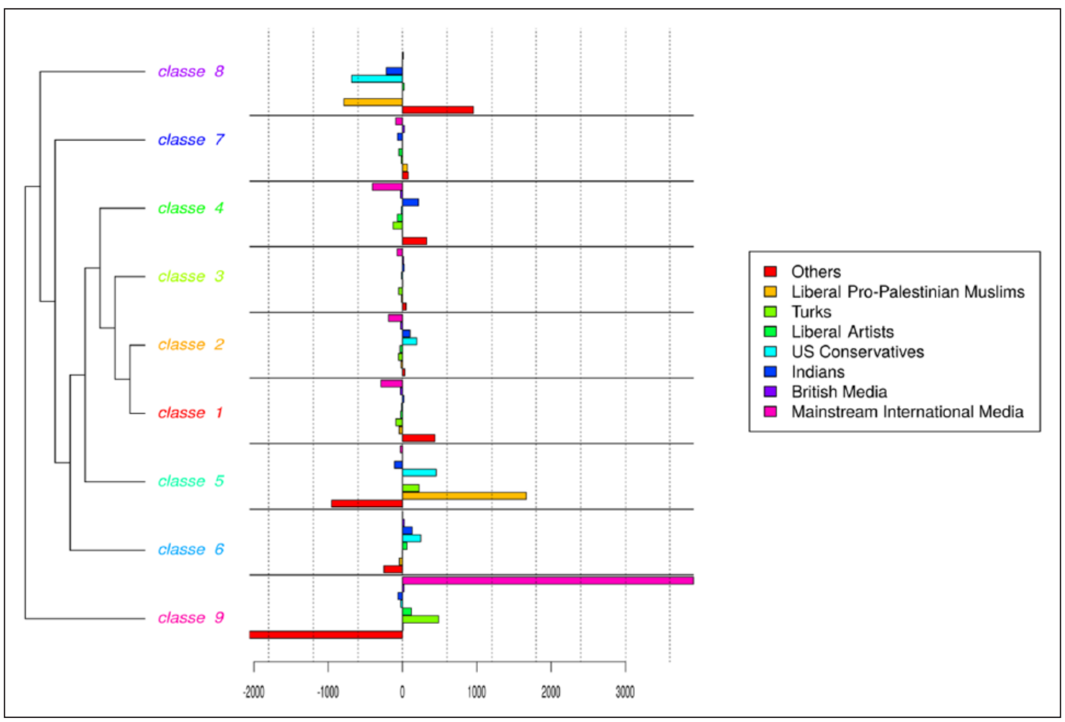


Chart 6. Links between lexical clusters and communities in the English sample (chi2).

words in the former, the most cited accounts in the latter). This doesn't mean that either lexical clusters or user communities are completely homogeneous, something that would be impossible in such a large sample. Nevertheless, we are convinced that our choice is pertinent given the scope and the finality of our study which is to monitor and analyze an online

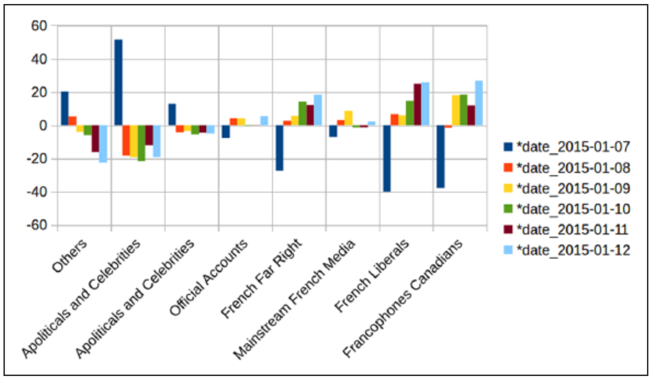


Chart 7. Standardized residuals of chi-square crossing French communities and dates.

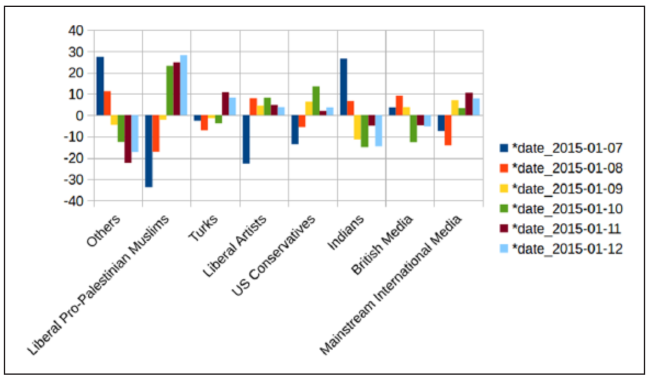


Chart 8. Standardized residuals of chi-square crossing English communities and dates.

political controversy with enough perspective to capture the overall picture, to identify the main issues and mechanisms of contestation as well as the main groups involved.

In this respect, several of our results are of interest. First of all we observed that the majority of the users who tweeted about Charlie Hebdo were not strongly connected to a particular network of other users so as to form a community (Others represent 58.95% of the tweets in the French sample and 59.98% in the English one). These people were overrepresented on the day of the attack on the Charlie Hebdo offices (7 January) and their dominant discourse was made of condolences and messages of sympathy to the families of the victims (Cluster 9 in French and 8 in English). On the other hand, in both samples, they were absent from the factual description of the hostage taking on 9 January and from the most polemical debates (Middle East and conspiracy theories in English, exploitation of the drama by politicians in French). In other words, these are "normal" people, without any particular affiliation, who were shocked by the attacks and immediately expressed solidarity to the victims and their families. But rather quickly they lost interest on ongoing events and didn't participate in politically polarizing discussions.

On the opposite, we found that smaller highly politicized and polarized groups such as French Liberals and French Far right (respectively, 11.12% and 6.37% of the tweets in the French sample) and Liberal Pro-Palestinian Muslims and US conservatives (respectively, 6.35% and 7.31% of the tweets in the English sample) had similar attitudes toward the events. They were less engaged immediately after the attacks in emotional expression of sympathy and shock (except French Liberals), but they participated vividly in the following days in polemical discussions or engaged themes: the Middle East and conspiracy theories for the English speaking, critique of mainstream politicians for the French Far right, and calls for protest for the French Left. The cases of the Liberal Pro-Palestinian Muslims, the US conservatives, and the French Far right are particularly interesting. The first two were both engaged in the same discursive theme that is strongly framed by the opposition between Pro-Palestinian and Pro-Israeli. It seems that the Charlie Hebdo attacks simply reactivated an arena of confrontation that pre-exists and an opposition that strongly characterizes US politics. The French Far right, on the other hand, is traditionally not fond of Charlie Hebdo because of its perceived leftist culture. Thus, what seems to interest this community in the context of the attacks is the opportunity to criticize (and insult) Muslims and also mainstream politicians such as François Hollande and Nicolas Sarkozy, who both participated in the "national unity" march of 11 January. This illustrates perfectly the argument used by Far right politicians that both the conservatives and the social democrats are the two sides of the same coin.

More generally, our results show clearly that even the most dramatic events such as a terrorist attack with innocent victims does not produce completely homogeneous reactions online. Rather, political engagement and cultural dispositions are keys to understand different attitudes on Twitter.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was supported by Agence Nationale de la Recherche (ANR-11-IDEX-0002-02) under grant Labex SMS (ANR-11-LABX-0066).

Notes

1. Public online expression and terrorism has been mainly addressed through the questions of propaganda and radicalization, that is, Archetti, Christina, "Terrorism, Communication and New Media: explaining radicalization in the digital era," Perspectives on Terrorism, Vol. 9, (online), http://www.terrorismanalysts.com/pt/index.php/pot/article/view/401/html

2. To address this problem, we choose to use only open software, the code of which is available and transparent.
3. In our case, we consider tweets as segments.
4. The Digital Methods Initiative Twitter Capture and Analysis Toolset (DMI-TCAT) is an open-source software developed by Erik Borra and Bernhard Rieder (2014), https://github.com/digitalmethodsinitiative/dmi-tcat
5. An important epistemological caveat when working on samples that are extracted from platforms such as Twitter comes from the fact that whatever the method used one can never be absolutely sure that a corpus is completely comprehensive.
6. Fox (2013).
7. Gephi is an open-source software for graph visualization and network analysis: http://gephi.org

References

An, J., Kwak, H., Mejova, Y., Alonso Saenz De Oger, S., & Gomez Fortes, B. (2016, May). *Are you Charlie or Ahmed? Cultural pluralism in Charlie Hebdo response on Twitter*. Proceedings of the 10th International AAAI Conference on Web and Social Media (ICWSM), Cologne, Germany.

Badouard, R. (2016). Je ne suis pas Charlie. Pluralité des prises de parole sur le web et les réseaux sociaux. In P. Lefébure & C. Sécail (Eds.), *Le défi Charlie. Les médias à l'épreuve des attentats* (pp. 187–220). Paris, France: Lemieux Editeurs.

Blondel, V., Guillaume, J.-P., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *10*. doi:10.1088/1742-5468/2008/10/P10008

Borra, E., & Rieder, B. (2014). Programmed method: Developing a toolset for capturing and analyzing tweets. *Aslib Journal of Information Management*, *66*, 262–278. doi:10.1088/1742-5468/2008/10/P10008

Callon, M. (1986). Eléments pour une sociologie de la traduction. La domestication des coquilles Saint-Jacques et des marins-pêcheurs dans la baie de Saint-Brieuc. *L'année sociologique*, *36*, 169–208.

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. P. (2010, May 23–26). *Measuring user influence in Twitter: The million follower fallacy*. Proceedings of the 4th International Conference on Weblogs and Social Media, Washington, DC.

Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Flammini, A., & Menczer, F. (2011, July 17–21). *Political polarization on Twitter*. Proceedings of the 5th International AAAI Conference on Weblogs and Social Media, Barcelona, Spain.

Fox, Z. (2013, December 18). Top 10 most popular languages on Twitter. *Mashable*. Available from http://mashable.com

Garraud, P. (1990). Politiques nationales: L'élaboration de l'agenda. *L'année sociologique*, *40*, 7–41.

Giglietto, F., & Lee, Y. (2015, May). *To be or not to be Charlie: Twitter Hashtags as a discourse and counter-discourse in the aftermath of the 2015 Charlie Hebdo shooting in France*. Proceedings of Microposts2015 5th Workshop on Making Sense of Microposts, Florence, Italy.

Marres, N. (2015). Why map issues? On controversy analysis as a digital method. *Science Technology Human Values*, *40*, 655–686. doi:10.1177/0162243915574602

Mousavi, R., & Gu, B. (2015). *The effects of homophily in Twitter communication network of U.S. House Representatives: A dynamic network study*. Retrieved from http://ssrn.com/abstract=2666052

Ratinaud, P. (2014). IRaMuTeQ: Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires (Version 0.7 alpha 2). Available from http://www.iramuteq.org

Reinert, M. (1983). Une méthode de classification descendante hiérarchique: Application à l'analyse lexicale par contexte. *Les cahiers de l'analyse des données*, *VIII*, 187–198.

Reinert, M. (1990). ALCESTE: Une méthodologie d'analyse des données textuelles et une application: Aurélia de Gérard de Nerval. *Bulletin de Méthodologie Sociologique*, *26*, 24–54.

Rieder, B. (2012). The refraction chamber: Twitter as sphere and network. *First Monday, 17*. Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/4199/3359

Rieder, B., & Röhle, T. (2012). Digital methods: Five challenges. In D. M. Berry (Ed.), *Understanding digital humanities* (pp. 67–84). Basingstoke, UK: Palgrave Macmillan.

Rouquette, M.-L., & Moscovici, S. (1994). *Sur la connaissance des masses: Essai de psychologie politique* (Vol. 8). Grenoble, France: Presses universitaires de Grenoble.

Schneider, S. M., & Foot, K. A. (2006). Web Sphere analysis: An approach to studying online action. In C. Hine (Ed.), *Virtual methods: Issues in social research on the Interne* (pp. 157–170). Oxford, UK: Berg.

Smyrnaios, N., & Ratinaud, P. (2014). Comment articuler analyse des réseaux et des discours sur Twitter: L'exemple du débat autour du pacte budgétaire européen. *tic&société*, *7*, 120–147. doi:10.4000/ticetsociete.1578

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011, April 18–21). *Comparing Twitter and traditional media using topic models*. Proceedings of the 33rd European Conference on IR Resarch, ECIR 2011, Dublin, Ireland.

Author Biographies

Nikos Smyrnaios (PhD University of Grenoble) is an associate professor at the University of Toulouse. His research interests include the political economy of online media, digital journalism, and the political use of social networking sites.

Pierre Ratinaud (PhD University of Toulouse) is an associate professor at the University of Toulouse. His research focuses on dynamics of social representations systems. He's also the developer of IRaMuTeQ, an open-source software for texts and questionnaires multidimensional analysis.