Article

Journal of Information Science
1–16

Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0165551510000000
jis.sagepub.com
SAGE

The megaphone of the people? Spanish SentiStrength for real-time analysis of political tweets

David Vilares
Grupo LYS, Departamento de Computación, Facultade de Informática, Universidade da Coruña

Mike Thelwall
Statistical Cybermetrics Research Group, School of Mathematics and Computing, University of Wolverhampton

Miguel A. Alonso
Grupo LYS, Departamento de Computación, Facultade de Informática, Universidade da Coruña

Abstract
Twitter is an important platform for sharing opinions about politicians, parties and political decisions. These opinions can be exploited as a source of information to monitor the impact of politics on society. This article analyses the sentiment of 2,704,523 tweets referring to Spanish politicians and parties from a month in 2014-15. The article makes three specific contributions: (1) enriching SentiStrength, a fast unsupervised sentiment strength detection system, for Spanish political tweeting; (2) analysing how linguistic phenomena such as negation, idioms and character duplication influence Spanish sentiment strength detection accuracy; and (3) analysing Spanish political tweets to rank political leaders, parties and personalities for popularity. Sentiment in Twitter for key politicians broadly reflects the main official polls for popularity but not for voting intention. In addition, the data suggests that the primary role of Twitter in politics is to select and amplify political events published by traditional media.

Keywords
Sentiment Analysis; Political Analysis; Twitter

1. Introduction

The online component of politics is widely recognised as important. Barack Obama is sometimes cited as the first major politician to effectively harness web networks for traditional political purposes [1] and Facebook, Twitter and YouTube played an important role in the Arabic Spring movement [2]. In Spain, the success of the popular 15M political protest movement was partly due to organising through social media [3]. It is therefore important to develop methods to analyse social media to gain insights into the online components of political participation. Although the traditional way to measure popularity or voting intentions is through opinion polls [4], these, even when reliable, are costly and time-consuming. This is a particular problem for regional or local elections, where exhaustive surveys are infeasible given the number of districts and candidates. In addition, polls are usually published every few months, so it is not possible to find out the impact that an individual act or decision has on society. As a partial solution, social media analyses may be able to track offline opinions through their online reflections.

As discussed below, there are many different automatic methods to extract opinions and trends from social media. In particular, the field of sentiment analysis [5], which is concerned with the automatic processing of subjective information, has made it possible to automatically detect opinions on a large scale. Whilst some studies have applied these methods to political topics, most research has focused on consumer reviews of products and most systems are

Corresponding author:
David Vilares, Universidade da Coruña, Facultade de Informática, Campus de Elviña, 15071 A Coruña, Spain
david.vilares@udc.es

designed exclusively for English text. There is therefore a need for political sentiment analysis systems as well as political sentiment analyses for languages other than English.

In response to the above gap, this article develops and applies a sentiment analysis system for Spanish political tweets with a case study of the main Spanish political representatives, and parties. The results are based on a set of 2,704,523 tweets mentioning 30 politicians and 6 political parties over 41 days. Due to the lack of free available real-time tools for sentiment analysis of Spanish tweets, we first enriched SentiStrength, a fast and widely used sentiment analysis algorithm, for Spanish political language. To evaluate the adaptation, we then created the first development and test Spanish corpora with dual scores representing both positive and negative sentiment strength. The results were used to analyse Spanish perceptions about their main political parties and representatives, relating them to important contemporary events. All the analyses in this article were conducted before the polling results were published and can therefore be considered to be real predictions.

2. Background and related work

Although most sentiment analyses of political communication have used general purpose sentiment analysis algorithms, these are sub-optimal for political communication. This is because the complexity of human communication cannot be easily described with a small set of rules, and so more elaborate techniques are needed for short political messages [6]. The results of most of the studies reviewed below may be affected to some extent by this issue.

2.1. Politics on social media

Some research has analysed the use of social media by politicians. For example, a Dutch study found that national election results correlated with politicians' use of social media, but the same was not true for local elections [7]. In contrast, an analysis of Australian politicians' use of Twitter argues that it is difficult to control, interpret or understand the benefit that they gain from it [8]. The remainder of this section discusses tweeting about politics by the electorate rather than by politicians.

Most political analyses on Twitter have focused on predicting electoral outcomes [9], but Twitter can also be used to identify political preferences [10] and for day-to-day monitoring of electoral campaigns [11, 12, 13]. One of the first studies analysed 104,003 Twitter messages mentioning the main German parties or politicians before the 2009 federal elections [14]. The German tweets were automatically translated into English for a LIWC keyword analysis [15]. The numbers of tweets mentioning parties or politicians were found to closely reflect voter preferences in traditional election polls. This study showed that Twitter may complement traditional polls as a political forecasting tool. Nevertheless, a Twitter sample may not be representative of the electorate, the general sentiment dictionaries used may not be optimal for politics, and replies to political messages may not be captured by keyword searches [14]. In support of the latter point, keyword-based searches for political tweets can aim for high precision (i.e., they generate few false matches) but not high recall (i.e., they will miss many relevant tweets) [16].

A time series analysis of the 2008 US presidential elections derived day-to-day sentiment scores by counting positive and negative messages: a message was defined as positive if it contained a positive word, and negative if it contained a negative word (a message can be both positive and negative). Although there were many falsely detected sentiments, these errors may tend to cancel out [17]. This approach missed sentiments in tweets using non-standard spellings and emoticons and needed smoothing to stabilise the results. The sentiment results correlated with presidential approval polls, but not with election polls, and message volume did not have a straightforward relationship with public opinion. Another study of US elections also found that sentiment results did not predict election outcomes, possibly due to the overrepresentation of young people and Democrats in Twitter [18]. Almost identically for the 2011 Irish General Election, sentiment did not predict voting patterns due to the overrepresentation of one party and the underrepresentation of another, although a simple volume measure was more accurate than sentiment [19]. Twitter bigrams (consecutive words) can also be used to predict daily approval ratings for US presidential candidates using a time series regression approach [20].

Election results have also been predicted in Twitter for many other countries, with varying degrees of success. For the 2013 general election in Italy tweet volume was a reasonable indicator for the final results, detecting a strong presence in Twitter of the (unexpected) winning party and the (also unexpected) relative weakness of another party, but failing to make accurate predictions for small parties, who were overrepresented in Twitter [21]. Small party overrepresentation in Twitter has also been found for German elections [22]. Tweet volumes were a reasonable indicator for the 2011 Nigerian Presidential election [23] and the Venezuelan, Paraguayan and Ecuadorian Presidential elections of 2013 – especially when counting tweets mentioning the full names of candidates or mentioning the aliases of candidates jointly with an electoral keyword [24]. In contrast, Twitter did not seem to be able to predict the 2011 Dutch Senate election outcomes [25].

One particularly comprehensive study analysed 542,969 tweets mentioning candidates together with data on 795 election outcomes in 2010 and 2012 US elections and socio-demographic and control variables such as incumbency, district partisanship, median age, percentage white, percent college educated, median household income, percentage female and media coverage [26]. There was a statistically significant association between the number of tweets mentioning a candidate and their subsequent electoral performance. The models under-performed in relatively uncompetitive or idiosyncratic districts, however.

Despite some of the positive results reported above, electoral predictions on Twitter data overall tend not to be better than chance [27]. When the predictions are better than chance (e.g., [28]), they are not an improvement on simply predicting that all incumbents would be re-elected (see also: [29]). It follows that sentiment analyses need to be sophisticated in order to make credible election predictions [30].

Several studies have applied sentiment analysis to social web politics and used the results to identify patterns of behaviour rather than to predict elections. It is possible to predict how Twitter users will vote by comparing the language of their tweets with that of the main parties in an election [31] and this technique has been used to show, unsurprisingly, that politically active users are less prone to changing their preferences. It is also possible to estimate the level of disaffection across society by counting negative tweets about politics in general, and this approach has shown that peaks in disaffection can correlate with important political news [32]. Twitter has also been used to study divisions within electorates. A sentiment analysis of Twitter in Pakistan, for example, found differences between expatriates and people living in the country and between urban and rural areas [33].

Finally, Twitter is also used by journalists to add direct quotes from politicians to stories, and so Twitter sometimes helps to generate the news in addition to reflecting it [34].

2.2. Twitter as a tool for political analysis in Spain

Twitter is extensively used in Spain for politics during elections [35]. An analysis of 370,000 tweets from over 100,000 users during the 2011 Spanish general elections found that half of the messages were posted by 7% of the participants, 1% of users were the target of half of the mentions, 78% of the mentions were of politicians, 2% of the users caused half of the retweets and the source of 63% of the retweeted messages were mass media accounts [36]. Moreover, 65% of the participants were men, with Madrid overrepresented but no significant differences were found between the behaviour of those living in large cities and in the rest of Spain; citizens with a strong party identification were particularly active [37].

A study of 84,387 tweets from Catalan regional elections found Twitter users to cluster by political affinity [38], corroborating similar results from other countries [39, 40, 41]. Despite this, it is difficult to predict the party of a Twitter user from the list of accounts that they follow [42]. Ideological groupings also occur on the web for political and media websites in Spain, highlighting the partisan nature of the media [43].

The number of times that Spanish political parties are mentioned on Twitter seems to correlate with electoral outcomes, but only for parties that obtained more than 1% of votes [36]. One study focused on predicting results for a new small, Spanish green party, eQuo, with an electoral strategy based mainly on social media [44]. For several days its proposals were trending topics on Twitter, and its Facebook page was more visited and had more “likes” than the pages of the other political parties. Nevertheless, this successful social media campaign did not translate into any elected politicians. Perhaps surprisingly, eQuo performed best in districts in which it used traditional activities, such as meetings and posters.

Twitter has been used for successful predictions for the Andalusian regional elections of 2012, counting the followers of the Twitter accounts of political parties and their leaders. For the two major parties, Partido Popular (PP) and Partido Socialista Obrero Español (PSOE), this simple method gave results that were closer to the final election outcomes than were traditional polls [45], although the polls were particularly inaccurate in these elections. The Twitter follower method was inaccurate for small and new parties, including those, such as the Izquierda Unida (IU), with leaders that were inactive on Twitter.

In an academic competition to classify the political tendency of public figures (not necessarily politicians) into left, centre, right or neutral [46], the best performing system considered a number of politicians related with the main political parties [47]. If messages from a user contained one of these politicians tended to be negative then the user was classified against that political orientation, and vice versa. Another competition was to classify the polarity of tweets mentioning one of the four main national parties. The best performing system assumed that the polarity of the whole tweet corresponded to the polarity of the party [48].

2.3. Sentiment classification on social media

A number of complex techniques have been proposed to try to overcome the main limitations of early simple sentiment analysis methods. Machine-learning approaches, although often effective, are highly domain dependent (i.e., they work well on tweets about one topic, but much less well on tweets about other topics), which is a drawback for multiple domain sentiment analysis tasks, where time-consuming or costly solutions might be necessary to build an accurate model [49]. Unsupervised lexicon-based methods [50] are an alternative way to obtain an acceptable level of performance for different domains. These approaches rely on dictionaries of subjective words labelled as positive or negative, perhaps also with a sentiment strength score. The algorithm SentiStrength [51], for example, is an unsupervised system to simultaneously estimate positive and negative sentiment strength in English short informal texts. SentiStrength applies a dual 5-point scheme for classifying texts: a score from 1 (no positivity) to 5 (very strong positivity) and a second score from -1 (no negativity) to -5 (very strong negativity). The system estimated positive and negative sentiment strength in MySpace comments with correlations of 0.60 and 0.56, respectively, with human judgements. SentiStrength was subsequently used to monitor popular events in Twitter, finding that the top media events tended to associate with small increases in average negative sentiment strength, even if they were positive [52]. An improved version of SentiStrength has been produced, with an expanded sentiment dictionary and idiom list as well as new rules to handle negation and other linguistic features. The new version gave good results across a variety of social media domains, including Twitter [53]. SOCAL is a similar system [54] that is intended for long reviews. Unlike SentiStrength, SOCAL has rules for *irrealis* (*e.g.*, conditional mood, subjunctive) and discourse structure features because some parts of a long text might be particularly important (e.g., the final sentences of a review). A Spanish version of SOCAL obtained an accuracy of 74.3% on a corpus of review texts [55]. Another unsupervised system for Spanish uses syntax rather than lexical knowledge [56]. Given a review, the syntactic structure of its sentences can be extracted with dependency parsing. A set syntactic rules can then identify intensification, negation and some adversative subordinated sentences, with an experimental accuracy of 78.5%.

3. Research questions

The aim of this article is to extend the state-of-the art with respect to harnessing Twitter to analyse Spanish political attitudes in the context of elections, driven by the following research questions.

- Is it possible to improve the accuracy of SentiStrength for Spanish tweets?
- Which are the key linguistic phenomena that impact on the accuracy of sentiment analysis in Spanish tweets?
- Can information from Twitter give insights into Spanish political reactions in a way that would complement traditional polls?

4. Spanish SentiStrength

The Spanish version of SentiStrength was used as the starting point for this study. This version was designed to help investigate political alignment and emotional expressions in Twitter [57, 46]. This article enriches SentiStrength and provides its first formal evaluation for Spanish. To achieve this goal, a human annotated Spanish corpus of tweets was created and the existing linguistic resources for SentiStrength were adapted.

4.1. Corpus description

In order to evaluate SentiStrength, a human-annotated corpus is needed, with each text given a positive (1 to 5) and a negative (-1 to -5) sentiment strength score. Although formal evaluations of SentiStrength exist for English [53] and German [58], there do not seem to be any published formal evaluations for Spanish, nor any human-coded Spanish corpora with positive and negative sentiment strength scores. In response, we created development and test sets with a random number generator from a large collection of Spanish tweets downloaded from the Twitter API (Applications Programming Interface) in September 2014.

Development set: A collection of 1,600 tweets was labelled by the first author. The corpus was used to explore ways of improving the performance of SentiStrength, as explained in the following section.

Test set: A collection of 1,600 tweets was manually labelled by Spanish linguists that were not part of this project. To identify reliable coders, we first asked seven people to annotate a common set of 160 tweets. We then selected the three linguists that coded most consistently against each other, with Krippendorf's alpha coefficient of inter-coder consistency varying from 0.630 to 0.701 for negative sentiment and 0.625 to 0.726 for positive sentiment. These three annotators were then asked to independently label the 1,600 tweets of the test set, obtaining consistency from 0.486 to 0.660 for negative sentiment and 0.503 to 0.678 for positive sentiment. Three different strategies were used to combine the scores: the average; the maximum (assuming that annotators tend to be conservative); and the average after removing the minimum. This resource has been released online for the research community.[1]

4.2. SentiStrength Adaptations

Sentiment dictionary: Lists of sentiment-related terms form the core of lexical sentiment analysis algorithms such as SentiStrength. Although there are several Spanish sentiment lexicons, they mainly distinguish between positive and negative words (e.g., [59, 60]) or do not use the same sentiment scale as SentiStrength. The existing SentiStrength Spanish dictionary contained 1,409 subjective terms, each annotated with a sentiment strength of 1 to 5 or -1 to -5. The terms were mainly derived from LIWC [15], which contains a psychological text analysis resource with a Spanish variant. Additional terms were also added by SentiStrength's commercial users[2] and from Spanish translations of other English resources [61, 62]. We extended this sentiment dictionary with the following resources.

A new dictionary of subjective adjectives, nouns, verbs and adverbs [55] was consulted for new terms to add. These terms were already scored 1 to 5 and from -1 to -5 for positive and negative sentiment strength. SentiStrength does not use typical natural language processing steps, such as lemmatisation, part-of-speech tagging or dependency parsing, because of its focus on short informal texts, such as tweets, in which non-standard spelling and grammar probably occur in the majority of texts. This is a problem for Spanish, however, because most nouns and adjectives vary with gender and number, which makes the direct matching of dictionary words to text words more difficult. For example, the Spanish dictionary might need to include four versions of most nouns to cope with their gender and number variations. Verb inflections differ even more than nouns but the problem is the same. In response, the dictionaries were expanded to include the main word form variants in each case. We did not include *irrealis* conditional and subjective forms, however, since they need to be processed differently [54, 63].

The improved version of SentiStrength with the added Brooke et al. dictionaries [55] was applied to the development set and discrepancies in the results analysed to identify additional modifications. This resulted in the addition of new terms as well as the removal of wrong terms, such as homonyms with sentiment and non-sentiment meanings, and modifications of some of the original term strengths. The final Spanish lexicon was almost 20 times larger, with 26,752 entries (5,728 word forms plus verb, noun and adjective inflections).

- **Emoticons:** SentiStrength uses a list of 115 traditional emoticons to help detect sentiment (see also: [64, 65]). We added 71 new emoticons (e.g. ♥.♥) from the development set.
- **Idioms:** SentiStrength's idiom list contains 3 stock Spanish phrases that have a different sentiment than the individual words (e.g., *Que tal* scores +2). It was expanded with 306 extra expressions, many involving verbs and their inflections.
- **Slang words:** Although the English version of SentiStrength employs a slang conversion table, the Spanish version did not. A list of common Spanish slang and abbreviations was therefore created from the development set with the help of the Ancora corpus [66]. Many of the abbreviations had subjective connotations, such as *tkm* (te quiero mucho - I love you) and *bs* (beso - kiss). Some frequent Anglicisms were also translated into Spanish (e.g., *VIP* - persona muy importante). The new slang list is larger than other Spanish slang collections in other studies [67].
- **Dictionary:** This collection of words that are part of the language is used within SentiStrength's spelling correction algorithm. It was expanded with the Ancora corpus.
- **Booster words:** Spanish SentiStrength contained ten booster words to amplify or diminish the sentiments of subjective words. We expanded it using another booster collection [55] and new terms from development set, mainly from South American Spanish (e.g., *re* and *so*), giving 169 terms.
- **Question words:** Spanish SentiStrength used five words (e.g., *qué*) to identify questions, which have modified sentiment rules. This list was extended to 20 terms. Acute accents are important for these because their absence

turns many question words into conjunctions but they are usually omitted in Twitter so the unaccented word forms were also included.

The new version of SentiStrength is available free for research from the first and second authors and can be tried out online at http://sentistrength.wlv.ac.uk/SpanishSentiDataDavidVilares.zip.

4.3. SentiStrength Evaluation

SentiStrength was evaluated by comparing it to the results of the human coders on the test set of tweets (i.e., the gold standard). The optimal metric for comparisons is the simple Pearson correlation because this reflects the closeness between the prediction and the true value in cases where the prediction is not perfect. For completeness and comparisons with other systems, we also report the percentage of scores that equal the gold standard (i.e., precision), the percentage of scores that differ by at most 1 from the gold standard (called +/-1), and the trinary accuracy. The trinary metric uses only three classes: positive, neutral/mixed (either no subjectivity or equal positive and negative scores) and negative.

The new version of SentiStrength substantially outperforms the original version both for positive and negative sentiment in terms of the key correlation metric, performs moderately better on the trinary metric and performs slightly better overall for precision but slightly worse for +/-1 (Table 1). The reason for the moderately worse performance on negative sentiment +/-1 is probably because the old version of Spanish SentiStrength had a relatively small set of negative terms and mostly assigned the minimum no negative score. In fact the strategy of assigning -1 to all tweets would get a high score in the +/-1 metric (89.6% on average test set) because of the relative scarcity of negative tweets.

Table 1. The performance of the two Spanish SentiStrength versions (default set up) on the test set of 1,600 human-coded tweets.

Test set	SentiStrength version	Positive correlation	Positive correct	Positive +/-1	Negative correlation	Negative correct	Negative +/-1	Trinary evaluation
Average	New	**0.437**	**51.4%**	**79.9%**	**0.421**	**63.4%**	86.2%	**54.5%**
	Old	0.304	47.6%	79.8%	0.351	63.1%	**89.0%**	50.9%
Maximum	New	**0.437**	**47.3%**	**78.1%**	**0.423**	**52.4%**	79.3%	**52.5%**
	Old	0,326	44,3%	76.1%	0.349	51,6%	**80.1%**	49,8%
Remove minimum	New	**0.437**	**51.3%**	79.0%	**0.417**	**63.6%**	84.5%	**55.2%**
	Old	0.294	46.7%%	78.7%	0.341	62.0%	**86.6%**	51.1%

SentiStrength includes a number of options to configure the behaviour of the algorithm [53], such as the maximum number of terms allowed between a negating word and a sentiment word. We used the development set to assess whether changes in these parameters could improve the overall SentiStrength results. We optimised the choice of parameters via a greedy search based upon performance (using the main correlation metric) on the development set. For example, the experiments on the development set indicated that flipping negated negative words to positive was better than neutralising them.

Table 2 details the performance of the new SentiStrength on the test set with different running configurations, activating or deactivating one option at a time. For simplicity, the remaining results use only the average of the scores of the three annotators. Most variations in performance are small. Table 2 also shows that the best configuration on the development set (the one activating the option *Negating negative neutralises emotion*) achieved the highest positive correlation and an acceptable performance on the rest of the metrics. This reinforces the competitiveness of this model for analysing real problems.

Table 2. The performance of Spanish SentiStrength on the test set (average scores) with a variety of options individually enabled.

Option disabled	Positive correlation	Positive correct	Positive +/-1	Negative correlation	Negative correct	Negative +/-1
Spelling correction	0.416	52.312	80.062	0.406	63.625	86.438
Questions reduce negative sentiment	0.436	51.437	79.812	0.409	**64.75**	**87.688**
Multiple letters boost sentiment	0.434	**52.375**	80.312	0.423	63.437	86.312
Booster list	0.433	51,750	**80.625**	0.415	63.462	86.625
Dictionary	0.427	51.812	79.875	0.410	63.562	86.375
Multiple positive words are boosted	0.437	51.562	80.125	0.421	63.375	86.188
Emoticon list	0.437	51.437	79.938	0.421	63.375	86.188
Negating positive neutralises emotion	0.435	50.500	79.500	**0.424**	63.937	86.375
Multiple negative words are boosted	0.437	51.437	79.938	0.421	63.375	86.25
Ignore booster words after negators	0.437	51.437	79.938	0.421	63.375	86.188
Exclamation marks count as +2	0.437	51.375	79.938	0.421	63.375	86.188
Negating negative neutralises emotion	**0.438**	51.500	79.938	0.417	62.375	85.625
Idiom list	**0.438**	51.500	79.875	0.416	63.250	86.188
None (default configuration)	0.437	51.437	79.938	0.421	63.375	86.188

Finally, political tweets from November 2014 were used to extend and adapt the new SentiStrength for politics. For example, *todos son iguales* (they are all the same) and *dimisión* (resignation) are common negative political expressions. The resulting political resources for Spanish SentiStrength are available free.[3]

5. Tweets about Spanish politicians

This section explores Spanish political tweets. The aim is not to predict elections, but to assess whether Twitter can reveal changing perceptions about politicians over time and the influence of individual events. For each of the six main political parties we selected five important politicians (see Appendix A). Since some parties have few widely recognised members [4], we took into account the number of Twitter followers to ensure that the selected politicians could be discussed on Twitter.

- *Partido Popular (PP, @Ppopular)*: The main conservative party and winner of the 2011 elections. Its leader and prime minister is Mariano Rajoy (@marianorajoy).
- *Partido Socialista Obrero Español (PSOE, @PSOE)*: The main social-democratic party and in government until 2011. Its secretary-general is Pedro Sánchez (@sanchezcastejon).
- *Izquierda Unida (IU)*: A left-wing party and usually third in general elections. Its current leader, Cayo Lara, (@cayolara) was set to step down and be replaced by Alberto Garzón (@agarzon).
- Unión, Progreso y Democracia (UPyD): A young political party founded in 2007. The leader, Rosa Díez, is the main politician without an official Twitter account.
- Ciudadanos (Cs): A non-regionalist centre party originally from Catalonia and led by Albert Rivera (@Albert_Rivera).
- Podemos: A new left wing political party from January 2014. The elected leader is Pablo Iglesias (@Pablo_Iglesias_) and at least one poll has rated them as the most popular Spanish party [4].

A number of steps were taken to filter out irrelevant tweets.

- Tweets just containing information without an opinion (i.e., classified as (-1,1)), were removed.
- Retweets of tweets from the parties or politicians analysed were removed. This step was taken because these messages tend to be retweeted many times due to their number of followers and the author of the message and therefore seem to create false peaks in activity.
- Messages involving two or more different political parties were removed.
- Phrases quoted in tweets were removed because these are often associated with titles or rhetorical devices, such as sarcasm or irony, that should be treated differently [53].

6. Experiments

For the political analysis experiment, we collected tweets from 3 December 2014 to 12 January 2015 via the Twitter Streaming API. After the above filtering steps there were 2,704,523 tweets and daily average positive and negative sentiment scores were computed for each politician and party.

Polls from the reputable Centro de Investigaciones Sociológicas (CIS, www.cis.es) [4] were taken as the primary source of public opinion in Spain and used as the reference point for the Twitter results. In order to cover a wider set of entities, additional well-known polls carried out by private companies were also included: Invymark (www.invimark.es), GESOP (www.gesop.net), DYM (www.gesop.net), Sigma-2 (www.sigmados.com) and Termómetro electoral (termometroelectoral.blogspot.com.es). The CIS poll used covered January 4 to January 12, 2015.

Table 3 compares the sentiment for the political leaders with the national poll results. Polls differ in their criteria and coverage, and the CIS poll includes only four leaders. The dual positive and negative SentiStrength scores make it possible to assess whether the most hated leaders are also the most loved, and whether some politicians attract particularly strong emotions. The ranking provided by SentiStrength for positivity matches exactly that provided by CIS. Surprisingly, SentiStrength's negativity rank is also similar to the one provided by CIS, switching only third and fourth place. The similarity between rankings can be compared with Hamming loss distance and the out-of-place measure [68] (Table 4) according to equations (2) and (3). For example, given a ranking, *R*, where the system only failed the classification for the entity *E: predicted(E) = 5* and *gold(E) = 2*, the score for the *Hamming-loss(R)* would be 2, since we need to make two changes to obtain the correct ranking, while the *out-of-place(R)* would be 3, because that was difference between the predicted and gold position. However, if *predicted(E)* is 3 then *out-of-place(R)* would be 1, although the Hamming loss would be the same.

$$f(x,y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases} \tag{1}$$

$$Hamming\ loss\ distance\ = \sum_{i=1}^{n} f\big(predicted(i), gold(i)\big) \tag{2}$$

$$Out\ of\ place = \sum_{i=1}^{n} |Predicted(i) - Gold(i)| \tag{3}$$

The SentiStrength scores were used to create daily rankings, and Mann-Whitney U tests were used to compare the relative rankings for pairs of politicians, as shown in Table 5. A significant result is evidence that the daily SentiStrength averages consistently indicate one of the two politicians as being the most popular, with a better than random chance.

Surprisingly, the leaders of the first and second ranked parties [4] were the only ones with stronger negative sentiment than positive sentiment. This suggests that negative expressions in Twitter do not imply less electoral support but could reflect other factors, such as the need for other parties to attack the leading contenders.

Table 3 also compares our ranking with the ones provided by the other polls collected. Table 4 measures how similar these rankings are to the ones provided by the main national polls, except for Sigma-2. Some of the polls were not published in January, so we substituted the ones released in December, which also fell within our period of analysis. Positive perception is a better indicator than negative perception in most of the cases to predict rankings similar to those of the traditional polls. The comparison shows that centre-right leaders (Rivera, Díez and Rajoy) are equivalently located in all polls (except Sigma-2). The differences in rankings are due to left-wing leaders (Sánchez, Garzón and Iglesias), an issue also observed when comparing traditional polls between them.

Table 3. Average SentiStrength scores and national Spanish poll scores. Reference polls ratings range from 0 to 10.

Leader	Positive SentiStrength	Negative SentiStrength	CIS (Jan.)	Termómetro Electoral (Jan.)	Invymark (Jan.)	GESOP (Dec.)	Sigma-2 (Dec.)	DYM (Dec.)
Albert Rivera	2.57	1.80	-	4.73	4.18	4.62	3.76	4.20
Pedro Sánchez	2.30	1.88	3.68	4.02	4.18	4.56	3.81	3.71
Alberto Garzón	2.25	2.03	-	4.52	3.89	4.60	3.80	3.90
Pablo Iglesias	2.22	2.23	-	4.51	3.96	4.51	3.93	4.00
Rosa Diéz	2.19	2.16	3.66	3.02	3.54	4.15	3.87	3.70
Cayo Lara	2.13	2.02	3.53	-	-	-	-	-
Mariano Rajoy	2.07	2.32	2.24	2.81	3.27	2.65	3.43	2.60

Table 4. Predicted and gold standard rankings compared to Hamming-loss distance and out-of-place measure.

Metric	SentiStrength	CIS	Termometro electoral	Invymark	GESOP	Sigma-2	DYM
Hamming-loss distance	Positive	0	3	3	2	5	2
	Negative	2	4	2	3	5	3
Out-of-place measure	Positive	0	4	2	2	12	4
	Negative	2	6	4	4	12	6

Table 5. Mann-Whitney U test, at a confidence level of 95%. (p < 0.05), for the main Spanish political leaders. This shows the complete results for the best two scored leaders (Albert Rivera and Pedro Sanchez), although Bonferroni corrections were applied to counteract the problem of multiple comparisons taking into account all possible comparisons between the pairs of leaders (p < 0.002381 to accept that differences in perception are significant). Cells marked with * indicate significant differences.

	Albert Rivera		Pedro Sanchez		Alberto Garzón		Pablo Iglesias		Rosa Díez		Cayo Lara		Mariano Rajoy	
	P	N	P	N	P	N	P	N	P	N	P	N	P	N
Albert Rivera	-	-	0.0026	0.0500	0.0011*	0.0091	0.0004*	0.0000*	0.0023*	0.0002*	0.0000*	0.0170	0.0000*	0.0000*
Pedro Sánchez	-	-	-	-	0.3050	0.0347	0.1069	0.0000*	0.1149	0.0015*	0.5843*	0.1360*	0.0000*	0.0000*

With respect to political parties, it does not make sense to predict poll results with sentiment because the two are different. For example, according to a CIS (2015) poll [4], Mariano Rajoy and the Partido Popular are the least popular leader and party, but would get the most votes. As shown in Table 6, sentiment scores are not reliable for predicting elections. The number of tweets naming either the political party or their leader is a better indicator, confirming similar results for other countries [14, 21].

In general, the more conservative the party is, the more negative tweets mention it and this may reflect a bias in the user base of Twitter, such as towards young people. Younger people may tend to be left-wing [69], which would explain the online hostility to the right. Left wing young people may also be more politically active [70], exacerbating the bias. In Spain, according to the CIS poll 40% of the population are left-wing and 21% are right-wing, and so negativity towards the right could be expected even without the youth bias.

Table 6. Average positive and negative sentiment strength in tweets mentioning the main Spanish political parties.

Party	Positive sentiment	Negative sentiment	Left- (0) right-wing (10) (CIS, 2014)	Vote intention + Sympathy (CIS, 2014)	Vote estimate (CIS, 2014)	Average daily party mentions in Twitter	Average daily leader mentions in Twitter
@CiudadanosCs	2.59	-1.63	5.14	1.9	3.1	963	893
@ahoraPodemos	2.22	-2.05	2.28	19.3	23.9	19495	5353
@UPyD	2.07	-2.22	5.34	2.5	4.6	2218	289
@PSOE	2.11	-2.13	4.62	18.1	22.2	4881	2957
@iunida	2.04	-2.15	2.62	4.2	5.2	542	623
@Ppopular	2.05	-2.48	8.17	14.6	27.2	2858	14007

Table 7 shows the rankings for all politicians. Traditional polls do not provide surveys for many of these, and so the ranking cannot be compared with other rankings. Nevertheless, they give information that cannot be obtained from traditional polls.

The Partido Popular politicians are last in both rankings, reinforcing the online sentiment agreement with traditional polls. Similarly, Ciudadanos politicians had the highest scores, which reflects the results for their leader, Albert Rivera. These results also show that politicians coming from these two parties attract similar sentiments to their party overall. The same is true for PSOE politicians, except that the party account is an outlier.

Politicians from Podemos, Izquierda Unida and UPyD were more scattered in rakings, but these might reflect specific news stories with wide media coverage. The low average positive sentiment for Tania Sánchez (IU) was perhaps reflected by her resignation shortly after the period of analysis. Negative press coverage about her management of a town council and disagreements with IU in December [71] and January [72, 73] seemed to trigger her departure.

Iñigo Errejón (Podemos) also had low sentiment rankings apparently as a consequence of negative press coverage. Figure 1 shows how during the beginning of the polling period he had very low scores, coinciding with news about alleged irregularities at his previous job from December 4, 2014 [74, 75]. A peak in the number of tweets at this time confirms that Twitter sometimes reflects popular political events.

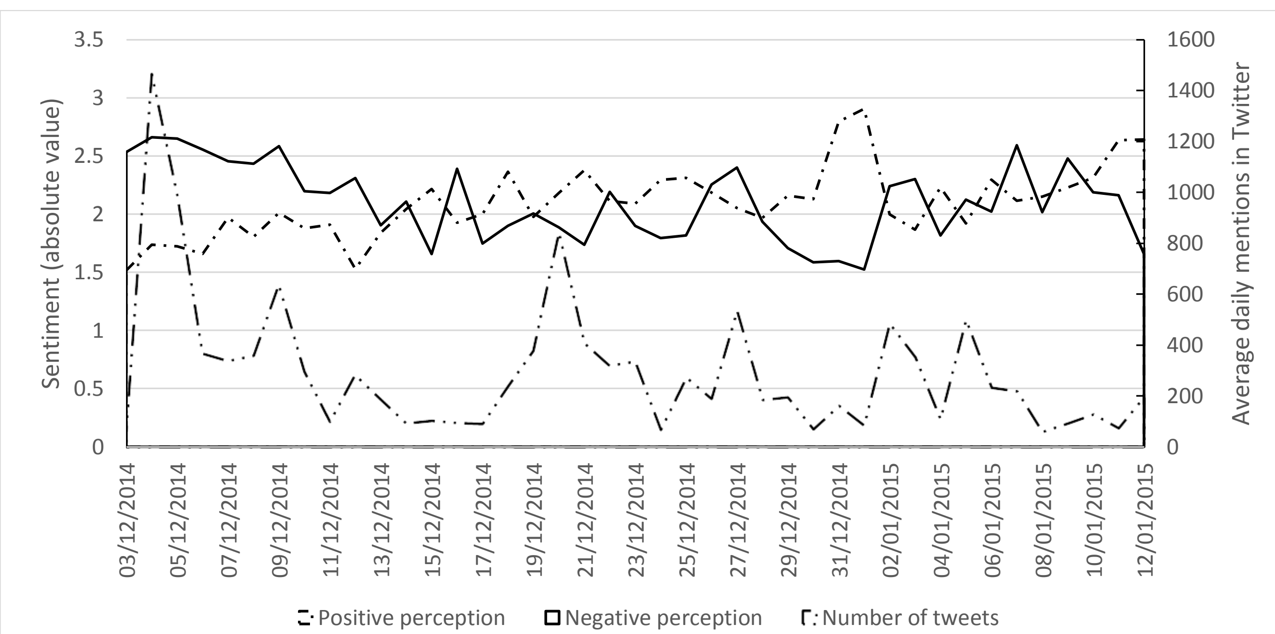

Figure 1. Variation of positive and negative perception of Iñigo Errejón during the period of polling.

Table 7. Positive and negative sentiment rankings from SentiStrength for tweets mentioning the politicians analysed.

Negative sentiment (from the lowest to the highest)			Positive sentiment (from the lowest to the highest)		
Luis Salvador	Cs	-1.39	Luis Salvador	Cs	2.84
Fran Hervias*	Cs	.1.42	Fran Hervias*	Cs	2.63
Ines Arrimadas*	Cs	-1.59	Ciudadanos	Cs	2.59
Ciudadanos	Cs	-1.63	Albert Rivera	Cs	2.57
Carme Chacón*	PSOE	-1.80	Ines Arrimadas*	Cs	2.43
Albert Rivera	Cs	-1.80	Pablo Echenique	Podemos	2.31
César Luena*	PSOE	-1.85	Pedro Sánchez	PSOE	2.30
Pedro Sánchez	PSOE	-1.88	Elena Valenciano*	PSOE	2.28
Susana Díaz	PSOE	-1.92	Cristina Cifuentes	PP	2.28
Elena Valenciano*	PSOE	-1.97	Carme Chacón*	PSOE	2.26
Pedro Echenique	Podemos	-2.02	Alberto Garzón	IU	2.25
Cayo Lara	IU	-2.02	Irene Lozano	UPyD	2.24
Alberto Garzón	IU	-2.03	César Luena*	PSOE	2.23
Irene Lozano	UPyD	-2.03	Teresa Rodríguez	Podemos	2.22
Podemos	Podemos	-2.05	Pablo Iglesias	Podemos	2.22
Carlos Martínez Gorriaran	UPyD	-2.08	Podemos	Podemos	2.22
Teresa Rodríguez	Podemos	-2.10	Esperanza Aguirre	PP	2.20
Iñigo Errejón	Podemos	-2.10	Juan Carlos Monedero	Podemos	2.20
Toni Cantó	UPyD	-2.10	Rosa Díez	UPyD	2.20
PSOE	PSOE	-2.13	Javier Nart	Cs	2.18
Gaspar LLamazares*	IU	-2.14	Carlos Martínez Gorriarán	UPyD	2.17
Izquierda Unida	IU	-2.15	Soraya Sáenz de Santamaría	PP	2.16
Juan Carlos Monedero	Podemos	-2.16	Hugo Martínez Abarca*	IU	2.16
Tania Sánchez	IU	-2.16	Toni Cantó	UPyD	2.15
Rosa Díez	UPyD	-2.16	Susana Díaz	PSOE	2.15
Javier Nart	Cs	-2.18	Cayo Lara	IU	2.13
Hugo Martínez Abarca*	IU	-2.19	PSOE	PSOE	2.10
Cristina Cifuentes	PP	-2.20	Iñigo Errejón	Podemos	2.08
Esperanza Aguirre	PP	-2.20	Mariano Rajoy	PP	2.07
UPyD	UPyD	-2.22	UPyD	UPyD	2.07
Pablo Iglesias	Podemos	-2.23	Partido Popular	PP	2.05
Soraya Sáenz de Santamaría	PP	-2.27	Izquierda Unida	IU	2.04
Mariano Rajoy	PP	-2.32	María Dolores de Cospedal	PP	2.04
Partido Popular	PP	-2.48	Gaspar Llamazares*	PP	2
María Dolores de Cospedal	PP	-2.65	Tania Sánchez	IU	1.98

* this politician has less than 120 mentions per day (20% of the maximum), and so may have an unreliable ranking.

7. Conclusions

This article extended the sentiment strength detection program SentiStrength for Spanish and for Spanish political tweets and then analysed the sentiments expressed in a month of tweets about the main Spanish politicians and parties. Mainly based on extending its lexical resources, the new SentiStrength gives substantially more accurate results. This is based on an evaluation using a new test set of Spanish tweets coded by three separate independent linguists.

The improved version of Spanish SentiStrength was used to analyse tweets about the main political parties and leaders of Spain. The sentiment scores obtained by SentiStrength were used to build ranks for the politicians and their parties, giving popularity ratings that are comparable with those provided by the classic polls, although tweet volume was a much better predictor of voting intentions. An advantage of sentiment analysis in Twitter is that is can be more comprehensive than traditional polls by covering more parties and politicians, although the results are less reliable for smaller parties. A deeper analysis of politicians that had sentiment scores that did not match those of their parties suggested that these had attracted negative media publicity that had been amplified in Twitter. This shows that the Twitter results may be useful to analyse the trajectories of individual politicians and perhaps even to evaluate the impact of negative press coverage on their popular perception.

With respect to future work in Spain, SentiStrength needs to be adapted for other Iberian languages in order to provide a more complete analysis of Spanish politics. Many regional parties have a presence in the Spanish parliament and play a key role in political decisions, but many tweets about them are in Basque, Catalan or Galician. Integrating language identification techniques for Iberian languages in Twitter [76] and machine translation techniques for lexicon translation should help with this. It would also be interesting to systematically analyse anomalies in the results, such as politicians that attract different sentiments to those of their parties, in order to assess their long term impact on politics. It would also be interesting to study the propagation of tweets bearing political content by applying a technique similar to the one described in [77].

Funding

This research is partially supported by the Ministerio de Economía y Competitividad y FEDER (FFI2014-51978-C2) and Xunta de Galicia (R2014/034). The first author is partially funded by the Ministerio de Educación, Cultura y Deporte (FPU13/01180) and Inditex-UDC 2014 grants for international visits. Thank you to the annotators for help with this research.

References

[1] Harfoush, R. (2009). Yes We Did! An inside look at how social media built the Obama brand. New Riders.

[2] Howard, P. N., Duffy, A., Freelon, D., Hussain, M., Mari, W., & Mazaid, M. (2011). Opening closed regimes: what was the role of social media during the Arab Spring?

[3] Borge-Holthoefer, J., Rivero, A., García, I., Cauhé, E., Ferrer, A., Ferrer, D., & Moreno, Y. (2011). Structural and dynamical patterns on online social networks: the Spanish May 15th movement as a case study. PloS one, 6(8), e23883.

[4] CIS. (2014). Barómetro de Octubre 2014. Technical Report 3041, Centro de Investigaciones Sociológicas.

[5] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, *2*(1-2), 1-135.

[6] Chung J., & Mustafaraj, E. (2011). Can collective sentiment expressed on Twitter predict political elections? Proceedings of the Twenty-Fifth Conference on Artificial Intelligence (AAAI-11) (pages 1170–1171), San Francisco, CA: AAAI.

[7] Effing, R., van Hillegersberg, J., & Huibers, T. (2011). Social media and political participation: are Facebook, Twitter and YouTube democratizing our political systems?. In Electronic participation (pp. 25-35). Springer Berlin Heidelberg.

[8] Grant, W. J., Moon, B., & Busby Grant, J. (2010). Digital dialogue? Australian politicians' use of the social network tool Twitter. *Australian Journal of Political Science*, *45*(4), 579-604.

[9] Ceron, A., Curini, L., & Iacus, S. (2014). Using social media to forecast electoral results. A meta-analysis. Italian Journal of Applied Statistics, Forthcoming.

[10] Golbeck, J., & Hansen, D.L. (2011). Computing political preference among Twitter followers. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011), (pp. 1105–1108), Vancouver, BC, Canada: ACM.

[11] Jensen, M.J., & Anstead, N. (2013). Psephological investigations: Tweets, votes and unknown unknowns in the Republican nomination process. Policy & Internet, 5(2), 161–182.

[12] Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S (2012). A system for real-time Twitter sentiment analysis of 2012 U.S. presidential election cycle. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL 2012) (pp. 115–120). Jeju, Republic of Korea: ACL.

[13] Ceron, A., Curini, L., & Iacus, S. M. (2015). Using Sentiment Analysis to Monitor Electoral Campaigns Method Matters—Evidence From the United States and Italy. Social Science Computer Review, 33(1), 3-20.

[14] Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *ICWSM*, *10*, 178-185.

[15] Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates*, *71*.

[16] Marchetti-Bowick, M., & Chambers, N. (2012). Learning for microblogs with distant supervision: Political forecasting with Twitter. Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, pages 603–612, Avignon, France: ACL.

[17] O'Connor, B., Balasubramanyan, R., Routledge, B., & Smith, N.A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (ICWSM 2010) (pp. 122–129), Washington, DC: AAAI.

[18] Gayo-Avello, D. (2011). Don't turn social median into another 'literary digest' poll. Communications of the ACM, 54(10), 121–128.
[19] Bermingham A., & Smeaton, A.F. (2011). On using Twitter to monitor political sentiment and predict election results. In S. Bandyopadhyay, & M. Okumurra (Eds.), Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011) (pp. 2-10). Chiang Mai, Thailand: Asian Federation of Natural Language Processing.
[20] Contractor, D., & Faruquie, T.A. (2013). Understanding election candidate approval ratings using social media data. In WWW 2013 Companion (WWW 2013) (189–190), Rio de Janeiro: ACM Press.
[21] Caldarelli, G., Chessa, A., Pammolli, F., Pompa, G., Puliga, M., Riccaboni, M., & Riotta, G. (2014). A multi-level geographical study of Italian political elections from Twitter data. PloS one, 9(5), e95809.
[22] Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment", Social Science Computer Review, 30(2), 229–234.
[23] Fink, C., Bos, N., Perrone, A., Liu, E., & Kopecky, J. (2013). Twitter, public opinion, and the 2011 Nigerian presidential election. In Proceedings of SocialCom / PASSAT / BigData / EconCom / BioMedCom 2013 (SocialCom 2013) (pp. 311–320), Washington, D.C., USA: IEEE Computer Society.
[24] Gaurav, M., Srivastava, A., Kumar, A., & Miller, S. (2013). Leveraging candidate popularity on Twitter to predict election outcome. Proceedings of the 7th Workshop on Social Network Mining and Analysis (SNA-KDD 2013) (Article No. 7), Chicago, IL: ACM.
[25] Sang, E.T.K., & Bos, J. (2012). Predicting the 2011 Dutch senate election results with Twitter. Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics, pages 53–60, Avignon, France: ACL.
[26] DiGrazia, J., McKelvey, K., Bollen, J., & Rojas, F (2013). More tweets, more votes: Social media as a quantitative indicator of political behavior. PLOS ONE, 8(11), e79449.
[27] Metaxas, P.T., Mustafaraj, E. & Gayo-Avello, D. (2011). How (not) to predict elections. Proceedings of 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing (PASSAT/SocialCom 2011) (pp. 165–171), Boston: IEEE Computer Society.
[28] Gayo-Avello, D., Metaxas, P.T., and Mustafaraj, E. (2011). Limits of electoral predictions using Twitter. Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011) (pp. 490–493), Barcelona, Spain: AAAI.
[29] Huberty, M. (2013). Multi-cycle forecasting of congressional elections with social media. Proceedings of the 2nd workshop on Politics, elections and data (PLEAD 2013) (pp. 23–29), San Francisco, CA: ACM.
[30] Gayo-Avello, D (2012). No, you cannot predict elections with Twitter. IEEE Internet Computing, 16(6), 91–94.
[31] Makazhanov, A., & Rafiei, D. (2013). Predicting political preference of Twitter users. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013) (pp. 298–305), Niagara, ON, Canada: ACM.
[32] Monti, C., Rozza, A., Zapella, G., Zignani, M., Arvidsson, A., & Colleoni, E. (2013). Modelling political disaffection from Twitter data. Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM 2013) (Article No. 3), Chicago, IL: ACM.
[33] Razzaq, M.A., Qamar, A.M., & Bilal, H.S.M. (2014). Prediction and analysis of Pakistan election 2013 based on sentiment analysis. In X. Wu, M. Ester, & G. Xu (Eds.), Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014) (pp. 700–703), Beijing, China: IEEE.
[34] 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.
[35] Criado, J.I., Martñinez-Fuentes, G. &Silván, A. (2013). Twitter en campaña: las elecciones municipales españolas de 2011. Revista de investigaciones Políticas y Sociológicas, 12(1), 93-113.
[36] Borondo, J., Morales, A.J., Losada, J.C., & Benito, R.M. (2012). Characterizing and modeling an electoral campaign in the context of Twitter: 2011 Spanish presidential election as a case study. AIP Chaos, 22(2), 023138.
[37] Barberá, P., & Rivero, G. (2012) ¿Un tweet, un voto? Desigualdad en la discusión política en Twitter. I Congreso Internacional en Comunicación Política y Estrategias de Campaña, Madrid: ALICE.
[38] Congosto, M.L., Fernández, M, & Moro Egido, E. (2011). Twitter y política: Información, opinión y ¿predicción? Cuadernos de Comunicación Evoca, 4, 11-15.
[39] Livne, A., Simmons, M.P., Adar, E., &Adamic, L.A. (2011). The party is over here: Structure and content in the 2010 election. Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011) (pp. 201–208), Barcelona, Spain: AAAI.
[40] Conover, M.D., Gonçalves, B., Flammini, A., & Menczer, F. (2012). Partisan asymmetries in online political activity. EPJ Data Science, 1(1), Article: 6.
[41] Conover, M.D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Flammini, A., & Menczer, F. (2011). Political polarization on Twitter. Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011) (pp. 89–96), Barcelona, Spain: AAAI.
[42] Barberá, P. (2012). A new measure of party identification in Twitter. Evidence from Spain. 2nd Annual General Conference of the European Political Science Association (EPSA), Berlin: EPSA.

[43] Romero-Frías, E. & Vaughan, L. (2012). Exploring the relationships between media and political parties through web hyperlink analysis: The case of Spain. Journal of the American Society for Information Science and Technology, 63(5), 967–976.
[44] Deltell, L. (2012). Estrategias de comunicación política en las redes sociales durante la campaña electoral del 2011 en España: el caso de eQuo. II Jornadas de la Asociación Madrileña de Sociología, Madrid: AMS.
[45] Deltell, L., Claes, F., & Osteso, J.M. (2013). Predicción de tendencia política por Twitter: Elecciones andaluzas 2012. Ambitos: Revista internacional de comunicación, 22, 91–100.
[46] Villena-Román, J., & García-Morera, J., (2013). TASS 2013-Workshop on Sentiment Analysis at SEPLN 2013: An overview. In Díaz-Esteban, A., Alegría I., & Villena-Román J. (Ed.), Proceedings of the TASS workshop at SEPLN 2013 Actas del XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural. IV Congreso Español de Informática (pp. 112–125). Madrid: SEPLN.
[47] Pla, F., & Hurtado, L. F. (2013). ELiRF-UPV en TASS-2013: Análisis de sentimientos en Twitter. In Díaz-Esteban, A., Alegría I., & Villena-Román J. (Ed.), Proceedings of the TASS workshop at SEPLN 2013 Actas del XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural. IV Congreso Español de Informática (pp. 112–125). Madrid: SEPLN.
[48] Gamallo, P., García, M., & Fernández-Lanza S. (2013) TASS 2013 – A Naive-Bayes strategy for sentiment analysis on Spanish tweets. In Díaz-Esteban, A., Alegría I., & Villena-Román J. (Ed.), Proceedings of the TASS workshop at SEPLN 2013 Actas del XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural. IV Congreso Español de Informática (pp. 126–132). Madrid: SEPLN.
[49] Aue, A., & Gamon, M. (2005). Customizing sentiment classifiers to new domains: A case study. In *Proceedings of recent advances in natural language processing (RANLP)* (Vol. 1, No. 3.1, pp. 2-1).
[50] Turney, P.D. (2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th annual meeting on association for computational linguistics (ACL'02) (pp. 417-424). Philadelphia: ACL.
[51] Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, *61*(12), 2544-2558.
[52] 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.
[53] Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*(1), 163-173.
[54] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, *37*(2), 267-307.
[55] Brooke, J., Tofiloski, M., & Taboada, M. (2009). Cross-Linguistic Sentiment Analysis:From English to Spanish. In Proceedings of International Conference on Recent Advances in Natural Language Processing (RANLP) (pp. 50–54). Borovets, Bulgaria.
[56] Vilares, D., Alonso, M. Á., & Gómez-Rodríguez, C. (2015). A syntactic approach for opinion mining on Spanish reviews. *Natural Language Engineering*, 21(1), 139-163.
[57] García, D., & Thelwall, M. (2013) Political alignment and emotional expression in Spanish Tweets. In Díaz-Esteban, A., Alegría I., & Villena-Román J. (Ed.), Proceedings of the TASS workshop at SEPLN 2013 Actas del XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural. IV Congreso Español de Informática (pp. 112–125). Madrid: SEPLN.
[58] Momtazi, S. (2012). Fine-grained German sentiment analysis on social media. In Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'12) (pp. 1215-1220). Instabul, Turkey.
[59] Saralegi X. & San Vicente I. (2013). Elhuyar at TASS 2013. In Díaz-Esteban, A., Alegría I., & Villena-Román J. (Ed.), Proceedings of the TASS workshop at SEPLN 2013 Actas del XXIX Congreso de la Sociedad Española de Procesamiento de Lenguaje Natural. IV Congreso Español de Informática. (pp. 143–150). Madrid: SEPLN.
[60] Martínez-Cámara, E., Martín-Valdivia, M. T., Molina-González, M. D., & Perea-Ortega, J. M. (2014). Integrating Spanish lexical resources by meta-classifiers for polarity classification. Journal of Information Science, 0165551514535710.
[61] Bradley, M. M, & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.
[62] Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior research methods*, *39*(3), 600-605.
[63] Trnavac, R., & Taboada, M. (2012). The contribution of nonveridical rhetorical relations to evaluation in discourse. *Language Sciences*, *34*(3), 301-318.
[64] Tang, D., Wei, F., Qin, B., Liu, T., & Zhou, M. (2014). Coooolll: A deep learning system for Twitter sentiment classification. *The 8th International Workshop on Semantic Evaluation (SemEval 2014). Proceedings of the Workshop (pp 208-2012).* Dublin, Ireland.

JIS Accepted for Publication
By the Journal of Information Science: http://jis.sagepub.co.uk

[65] Silva, N. F., Hruschka, E. R., & Hruschka Jr, E. R. (2014). Biocom Usp: Tweet Sentiment Analysis with Adaptive Boosting Ensemble. *The 8th International Workshop on Semantic Evaluation (SemEval 2014). Proceedings of the Workshop (pp 123-128)*. Dublin, Ireland.
[66] Taulé, M., M. A. Martí & M. Recasens. (2008). AnCora: Multilevel Annotated Corpora for Catalan and Spanish. In Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08), Marrakech, Morocco: ELRA.
[67] Vilares, D., Alonso, M. Á., & Gómez-Rodríguez, C. (2013). Supervised polarity classification of Spanish tweets based on linguistic knowledge. In *Proceedings of the 2013 ACM symposium on Document engineering* (pp. 169-172). ACM.
[68] Cavnar, W. B., & Trenkle, J. M. (1994). N-gram-based text categorization. Ann Arbor MI, 48113(2), 161-175. www.eldiario.es. Retrieved from http://www.eldiario.es/politica/Tania-Sanchez-IU-Madrid-primarias_0_334216900.html
[69] Pew Research Center "Litle Change in Public's Response to 'Capitalism', 'Socialism'. (2011). Washington, D.C, USA: Pew Research Center.
[70] Oswald, H., & Schmid, C. (1998). Political participation of young people in East Germany. *German Politics*, *7*(3), 147-164.
[71] Riveiro, A. (2014). Tania Sánchez reclama una nueva dirección de IU de Madrid que refleje su triunfo en las primarias.
[72] Silió, E (2014). La guerra interna fractura IU. www.elpais.com. Retrieved from http://ccaa.elpais.com/ccaa/2015/01/12/madrid/1421067353_915483.html
[73] Gil, I. (2014). Auge y caída de Tania Sánchez: de ganar las primarias de IU a peligrar su candidatura. www.elconfidencial.com. Retrieved from http://www.elconfidencial.com/espana/2015-01-10/auge-y-caida-de-tania-sanchez-de-ganar-las-primarias-de-iu-a-peligrar-como-candidata_619370/#
[74] Granado, O. (2014). Iñigo Errejón, suspendido de empleo y sueldo por la Universidad de Málaga, www.eldiario.es. Retrieved from http://www.eldiario.es/andalucia/Errejon-suspendido-empleo-Universidad-expediente_0_331417056.html
[75] RTVE (2014). La Universidad de Málaga suspende cauterlarmente de empleo y sueldo a Iñigo Errejón. RTVE.es. Retrieved from http://www.rtve.es/noticias/20141204/universidad-malaga-suspende-cautelarmente-empleo-sueldo-inigo-errejon/1061202.shtml
[76] Zubiaga, A., San Vicente, I., Gamallo, P., Pichel, J. R., Alegria, I., Aranberri, N., & Fresno, V. (2014). Overview of tweetlid: Tweet language identification at sepln 2014. *TweetLID@ SEPLN*.
[77] Trung, D. N., & Jung, J. J. (2014). Sentiment analysis based on fuzzy propagation in online social networks: a case study on TweetScope. *Computer Science and Information Systems*, *11*(1), 215-228.

Notes

1. http://sentistrength.wlv.ac.uk/SpanishTweetsTestAndDevelopmentSetsDavidVilares.zip
2. https://www.inbenta.com/
3. http://sentistrength.wlv.ac.uk/SpanishSentiDataDavidVilares.zip

Appendices

Appendix A. Description of the parties and their politicians.

Partido Popular (@PPopular) 189,000 followers

Mariano Rajoy (@marianorajoy): The Prime Minister with 650,000 followers.
Soraya Sáenz de Santamaría (@Sorayapp): The Deputy Prime Minister with 154,000 followers.
María Dolores de Cospedal (@mdcpospedal): The PP Secretary-General with 88,900 followers.
Esperanza Aguirre (@EsperanzAguirre): President of the PP Madrid federation with 245,000 followers.
Cristina Cifuentes (@ccifuentes): Delegate of the Spanish government in Madrid with 65,500 followers.

Partido Socialista Obrero Español (@PSOE) 195,000 followers

Pedro Sánchez (@sanchezcastejon): The current leader and Secretary-General of PSOE with 112,000 followers.
César Luena (@cesarluena): The PSOE Secretary and deputy leader with 9,848 followers.
Susana Díaz (@_susanadiaz): President of Andalucia with 44,400 followers.
Carme Chacón (@carmechacon): Former Minister of Defence with more than 87,900 followers.
Elena Valenciano (@ElenaValenciano): Head of the PSOE 2014 European election list with 21,300 followers.

Podemos (@ahorapodemos) 482,000 followers

Pablo Igleias (@_Pablo_Iglesias): Leader and current Secretary-General of the party with 739,000 followers.
Juan Carlos Monedero (@MonederoJC): Program Secretary of Podemos with 128,000 followers.
Iñigo Errejón (@ierrejon): Secretary of Politics with 145,000 followers.
Pablo Echenique (@pnique): Representative of Podemos in the European parliament with 95,800 followers.
Teresa Rodríguez (@TeresaRodr_): A European parliamentary member with 64,700 followers.

Izquierda Unida (@unida) 124,000 followers

Cayo Lara (@cayo_lara): Current coordinator of Izquierda Unida and member of the Spanish parliament with 170,000 followers.
Alberto Garzón (@agarzon): Member of the Spanish parliament and future IU leader with 282,000 followers.
Tania Sánchez (@Ainhat): IU candidate for Mayor of Madrid with 84,700 followers.
Gaspar Llamazares (@GLlamazares): Former head of the party with 227,000 followers.
Hugo Martínez Abarca (@hugomabarca): Member of IU with a high Twitter profile and 26,200 followers.

Unión, Progreso y Democracia (@UpyD) 106,000 followers

Rosa Díez: Party leader without a Twitter account during the polling period.
Toni Cantó (@ToniCanto1): Spanish actor and member of the Spanish parliament with 169,000 followers.
Irene Lozano (@lozanoirene): Deputy of the Spanish parliament with 16,700 followers.
Carlos Martínez Gorriarán (@cmgorriaran): Member of the Spanish parliament with 21,300 followers.
Beatriz Becerra: European Deputy with 6,700 followers.

Ciudadanos (@CiudadanosCs): 73,800 followers

Albert Rivera (@Albert_Rivera): Founder and president with 141,000 followers.
Luis Salvador (@luissalvador): Candidate for Mayor of Granada and a high Twitter profile with 63,000 followers.
Fran Hervias (@FranHervias): Secretary with 5,000 followers.
Inés Arrimadas (@InesArrimadas): Catalan member of parliament with 7,000 followers.
Javier Nart (@JavierNart): European Deputy with 18,400 followers.