Dr. Marc Blasco-Duatis
Universitat de Girona. Spain. marc.blasco@udg.edu

Dr. Germà Coenders
Universitat de Girona. Spain. germa.coenders@udg.edu

Sentiment analysis of the agenda of the Spanish political parties on Twitter during the 2018 motion of no confidence. A compositional data approach

Análisis de sentimiento de la agenda de los partidos políticos españoles en Twitter durante la Moción de Censura de 2018. Un enfoque de datos composicionales

Abstract

This paper provides a sentiment analysis of the political parties' discussion on Twitter about the motion of no confidence in the Spanish government in 2018. In particular, 2824 tweets from the official accounts of the 13 political parties represented in the Congress of Deputies were extracted and analysed. The methodology applied compositional data analysis and its visualisation employing Biplot (a visualisation tool that enables contrasting the relative importance of the elements under study). Unlike traditional approaches, this study emphasises the relative importance of the issues within the agenda while incorporating a third category, the sentiment analysis. The main findings concern the reliability of the method to represent compositionally the agenda and the agenda-setters, as well as the sentiment analysis. This confirms that the issues most notably associated with certain parties also reflect their projection on sentiment. In sum, this analysis sheds light on sentiment representation in the agenda-setters (attribute agenda), especially in the field of political communication.

Resumen

Este artículo presenta el análisis de sentimiento de la discusión de partidos políticos en Twitter, en el contexto de la Moción de Censura al gobierno español de 2018. En particular, extrajimos y analizamos 2824 tweets de las cuentas oficiales de las 13 formaciones políticas representadas en el Congreso de los Diputados. En su desarrollo metodológico aplicamos el análisis composicional de datos y su visualización a través del biplot (una herramienta de visualización que permite contrastar la importancia relativa de los elementos en estudio). A diferencia de los enfoques tradicionales, nuestro estudio enfatiza la importancia relativa de los temas dentro de la agenda, a la vez que incorpora un tercer componente, el análisis de sentimiento. La investigación concluye sobre la fiabilidad del método para representar composicionalmente la agenda y los agenda-setters, así como el análisis de sentimiento, constatando que los temas que se asocian de forma más notable con determinados partidos, también lo hacen con su proyección sobre los sentimientos. El análisis arroja luz sobre la representación de los sentimientos en los agenda-setters (agenda de atributos), especialmente en el campo de la comunicación política.

1. Introduction

The last decade was a testimony of the consolidation of social networks as an important channel through which the political leaders convey their discourses, making them more accessible to their potential voters (Golbeck et al., 2010; Conway et al., 2015; Graham et al., 2016). The professionalization of political discourse in the networks was not only meant as another form of political communication capable of influencing the social sphere, but also, at the same time, it has proven its capacity to influence the political and media agenda. Within this media-social ecosystem, Twitter was established as one of the main political spaces of spreading information to a growing audience (Conway et al., 2015), appearing as a reference platform in order to gather information, efficiently connect with potential voters and follow the activity of the sources (Parmelee, 2013). In this way, Twitter claims to be the standardised tool by which connections among parties, journalists and audiences are established, especially during the election campaign processes (Jungherr, 2016), thus being an important channel to measure the reactions of the public opinion on behalf of the political action. For this reason, the inherent interactivity of the platform has significantly fostered the exposition of citizens in the political discussion and the ideological confrontation (Guo et al., 2016).

Twitter has been established as one of the most important global platforms in terms of its immediacy in diffusing information among the society (Yousuf Al-Aama, 2015). The most outstanding socio-political events of the last years like the Egyptian revolution in 2011, the election of Trump, the referendum for peace in Colombia, Brexit, or the most recent movements like #MeToo or #FridaysForFuture, constitute a current evidence. The new dynamics in the relation among the media and the audiences propose a greater pluralism in the media sphere, carrying the traditional model of political communication through an unprecedented transformation of the interaction among their principal agents. All of this is part of a systemic reformulation process of the dynamics of power and influence, where the message is not transferred to the public unidirectionally from the traditional, but it is understood as a construct which is fed and developed through the horizontal and convergent interaction of the various actors and platforms (Scolari et al., 2019). So much so, that in the last decade, from the academy, the investigation in the field of political, media and public agendas has been especially focused in the study of digital media as new spaces to influence the construction of the contents that are dealt within (Meraz, 2009).

During the second quarter of 2018, the Twitter platform registered 338 million profiles worldwide and 4.9 million in Spain. Their users published an average of 500 million tweets per day in the whole world (Statistic Brain, 2018). Moreover, it is significant that the number of active users on Twitter exceeds 22% of internet users in the world (Kayser and Bierwisch, 2016). These statistics reflect the global spread of this social network and its potential impact. Besides having a global coverage of the main daily issues, it provides a space that allows to share opinions easily, using various content resources, including text, images, and links, in contrast to many other social media platforms. It also provides almost real-time access to the posts through the API, boosting it as an efficient platform to contrast large scale states of opinion almost in real time. In this sense, we decided to focus on Twitter as the main source of data, among other social networks, in order to carry out the analysis proposed in the study.

In this context, the emotional dimension of a political discussion in the social media turns out to be of particular importance, since an emotional debate on a controversial issue often develops in a more dynamic and unpredictable way than an objective discussion (Kušen et al., 2017). Thematic analysis methods have contributed to classifying and comprehending the emerging interests among the public on the basis of the various agendas. However, the complexity of the studies referring to the dynamics of emission-reception make the analysis of the generated sentiments in these discursive actions a difficult task (Fan and Gordon, 2014). This is why case studies about real-world political events are of particular interest, since they contribute to the understanding of human behaviour, to detecting patterns and to identifying generic approaches in order to analyse the behaviour of the actors involved in managing the on-line social networks (see, for example, Ahmadian et al., 2017; or Ferrer-Rosell et al., 2019, 2020b).

The sentiment analysis on Twitter has been used in a wide spectrum of areas related with governance and public trust, that range from the prediction of resentment against the government politics to the prediction of general elections' results (Tumasjan et al., 2010; Calderon et al., 2015). This emerging field aims to comprehend and predict such behaviour, and even though this area of study is still evolving and generates a lot of enthusiasm, there is a constant debate about the effectiveness of the use of sentiment analysis on Twitter to predict elections and other events of the real-world.

Regarding research on the use of Twitter in the political agenda, scholars have been initially interested in the way in which the platform influences in determining the creation of such thematic agenda for the main communication media (Jungherr, 2014; Wallsten, 2014), but it has gone far beyond that (Calvo and Aruguete, 2020). The analysis has also centred the interest of researchers in the study of the second level agenda (or attribute agenda), where the objective is to decipher the given focus on the subject, which

relates undoubtably with sentiment analysis. A recent example is that of Lee and Xu (2018). Another field of study is that of dynamic relations among agendas (Neuman et al., 2014). In the same way, by the method of social networks analysis, it has been studied what has arrived to be termed third level agenda, in which the associations between perceived themes for different audiences have been revealed (Guo and Vargo, 2015).

From the beginning of the study of the agenda-setting theory it is supported that certain media or social agents rise, or intend to rise, as mediators of information, in an praxis based on including or emitting certain themes in a scenario of content hierarchy (McCombs and Shaw, 1972; McCombs, 2004). As a result, the relations between those that we can call agenda-setters, or at least they pretend to be, and the contents (themes on the agenda), were traditionally studied from a descriptive standpoint, by counting repetition frequencies of a theme, or through the Spearman correlations, where only the order in the theme's ranking is registered and any other kind of information is omitted.

Unfortunately, the traditional methods of analysis mentioned above do not actually centre the analysis on the relative importance of the themes of the agenda, nor do they provide visualisation tools. To resolve this limitation, in the last years work has begun with the application of compositional data analysis (CoDa) in the field of political communication (see Blasco-Duatis et al., 2018a, 2018b, 2018c, 2019). CoDa is the standard statistical methodology which is applied when the researcher is interested in the relative importance of the parts of a whole (Aitchison, 1986).

In our research, we focus especially on a well-known data visualisation tool, the compositional biplot, that allows us to see which themes are prioritised by which political parties (that act or try to act as agenda-*setters* when they use their Twitter accounts), which topics change their presence proportionally, and which parties have a similar behaviour according to the themes, having in mind that what is important does not reside in the volume of information but in the relative weight of each theme for each party.

The innovative contribution of this article lies in posing a visualisation of compositional data that allows the attribute agenda to be represented, including the themes suggested by the political parties and the sentiments linked with the published messages on Twitter. In the first place, we employ the tool "Tweet Binder" to capture the published tweets through the official accounts of the political parties in study. Next, we apply a methodology based on content analysis to detect, on one hand, the themes that form the agenda and, in a second phase, the sentiments derived from the semantic and contextual constructions of those tweets. This serves in applying, in a third phase, a new type of compositional biplot which permits us to deal with more than one composition at the same time (Ferrer-Rosell et al., 2020c; Filzmoser et al., 2018; Kynčlová et al., 2016), where the focus lies in the proportions among the various themes of the agenda, sentiments, the relative importance of both for the agenda-setters and the relation between themes and sentiments.

In sum, in this article we not only provide a compositional study of the agenda and the agenda-setters in Twitter (a theme that has been already been developed in the seminal work of Blasco-Duatis et al., 2019) but we extend the compositional analysis in a second level, the sentiment analysis, which carries a value, from what we modestly know, that has not been explored so far in a compositional way in the field of political communication. We specifically compile and analyse 2824 tweets from the 13 political formations represented in Spanish Deputy Congress in 2018. We gather a data set of thematic character, of the transmission and information about the sentiments of the agenda-setters, that are expressed within the content of the messages. The rest of this work is structured in a section presenting the methodology, which is divided in two sub-sections consisting of a first process of codification-categorisation, and a second one of compositional visualisation of the data. Further on, the results over the sentiment analysis of the agenda on Twitter by the political parties in the context of the no confidence motion to the government in 2018, will be presented.

2. Methodology

The methodological proposal that we present constitutes a novel contribution in the field of sentiment analysis of political communication, since it applies the analysis and visualisation of compositional data in a study that, along with representing the mapping of the agenda, incorporates a third element, the sentiment analysis (attribute agenda). The method is developed in a first phase through the content analysis of the set of themes processed during the motion of no confidence to the government in 2018 by the main Spanish political parties on Twitter (first level agenda), and the sentiment codification derived from the semantic constructions of each tweet (second level agenda). In a second stage, and since the quantity is not comparable because the number of tweets and analysed sentiments is different in each party, we are interested in relative measurements. What is actually important in this assumption is the

appearance percentage of a theme of the agenda or sentiment (category) in each political party (or agenda-setter), as well as the relations among them. In the literature this compositional approach is used to analyse data that entail relative information, when the focus lies in the proportions between the different vectors-categories, or themes of the agenda and sentiments, and their agenda-setters or political parties (Aitchison, 1986).

2.1. Selection and codification of the tweets

There is a common affirmation in the field of sentiment analysis, that successful results depend, to a large degree, on the specific development that has been used to gather the set of data to be studied. In this sense, most of the methodological approaches in this field employ systems based on supervised techniques of automatic and statistical learning. Even though they have proven to be quite successful in the past-and not only in the field of sentiment analysis, but also in the majority of text extraction and information retrieval applications- their obvious disadvantage, in terms of functionality, is their limited applicability in domains of data different from those they were designed for (Pang and Lee, 2004). Also, the codification of sentiments requires a system that provides a measurement of the semantic orientation of the text, and not only about the particular terms, limiting a lot the use of traditional models based on text mining, that are usually found in the use of manually developed dictionaries to which a valence modifier is associated.

For this reason, the present article considers that the manual codification of each tweet not only ensures an assertive thematic clarification (theme of the agenda), but at the same time, it permits to define with higher precision the sentiment derived from the semantic context. Even though it is true that in some cases a content can be associated with certain key words easy to detect with this method (i.e. the concept *machismo*, clearly associated with gender politics), in other cases it is not possible (i.e. the concept *dialogue*, that according to context can refer to the necessity of discussing about the Catalan independence process, or against dialoguing with certain leaders or parties involved in said process).

Starting from this conception, we apply content analysis (Berelson, 1952; Krippendorff, 2004; Wimmer and Dominick, 2006) adapted to the Twitter environment. We collect a data base with the sequence of the set of tweets published on the official Twitter accounts of the parties represented in the Parliament in the time of the motion, referred to as: PP (Partido Popular); PSOE (Partido Socialista Obrero Español); UP (Unidas Podemos, the regional "confluencias" and also the integration of IU); C's (Ciudadanos); EAJ-PNV (Euzko Alderdi Jeltzalea - Partido Nacionalista Vasco); ERC (Esquerra Republicana de Catalunya); PDeCAT (Partit Demòcrata Català); Compromís (Compromís); Bildu (Euskal Herria Bildu); CC (Coalición Canaria); NC (Nueva Canaria); FA (Foro Asturias); and UPN (Unión del Pueblo Navarro). The total number of tweets gathered during the 6 days of the study (from 29 of May to 3 of June 2018), that included the two days prior to the first session of the parliament debate (1st of June) until the two days after the second debate and the approval of the motion (3rd of June), was: PP, 247 tweets; PSOE, 463 tweets; UP, 180 tweets; C's, 491 tweets; EAJ-PNV, 61 tweets; ERC, 493 tweets; PDeCAT, 226 tweets; Compromís, 184 tweets; Bildu, 216 tweets; CC, 130 tweets; NC, 91 tweets; FA, 19 tweets; and UPN, 23 tweets. The last two regional parties (FA and UPN) were discarded from the compositional study due to lack of data representation in comparison with the sample of parties in the study.

Following Blasco-Duatis et al. (2019), a list of eighty topics related with the coding of the set of tweets analysed, was compiled. Subsequently, and for conceptual similarity, the topics were grouped into 12 categories from which only 9 with sufficient representative entity were selected, hereinafter named top9: (1) Catalan independence (application of the article 155 of the Spanish constitution, referendum 1-O, separatism, independence, TV3, Bildu and Otegui on the Catalan independence, parallelism with the Basque process, formation of the Government of Generalitat de Catalunya, President of Parliament Carme Forcadell's birthday, appointment of the Consellers of Generalitat); (2) territorial policy (autonomy policy, strictly territorial questions, regional competences, multinationalism, Canary Islands' Day, the fall of the Podem municipal government in the city of Badalona); (3) economy-employment (General Budgets of Spain, regional funds, pensions, taxes, municipal capital gains tax, PNV against UPN in the blocked European investments, global economic crisis); (4) corruption (commission of inquiry into the irregular financing of Partido Popular, the Bárcenas case, positions of trust, the Dolores de Cospedal case, the RTVE Court of Auditors case, the Orihuela mayor case, the Daniel De Alfonso case, the commission for the corruption of Partido Popular at the Congress, informative manipulation on corruption); (5) gender (gender violence, La Manada case, equality, family conciliation, cases of church abuse); (6) European policy (steel agreement, debates at the European Parliament, relations of the EU countries with Spain, European policy associated with the Catalan case); (7) education-culture (pact on Basque education, education commission at the Congress, new Valencian toponyms, Sant Andreu de la Barca Secondary School case, universities, research, university grants); (8) social policies-health (Menas, surrogacy, dependency,

environment, sustainability, adapted sport, freedom of expression, social emergency, vote to repeal the Ley Mordaza, health, drug trafficking in Campo de Gibraltar); (9) regeneration policies (change of government, two-party system, health of democracy, health of institutions). It should be noted that in the process of coding the tweets it was deemed appropriate to categorise each one with as many topics (categories) as identified in their text. That is how there are various tweets that have been categorised in two or more categories (i.e. a tweet about the implications of the Catalan independence in an EU country has been coded with two themes (6) European politics and (1) Catalan independence).

At the same time, and related with the sentiments' codification process, most of the published works in the field have focused on the dual approach of "approval or rejection", that is to obtain a positive or negative qualification (Turney, 2002). An additional step of this model implies an intention to determine the textual indicators not only by a binary classification, but also establishing a scale of intermediate sentiments that are associated with certain semantic and context constructions. In this sense, and following the pioneer work of Moreno-Ortiz and Pérez Hernández (2013) about the sentiment analysis of the messages in Spanish on Twitter, we decided to establish the following categories: VP (positive, in favour); P (more positive than negative, more in favour than against); NEUT (neutral, neither positive nor negative, of informational character); N (more negative than positive, more against than in favour); VN (negative, against); NONE (no sentiment is associated, usually derived from tweets linked to the activities' agenda of the party or its representatives). So, unlike the process of codification-categorisation described before about the themes of the agenda, we attribute a single sentiment to each tweet. In a compositional level, the application of this method is particularly important in its representational phase, since it does not only permit the study of the themes by their issuers (agenda-setters), but it also transforms the analysis in a three-component study and adds the sentiment composition to the specific interactions of the agenda.

2.2. Analysis and visualisation of data

Following Marine-Roig and Ferrer-Rosell (2018), Ferrer-Rosell et al. (2019, 2020a, 2020b, 2020c), and Ferrer-Rosell and Marine-Roig (2020), the analysis of contents and communication styles centres its interest in the relative importance of such contents and styles, that is their proportions, which turns such analysis into a compositional analysis (Aitchison, 1986; Van den Boogaart et al., 2013; Egozcue and Pawlowsky-Glahn, 2016; Filzmoser et al., 2018; Greenacre, 2018; Pawlowsky-Glahn et al., 2015).

The *analysis of compositional data* is defined as the analysis of the parts of a whole when the interest lies on the relative size of their *D* elements, parts or components. A composition is expressed as a vector of positive parts:

$$\mathbf{x} = \left(x_1\ x_2\ ...x_D \right)\ \text{with}\ \ x_j > 0,\ j = 1, 2, ..., D \cdot \tag{1}$$

In our case, the part x_j represents the number of times that the content or theme appears in the tweets of a specific party. Compositional methods presuppose that two parties distinguished from the fact that one emits twice as much of each theme than the other, constitute of two identical parties concerning the research questions. Thus, such questions consider it to be trivial and of minor interest that some parties are more active in the networks than others, globally speaking, and therefore emit more of each theme.

To calculate a measurement of central tendency, taking into consideration such relative and not absolute importance of the themes, the geometric mean is used, and it is rescaled to an arbitrary total that is easily understood, for example 1 (proportions) or 100 (percentages).

To calculate a measure of association, in this type of data correlations are disregarded and replaced by the concept of *proportionality*. If two themes *j* and *k* move proportionally, their ratio remains stable and the logarithm of their ratio has zero variance.

$$\text{var}\left(\ln \frac{x_j}{x_k} \right) = 0. \tag{2}$$

Thus, the correlation matrix is substituted by the so-called *variation matrix,* which contains such variances of logarithms of ratios calculated for every possible theme pair *j*=2,3,...,*D*; *k*=1,2,...,*j*−1. Like the correlation matrix, it is symmetrical. As commented, the value zero implies a perfect direct association between two themes. Instead, there is not any upper threshold that indicates a perfect inverse association. High values of the variance indicate that the two themes do not vary proportionally.

As stated in (2), the logarithms of ratios, termed log-ratios, play a fundamental role in the analysis of compositional data. The *transformation with centred log-ratios (clr)* indicates the relative importance of

each theme, regarding the other themes present in the composition. One way to understand this transformation is comparing each part with the geometric means of the rest (Filzmoser et al., 2018):

$$clr(\mathbf{x}) = \frac{D-1}{D}\left(\ln\left(\frac{x_1}{\sqrt[D-1]{x_2 x_3 ... x_D}} \right) \quad \ln\left(\frac{x_2}{\sqrt[D-1]{x_1 x_3 ... x_D}} \right) \quad ... \quad \ln\left(\frac{x_D}{\sqrt[D-1]{x_1 x_2 ... x_{D-1}}} \right) \right). \tag{3}$$

When the data contain zeros (in our case, themes not often recurring that are not treated in any message of a specific party) the log-ratios cannot be calculated, therefore it is necessary to somehow deal with the zeros before the analysis. In content analysis this corresponds to the so-called *count zeros,* and is usually treated by Bayesian imputation methods, as it is recommended in Martín-Fernández et al. (2015).

In the same way as classical statistical analysis, the analysis of compositional data requires graphic tools for data visualisation. A very useful tool is the *compositional biplot* (Aitchison and Greenacre, 2002; Egozcue and Pawlowsky-Glahn, 2016), and more specifically the *compositional covariance biplot.* It is the standard biplot resulting from the *principal component analysis* based on the covariance matrix of the transformed data with centred log-ratios. The biplot represents jointly the relative importance of the *D* contents or themes subjected to the *clr* transformation and the agenda-setters (in this case the political parties), in the first two dimensions of the principal component analysis. The themes are usually represented as vectors with origins in the centre of the coordinates and the agenda-*setters* as points. The biplot constitutes the best possible approximation of the original data in two dimensions. The quality of this approximation is indicated by the variance percentage explained by those first two dimensions.

In the *covariance biplot* the main interpretational element is the *between the extremes of the theme vectors.* Such distance is approximately proportional to the standard deviation of the log-ratio between the two involved themes, in other words, the square root of (2). Extremes of the vectors that are close to each other indicate themes that keep an approximate proportionality, in the sense, for example, of theme pairs that if the proportion of one is doubled in certain political parties, the proportion of the other is doubled as well. Extremes far from each other indicate the contrary, in the sense, for example, of theme pairs that if the proportion of one increases in certain parties, the proportion of the other decreases.

The biplot also permits to visualise the approximate importance of each theme for each party, in relative terms. If we *project the parties orthogonally* (we "drop" them forming a 90-degree angle in the direction defined for each theme vector), the direction indicated by the theme vector points towards the parties with a higher content of the theme, relatively speaking, identifying, in this way, which themes contribute to distinguish each party from the rest.

More recently, biplots with more than one composition have been developed (Ferrer-Rosell et al., 2020c; Filzmoser et al., 2018; Kynčlová et al., 2016) that allow to complement the traditional interpretation which has been explained within the same composition, with the relationship between compositions. The present article constitutes the first compositional content analysis in using this tool.

Given the centred log-ratios of a second composition y (in our case the number of times that each one of the *S* sentiments is expressed in the tweets of each party):

$$clr(\mathbf{y}) = \frac{S-1}{S}\left(\ln\left(\frac{y_1}{\sqrt[S-1]{y_2 y_3 ... y_S}} \right) \quad \ln\left(\frac{y_2}{\sqrt[S-1]{y_1 y_3 ... y_S}} \right) \quad ... \quad \ln\left(\frac{y_S}{\sqrt[S-1]{y_1 y_2 ... y_{S-1}}} \right) \right) \tag{4}$$

the biplot of two compositions is constructed from an analysis of principal components based on the covariance matrix between the *D+S* centred log-ratios. In this biplot, the distances between the extremes of the vector of the same composition, are interpreted on the same way as described. The parties can be projected orthogonally upon any vector of any composition. Additionally, the cosines of the angles between the vectors of different compositions are approximately equal to the correlations between their corresponding centred log-ratios, and therefore between their relative importance. Let us remember that two parallel arrows in the same direction correspond to an angle of zero degrees, whose cosine is one and indicates the maximum positive correlation; that two parallel arrows in the opposite direction correspond to an angle of 180 degrees, whose cosine is minus one and indicates the maximum negative correlation, and that two perpendicular arrows correspond to an angle of 90 degrees, whose cosine is zero and indicates no correlation. Thus, acute angles indicate positive correlation and obtuse angles negative.

It is advisable that *D* and *S* are of the same type of magnitude, otherwise one of the two compositions could dominate the analysis, which can result in poor quality of the representation of the other.

The replacement of the zeros was done with the R package "zCompositions" (Palarea-Albaladejo and Martín-Fernández, 2015) with the command cmultRep default options, which is the so-called *Geometric*

Bayesian Multiplicative method (GBM) of Martin-Fernandez et al. (2015). The percentage of zeros, as recommended, was reduced, 6.1% in the contents' composition and 3.0% in the sentiments' composition. The rest of the analyses were done with the programme CoDaPack (Thió-Henestrosa and Martín-Fernández, 2005).

The advantages of using compositional data analysis, the biplots and the transformation by centred log-ratios have been treated both generally (Aitchison, 1986; Van den Boogaart et al., 2013; Filzmoser et al., 2018; Greenacre, 2018; Pawlowsky-Glahn et al., 2015) and also in the specific case of political communication studies (Blasco-Duatis et al. 2018a; 2018b; 2018c; 2019).

Table 1: Example of the differences between the traditional approaches of relative frequencies (above) and ranks (centre), with the compositional data approach with log-ratios (below)

	VP	P	NEUT	N	VN	NONE
Original data out of 100						
Bildu	46.3	10.7	7.0	8.4	22.0	5.6
PSOE	60.7	10.4	7.1	3.0	17.5	1.3
UP	54.2	9.5	6.7	2.8	26.3	0.6
Ranks						
Bildu	6.0	4.0	2.0	3.0	5.0	1.0
PSOE	6.0	4.0	3.0	2.0	5.0	1.0
UP	6.0	4.0	3.0	2.0	5.0	1.0
Centred log-ratios						
Bildu	1.3	-0.1	-0.6	-0.4	0.6	-0.8
PSOE	2.0	0.2	-0.1	-1.0	0.8	-1.8
UP	2.0	0.3	-0.1	-0.9	1.3	-2.6

Source: Own elaboration

We summarise here some of the most important ideas with a small example. Compared with the traditional approaches of using the original data referred to a common total, and transforming the data as ranks, the approach presented here is one that takes into consideration the differences in the relative importance of the themes and sentiments. Table 1 shows a fragment of the sentiments' composition for three parties. The biggest difference in absolute value is encountered between Bildu and PSOE regarding the VP sentiment. Even so, the difference in relative terms between 60.7 and 46. 3 is only 31%. The biggest difference in relative terms is encountered between Bildu and UP regarding the NONE sentiment. The difference in relative terms between 5.6 and 0.6 is 833%.

On the original data, the Euclidean distance Bildu-PSOE is 16.6 and the distance Bildu-UP 11.8, considering the bigger differences in absolute value. On ranks, the distances Bildu-PSOE and Bildu-UP are both 1.4 because even though PSOE and UP are not identical in their values, they are so in their ranks. On the centred log-ratios, the distance Bildu-PSOE is 1.5 and the distance Bildu-UP is 2.2, taking into account the bigger differences in relative terms, that constitute the analysis objective.

Another aspect to have in mind are the spurious correlations since the rows of Table 1 have always a fixed sum. In relative terms one part can only increase if at least another part is reduced, which leads to mainly negative and uninterpretable correlations. Of the 15 correlations between pairs of sentiments, 7 are negative on the original data, 11 are negative on the ranks and 12 are negative on the centred log-ratios. For this reason, compositional analysis avoids using the correlations among the parts of the same composition, and it substitutes them with the variances of the log-ratios between pairs of parts (2), that do not bear this problem. The correlations between parts of different compositions are valid and they are the only ones that we use in this article.
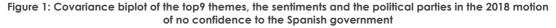
3. Results

According to the geometric means expressed over a total equal to 100, the composition centres show that the most recurring themes are, in this order: territorial policy (21.2%), economy-employment (17.6%), corruption (16.1%), Catalan independence (15.9%), regeneration policy (9.3%), social policies-health (6.3%), education-culture (5.6%), gender (4.2%) and European policy (3.9%); and the sentiments more frequently expressed are, in this order: VP (40.1%), P (19.2%), VN (16.4%), N (11.9%), NEUT (9.8%), and NONE (2.1%).
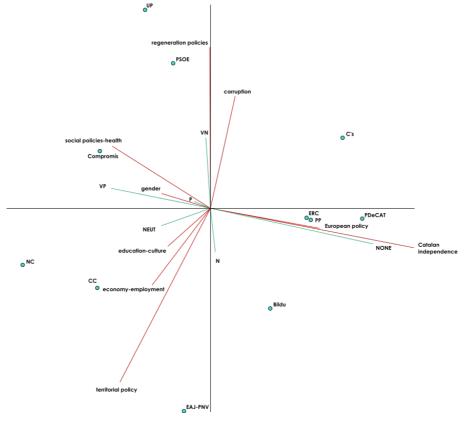
The variance explained by the first two dimensions of the principal component analysis is 55%. A third dimension would allow to reach 71% at the cost of the visualisation of a two-dimensional graphic. As in any principal component analysis, the representation quality of each vector depends on its individual variance percentage explained by the first two dimensions, of which the mentioned 55% is only an average (Daunis-i-Estadella et al., 2011). In this case, the contents worst represented are gender (24.6%) and education-culture (21.4%). The sentiments worst represented are P (3.4%) and N (16.7%). Their extremes and directions in the graphic must be interpreted with extreme caution. On the contrary, the following vectors are represented in an especially reliable way: the Catalan independence (72.6%), the territorial policy (77.5%), the corruption (60.8%), the social policies and health (66.9%), the regeneration policy (67.6%), the sentiments VP (56.3%) and the absence of sentiments -NONE (74.7%).

According to the proximities of the theme vector extremes, the pairs of themes that tend to move together in the different political parties are regeneration policy with corruption, European policy with Catalan independence, economy and employment with social policies and health, and economy and employment with territorial policy.

According to the proximities of the sentiment vector extremes, the sentiments that tend to move together in the different political parties are the central ones, that is, P, NEUT and N, even though, as we have mentioned, some of them are represented in the biplot with poor quality. The most prominent opposition is found between NONE and the rest all the others.

According to the orthogonal projections on the vectors represented with higher quality, the most notable associations between parties and themes are Cs, ERC, PDeCAT and PP with the Catalan independence; UP and PSOE with regeneration policies and corruption; and NC, CC, Bildu and EAJ-PNV with economy-employment and territorial policy. The most notable theme absences that are generally popular in particular parties are corruption in NC, CC, Bildu and EAJ-PNV; Catalan independence in NC, UP and Compromís; territorial policy in C's and UP; and economy-employment in PSOE, UP and C's.

Figure 1: Covariance biplot of the top9 themes, the sentiments and the political parties in the 2018 motion of no confidence to the Spanish government

Source: Own elaboration

According to the orthogonal projections, the most notable associations between parties and sentiments are C's, ERC, Bildu and PP with NONE; Compromís, UP, C's and PSOE with VN; and NC, UP, Compromís, PSOE and CC with VP. The most notable sentiment absences in parties are VP in C's, ERC, PDeCAT and PP; and VN in NC, CC, Bildu and EAJ-PNV.

As we have indicated, in this article we present for the first time a biplot of two compositions in the content analysis of political communication. As a novelty, this approach opens the door of interpreting as correlations the angles between parts of one composition and parts of another, and in our case to answer the question concerning which sentiments tend to be expressed by the parties in relation to which themes during their communication on Twitter.

Acute angles between vectors show the following notably positive correlations between themes and sentiments: VP with social policies-health; NONE with the Catalan independence and European policy; and VN with regeneration and corruption. Obtuse angles between vectors show the following notably negative correlations between themes and sentiments: VP with the Catalan independence and with European policy; VN with territorial policy; NEUT with corruption and NONE with social policy-health. The angles that are almost right indicate absence of relation. For example, to express VP sentiments is independent of whether the theme refers, or not, to corruption or regeneration, to express VN sentiments is independent of whether the theme refers, or not, to European policy.

4. Discussion

As we pointed out before, there are several studies that have analysed the Twitter transmission data as a source of identifying current news and real-world events (Broersma and Graham, 2012; Thelwall et al., 2011; Neuman et al., 2014; Guo and Vargo, 2015; Calvo and Aruguete, 2020; Lee and Xu, 2018), concluding that the tendencies on Twitter are usually the most important events of the day and can be used to predict the news headlines. In our study, we detect that the themes that centered the debate of all the parties were, essentially, those that impelled PSOE to lead the 2018 motion of no confidence to the government. That is how the issues related to territorial policy, economy/employment, corruption and the Catalan independence were the central thematic axis of the parties' conversation, with approximately 70% of thematic coverage. In this way, we can confirm that the framework set by PSOE for the discussion on the motion of no confidence was, at the same time, the mark that centered the discussion of the political parties on Twitter.

At the level of the parties' sentiment (attribute agenda), the messages of positive or very positive sentiment (close to 60%) stand out clearly above the negative or very negative ones (a little over of 28%). A possible explanation to these data would be the expected supports that PSOE had in that moment in order to forward the motion, since it had potentially more parties in favour of it (UP, ERC, PDeCAT, PNV, Compromís, Bildu, NC), than against (PP, C's, UPN, FA). Yet, and as we noted in the previous chapter of the results, it is worth highlighting that the most representative sentiments among the political parties (that are usually associated and, therefore they move together) are those in the centre (P, NEUT and N).

In our study about the themes that focused the agenda it is also interesting to see the association established among certain themes. That is how we can clearly identify pairs of themes that are proportional between them, meaning that the more you talk about one theme, the more you talk about the other. This phenomenon is especially visible with the following pairs: regeneration policies and corruption, central axis of the motion debate; European policy and Catalan independence, clearly associated with the internationalisation of Catalan political conflict; economy/employment and social policy/health, focused in reversing the austerity policies of the government of Mariano Rajoy; the same as with economy/employment and territorial policy, associated with territorial inequalities and lack of regional investment.

In the link established between specific themes and political parties, it is interesting to comment the following associations: C's and PP vs. ERC and PDeCAT with the Catalan independence, where the first two parties maintain a position clearly contrary to the other two, but they are highly equivalent in terms of dedication on the theme; NC, CC, Bildu and EAJ-PNV with economy/employment and territorial policy, an association clearly connected with the traditional territorial demands of the nationalist parties about the investment in the region they represent; and finally, UP and PSOE with regeneration and corruption, which was not only the central programmatic axis of PSOE in defending the motion to the government, but it also was one of the elements of the opposition in the years of the government of Mariano Rajoy.

The notable correlations between sentiments and themes also contribute with some interesting results to comment about the development of the second level agenda. While it is true that the general perception

about the Catalan independence theme is such that in the last years has raised highly polarised debates as far as the derived sentiments are concerned, in the case that concerns us of the agenda-*setters* on Twitter, a tone of neutrality predominated among the parties during the motion. On the contrary, themes like social policies-health (VP) or regeneration and corruption (VN), maintained their logic -as established by the socio-political debate of the country- in capturing extreme sentiments. This phenomenon is strengthened by certain absences of relation as for example NEUT with corruption.

Precisely in the theme of corruption (as well as that of regeneration, that as we have commented they appear to be closely related) it is interesting to have a pause for thought in the analysis of its components, since the representation of the leitmotif that impulses the motion on behalf of PSOE can be observed in it. The parties associated in the most notable way with this theme and with a sentiment of VN are UP and PSOE, followed by C's and Compromís. On the other hand, PSOE and some parties that would support the motion (UP and Compromís) would also wish that the substitute government would be more sensitive to social policies, health, gender and equality, themes on which the sentiment VP was expressed. On the contrary, those who are further away from this alliance are the nationalist parties like Bildu, CC, NC and PNV (especially focused on the themes of territorial policy and with a sentiment more N than VN concerning corruption). Meanwhile, PP has a very close position (both at the thematic and sentiment level) to the Catalan independence formations of ERC and PDeCAT. In this sense, corruption is not a priority theme for those three parties, but they focus, and from NONE sentiment, in the themes of Catalan Independence and European policy. From all this, we are able to verify that while the motion was presented to suspend the functions of a government, whose party had just been condemned for corruption, the same governing party (PP) hid behind the Catalan independence frame in order to avoid the main theme of the motion of no confidence.

These results respond to our initial intention to achieve a visualisation of compositional data that would allow us to represent the attribute agenda, by advancing the representation of the agenda and the agenda-*setters* as political parties, towards the visualisation of the sentiments connected to the messages published on Twitter. With that, not only have we managed to map which parties emphasize which contents or which themes or parties are more or less similar to each other, but we have also incorporated, in this compositional study concerning the agenda, the sentiment analysis. Therefore, we consider the method's capacity of sentiment visualisation and study in the scene of the attribute agenda proven, being especially relevant for its ability to prioritise the relativity of the elements under study.

5. Conclusions

In this article we present a sentiment analysis on the discussion over Twitter, by the main political parties concerning the 2018 motion of no confidence to the government of Spain. We extract and analyse compositionally 2824 tweets, as well as their associated sentiments, that were posted by the political parties represented at the Deputy Congress. In particular, we develop, document and apply a methodological approach which is innovative in the field of sentiment analysis concerning political communication, systematising its application in order to achieve the visualisation of those sentiments in the narration of the political parties, and on the various topics of the agenda exposed on Twitter.

In the of 2018 motion of no confidence to the Spanish government, we find that:

- 70% of the parties' conversation on Twitter on the motion of no confidence was centred on only four themes (territorial policy, economy/employment, corruption and Catalan independence). Those themes, that served PSOE as "red" lines in order to put forward this motion, were the framework of the discussion on Twitter by the political parties.

- The predominant sentiment in the parties' messages was more positive (60%) than negative (28%), an atmosphere clearly associated with the fact that there are more parties potentially in favour of the motion and, therefore, wishing that it would all end up moving forward.

- There are certain groups of themes that are more associated among them than with others, and at the same time they are also associated with certain groups of parties. In this line the following stand out: C's, PP, ERC and PDeCAT with Catalan independence; NC, CC, Bildu and EAJ-PNV with economy/employment and territorial policy; and finally, UP and PSOE with regeneration and corruption.

- The sentiments more equitably distributed in the messages of the parties tend more towards the centre (P, NEUT and N), than the extremes (VP, VN).

- It is remarkable that the theme "Catalan independence" which has raised big sentiment polarities in the public debate, would reach a clear tone of neutrality in the case of the parties' debate for the motion of no confidence on Twitter.

- Regarding the previous point, PP has a very close position (both at the thematic and sentiment level) to the Catalan independence formations of ERC and PDeCAT. In this sense, corruption is not the desirable frame for those three parties, but they displace the agenda towards the themes of Catalan independence and European policy.

Likewise, by applying CoDa to the sentiment analysis of the agenda on Twitter, we have been able to test the viability of the method concerning its visualisation and study, focusing the analysis on the relative importance of each theme of the agenda (content), about each party (agenda-*setter*) and for each sentiment (attribute agenda).

In this way, the application of the method to the present case study, has allowed us to visualise and study the second level agenda (or attribute agenda). Through the sentiment analysis we have been able to confirm how the main attributes (VP, P, NEUT, N or VN) are linked to the themes that shaped the agenda and for each party. Thus, and in line to what was presented in the methodological possibilities, the application of CoDa upon two compositions -which for the case we are engaged with is the second level agenda- has proved its valuable contribution to visualise and learn in a compositional way the relations among agenda-*setters*/parties (points), themes of the agenda (vectors) and sentiments (angles). For this last aspect in particular, the sentiments analysed through angles, we note that the acute angles indicate a positive correlation between the theme and the sentiment, and the obtuse angles a negative one.

It is noteworthy, and as a limitation that we have already described in the methodological development, that the studies in this field should centre, as a priority, in the main themes of the agenda, since those themes with little or no recurrence, increase the zeros and, for that reason, hinder the calculation of log-ratios. In fact, the Bayesian imputation for count zeros is no longer reliable when the percentage of zeros is elevated (Martin-Fernandez et al., 2015) and other methods like correspondence analysis become preferable to the compositional methods (Greenacre, 2018). Also, it is highly recommended that the two compositions should be itemised in numbers of parts of the same order of magnitude, since otherwise one of the two compositions would dominate the analysis and distort the visualisation.

This article represents the first proven experience of CoDa of two compositions of the agenda (second level study, or attributes, through sentiments) in the field of political communication. The CoDa methodology goes far beyond data visualisation. For example, a compositional cluster analysis can classify both themes and agenda-*setters* when there is large number of them or cross the agenda-*setter* classification by themes with the classification by sentiments. If longitudinal data is available, the agenda-*setters* may appear in the biplot as trajectories depending on the evolution of their predominant themes, or the agenda can be statistically modelled at a given moment from the agenda at an earlier moment. The repertoire of compositional methods expands rapidly as most of the statistical methods now have their compositional equivalent (Van den Boogaart et al., 2013; Egozcue and Pawlowsky-Glahn, 2016; Filzmoser et al., 2018; Greenacre, 2018; Pawlowsky-Glahn et al., 2015), and with it the further research opportunities focused on the relative importance of themes, issuers and/or attributes.

6. Acknowledgement

7. References

[1] Ahmadian, S., Azarshahi, S. & Paulhus, D.L. (2017). Explaining Donald Trump via communication style: Grandiosity, informality, and dynamism. *Personality and Individual Differences*, *107*, 49–53. http://doi.org/f9pgfd

[2] Aitchison, J. (1986). *The statistical analysis of compositional data. Monographs on statistics and applied probability*. London: Chapman and Hall.

[3] Aitchison, J. & Greenacre, M. (2002). Biplots of compositional data. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 51*(4), 375–392. http://doi.org/bck69d

[4] Berelson, B. (1952). *Content analysis in communication research*. New York: Hafner.

[5] Blasco-Duatis, M., Coenders, G. & Saez, M. (2018a). Compositional visualization of intermedia agenda setting by the main media groups and political parties in the Spanish 2015 General Elections. *Revista Latina de Comunicación Social, 1*(73), 264–292. http://doi.org/dpgb

[6] Blasco-Duatis, M., Coenders, G. & Sáez, M. (2018b). Representación composicional (CoDa) de la agenda-setting de los opinantes sobre política en la prensa escrita española durante las Elecciones Generales de 2015. *Estudios sobre el Mensaje Periodístico, 24*(2), 1121–1146. http://doi.org/dpgc

[7] Blasco-Duatis, M., Sáez-Zafra, M. & Fernández-García, N. (2018c). Compositional representation (CoDa) of the agenda-setting of the opinion makers on politics in the main Spanish media groups in the 2015 General Elections. *Communication & Society, 31*(2), 1–23. https://doi.org/10.15581/003.31.2.1-23

[8] Blasco-Duatis, M., Coenders, G., Sáez, M., Fernández.García, N. & Cunha, I. (2019). Mapping the agenda-setting theory, priming and the spiral of silence in twitter accounts of political parties. *International Journal of Web Based Communities, 15*(1), 4–24. http://doi.org/dpgd

[9] Broersma, M. & Graham, T. (2012). Social media as beat: Tweets as a news source during the 2010 British and Dutch elections. *Journalism Practice, 6*(3), 403–419. http://doi.org/gdqc56

[10] Van den Boogaart, K.G. & Tolosana-Delgado, R. (2013). *Analyzing Compositional Data with R*. Berlin: Springer.

[11] Calderon, N. A., Fisher, B., Hemsley, J., Ceskavich, B., Jansen, G., Marciano, R. & Lemieux, V.L. (2015). Mixed-initiative social media analytics at the World Bank: Observations of citizen sentiment in Twitter data to explore "trust" of political actors and state institutions and its relationship to social protest. *2015 IEEE International Conference on Big Data*, (pp. 1678–1687). http://doi.org/dpgf

[12] Calvo, E. & Aruguete, N. (2020). *Fake news, burbujas, trolls y otros encantos. Cómo funcionan (para bien y para mal) las redes sociales*. Buenos Aires: Siglo veintiuno editores.

[13] Conway, B., Kenski, K. & Wang, D. (2015). The rise of Twitter in the political campaign: Searching for intermedia agenda-setting effects in the presidential primary. *Journal of Computer-Mediated Communication, 20*(4), 363–380. http://doi.org/f7kr8b

[14] Daunis-i-Estadella, J., Thió-i-Fernández-de-Henestrosa, S. & Mateu-i-Figueras, G. (2011). Two more things about compositional biplots: quality of projection and inclusion of supplementary elements. In J.J. Egozcue, R. Tolosana-Delgado & M.I. Ortego (Eds.), *Proceedings of the 4th International Workshop on Compositional Data Analysis* (pp. 1-14). Available at http://bit.ly/2W0azYK

[15] Egozcue, J. J., & Pawlowsky-Glahn, V. (2016). What are compositional data and how should they be analyzed? *Boletín de Estadística e Investigación Operativa, 32*(1), 5-29.

[16] Fan, W. & Gordon, M.D. (2014). The Power of Social Media Analytics. *Communication of the ACM, 57*(6), 74–81. http://doi.org/dpgg

[17] Ferrer-Rosell B. & Marine-Roig, E. (2020). Projected versus perceived destination image. *Tourism Analysis*, 25. http://doi.org/dpgh

[18] Ferrer-Rosell, B., Martin-Fuentes, E. & Marine-Roig, E. (2019). 'Do hotels talk on Facebook about themselves or about their destinations?', In J. Pesonen & J. Neidhardt (Eds.), *Information and Communication Technologies in Tourism 2019* (pp. 344-356). Cham: Springer.

[19] Ferrer-Rosell, B., Coenders, G. & Martin-Fuentes, E. (2020a). Compositional data analysis in e-tourism research. What TripAdvisor reviews complain about hotels in Barcelona?. In X. Zheng, M. Fuchs, U. Gretzel & W. Höpken (Eds.) *Handbook of E-Tourism*. Deutschland: Springer Nature.

[20] Ferrer-Rosell, B., Martin-Fuentes, E. & Marine-Roig, E. (2020b). Diverse and emotional: Facebook content strategy by Spanish hotels. *Journal of Information Technology & Tourism, 22*(4), 53–74. http://doi.org/dw85

[21] Ferrer-Rosell, B., Martin-Fuentes, E., Vives-Mestres, M. & Coenders, G. (2020c). 'When size does not matter: compositional data analysis in marketing research', In R. Nunkoo, V. Teeroovengadum & C. Ringle (Eds.), *Handbook of Research Methods for Marketing Management*, in press.

[22] Filzmoser, P., Hron, K. & Templ, M. (2018). *Applied Compositional data Analysis with Worked Examples in R*. New York: Springer.

[23] Graham, T., Jackson, D. & Broersma, M. (2016). New platform, old habits? Candidates use of Twitter during the 2010 British and Dutch general election campaigns. *New Media & Society*, *18*(5), 765–783. http://doi.org/cv2q

[24] Greenacre, M. (2018). *Compositional data analysis in practice*. New York: Chapman and Hall/CRC press.

[25] Golbeck, J., Grimes, J.M., & Rogers, A. (2010). Twitter use by the U.S. Congress. *Journal of the American Society for Information Science and Technology*, *61*(8), 1612–1621. http://doi.org/czr5b8

[26] Guo, L., Rohde, J. A., & Wu, H. D. (2018). Who is responsible for Twitter's echo chamber problem? Evidence from 2016 U.S. election networks. Information. *Communication & Sociery*, *21*(12), 234–251. http://doi.org/gfdbqn

[27] Guo, L. & Vargo, C. (2015). The power of message networks: A big-data analysis of the network agenda setting model and issue ownership. *Mass Communication and Society, 18*(5), 557-576. http://doi.org/dxb2

[28] Jungherr, A. (2014). Twitter in politics: A comprehensive literature review. *Social Science Research Network*. Available at http://bit.ly/2TOsOgZ

[29] Jungheer, A. (2016). Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology & Politics*, *13*(1), 72–91. http://doi.org/gf6bss

[30] Kayser, V. & Bierwisch, A. (2016). Using Twitter for Foresight: An Opportunity?. *Futures*, *84*(Part A), 50–63. http://doi.org/f9jc6z

[31] Krippendorff, K. (2004). Content qnalysis: An introduction to its methodology. Sage: Thousand Oaks.

[32] Kušen, E., Strembeck, M., Cascavilla, G. & Conti, M. (2017). On the influence of emotional valence shifts on the spread of information in social networks. *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Net-works Analysis and Mining 2017*, 321–324. http://doi.org/dpgj

[33] Kynčlová, P., Filzmoser, P. & Hron, K. (2016). Compositional biplots including external non-compositional variables. *Statistics*, *50*(5), 1132–1148. http://doi.org/dpgk

[34] Lee, J. & Xu, W. (2018). The more attacks, the more retweets: Trump's and Clinton's agenda setting on Twitter. *Public Relations Review, 44*(2), 201-213. http://doi.org/dxb3

[35] Marine-Roig, E. & Ferrer-Rosell, B. (2018). Measuring the gap between projected and perceived destination images of Catalonia using compositional analysis. *Tourism Management*, 68, 236–249. http://doi.org/dpgm

[36] Martin-Fernandez, J.A., Hron, K., Templ, M., Filzmoser, P. & Palarea-Albaladejo, J. (2015). Bayesian-multiplicative treatment of count zeros in compositional data sets. *Statistical Modelling*, *15*(2), 134–158. http://doi.org/dpgn

[37] McCombs, M. & Shaw, D. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, *36*(2), 176–187. http://doi.org/dnd7p4

[38] McCombs, M. (2004). Setting the agenda. Cambridge: Polity Press.

[39] Meraz, S. (2009). Is there an elite hold? Traditional media to social media agenda setting influence in blog networks. *Journal of Computer Mediated Communication*, *14*(3), 682–707. http://doi.org/d79bmf

[40] Moreno-Ortiz, A. & Pérez Hernández, C. (2013). Lexicon-Based Sentiment Analysis of Twitter Messages in Spanish. *Procesamiento del Lenguaje Natural*, 50, 93–100. http://hdl.handle.net/10045/27869

[41] Neuman, W., Guggenheim, L., Jang, S. & Bae, S.Y. (2014). The dynamics of public attention: Agenda-setting theory meets Big Data. *Journal of Communication*. *64*(2), 193-214. http://doi.org/f5x3nb

[42] Palarea-Albaladejo, J. & Martín-Fernández, J.A. (2015). zCompositions—R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems*, 143, 85–96. http://doi.org/f696n3

[43] Pang, B. & Lee, L. (2004). A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics,* (pp. 271–278). Barcelona: Association for Computational Linguistics. http://doi.org/cdz7k3

[44] Parmelee, J. (2013). Political journalists and Twitter: Influences on norms and practices. *Journal of Media Practice*, *14*(4), 291-305. http://doi.org/ggb6jm

[45] Pawlowsky-Glahn, V., Egozcue, J.J. & Tolosana-Delgado, R. (2015). *Modeling and Analysis of Compositional Data*. Chichester: Wiley.

[46] Scolari, C. A., Lugo Rodríguez, N. & Masanet Jordà, M. (2019). Educación transmedia: de los contenidos generados por los usuarios a los contenidos generados por los estudiantes. *Revista latina de comunicación social*, *1*(74), 116–132. http://doi.org/dpgp

[47] Statistic Brain (2018). STATS. Twitter company statistics. Available at http://bit.ly/3cJwoRY

[48] Thelwall, M., Buckley, K. & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, 62(2), 406–418. http://doi.org/bb3n8g

[49] Thió-Henestrosa, S. & Martín-Fernández, J.A. (2005). Dealing with compositional data: The freeware CoDaPack. *Mathematical Geology* 37(7), 773–793. http://doi.org/fv8tj2

[50] Tumasjan, A., Sprenger, T. O., Sandner, P. G. & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. *International Conference on Weblogs and Social Media*, 10, 178–185. Available at http://bit.ly/3cPuvn9

[51] Turney, P.D. (2002). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, 417–424. Available at http://bit.ly/334ibL2

[52] Wallsten, K. (2014). Microblogging and the news: political elites and the ultimate retweet. In K. Wallsten (Ed.), *Political campaigning in the information age*, (pp. 128–147). http://doi.org/dpgq

[53] Wimmer, R. & Dominick, J. (2006). Mass Media Research: an introduction. Wadsworth: Thomson.

[54] Yousuf Al-Aama, A. (2015). The use of Twitter to promote e-participation: connecting government and people. *International Journal of Web Based Communities*, *11*(1), 73-96. http://doi.org/dpgr