Twitter, language and politics: What language do people use to make arguments about DACA?

Mireia Triguero Roura
Columbia University
mt3197@columbia.edu

Abstract

What are the implications of language when people make arguments towards immigrants? This research looks at the cognitive frameworks people use to understand membership to the nation and justify anti-immigration sentiment. The debate around DACA provides an opportunity to look at what schemas people deploy to make sense of and argue their position on the issue in an open-ended medium. Using Twitter data and computational methods of text analysis I look at what arguments people make and categorize the discursive frameworks that people deploy in their argumentation. Analyzing these frameworks allows us to understand the underlying notions of citizenship and membership to the nation.

1 Introduction

Academics in many different fields have studied polysemy in language and its social and political implications. Similar words, and even the same exact word, can have very different meaning for different people and in different social contexts (Hall, 2001), the word welfare will mean and be associated with different words if you are a Republican or a Democrat. Similarly, sociologist have long looked at what Haney-Lopez dubbed dog-whistling or the ability to signal some racist message with seemingly innocuous, non-racial, language (López, 2015; Bonilla-Silva, 2017) . This project aims at looking precisely how language is deployed to defend opposing political positions on the Deferred Action for Childhood Arrivals debate. What language do people use to make arguments about the contentious immigration policy? My project looks at the language people deploy on Twitter, how that language is used and what discursive frameworks people access to.

The Deferred Action for Childhood Arrivals was a policy implemented as executive order by former president Obama in 2012. It's stated goal was to give temporary relief to children (under 16) that had been brought into the country without proper documentation before 2007. These children are also known as Dreamers, because they were also the target of a similar legislative proposal, the Dream Act. DACA offered access to working permits as well as a temporary freeze on deportation. DACA permits last two years and are renewable. In October 2017, Donald Trump, who ran on an anti-immigration platform but has also stated elsewhere that he supports the DACA recipients, said that he would end DACA on March 5th. This move, it was argued, was supposed to be a deadline for Congress to come up with a more permanent solution that could be voted on and passed by the legislative branch. March 5th came with no resolution, but state judges have declared that it was illegal to end the policy to begin with. Amid of all this turmoil, concerned citizens have been arguing about their positions (for or against) on Twitter and elsewhere. This paper uses precisely the debate on Twitter as data.

2 Previous Research

Although a lot has been done in terms of sentiment analysis classification, framing and ideology classification is a newer field. One of the particular challenges of the literature is that it is hard to detect for and against positions when they are not inscribed in a particular discussion (i.e. people in twitter arent responding to someone else specifically for the most part) (Tan et al., 2016) or have a clear signal word attached to them. For instance, Celli et al. (2016) were able to accurately predict the Brexit vote by looking at agreement on Twitter. However, in that specific case, the debate around Brexit had clear single words (i.e. remain for those against and leave for those against).

Moreover, most state-of-the-art techniques need some supervision and hand-annotation of tweets. Considering the economic and time constraints of this class, outsourced hand-annotation of my data is beyond the scope of this project. For a future iteration of this project, I will use Amazon Turk to hand-annotate tweets as for or against and then use that to train an algorithm to properly classify tweets (Greene and Resnik, 2009; Iyyer et al., 2014). This will both give me a more accurate classification of the tweeters stance on the debate and also a gold standard to evaluate the effectiveness of the method.

For the scope of this project, however, I will use a more traditional sentiment analysis classification approach. Some of the most common approaches to sentiment analysis are dictionary-based, which means they rely on sentiment lexica to propagate polarity (Fellbaum, 1998; Baccianella et al., 2010). One of the biggest limitations with these approaches is that these dictionaries tend to be very general and not domain-specific. However, we know semantic meaning varies a lot in different situations, and therefore they risk misclassification. Since not only I am looking at domain- but also topic- and subtopic-specific documents, general lexica might not necessarily be fine-grain enough for the task. Other approaches that try to go beyond this limitation and look at corpus-specific sentiment require some form of supervision (e.g. labeled data), which, as I wrote above, is something beyond the scope of this project. Specifically, for Twitter, other methods have used emoticons as signals (Mohammad and Turney, 2010). Because of the specificity of the topic, limiting it to emoticon use might miss a lot of information.

Hamilton et al. (2016) propose a corpus-based approach that does not require distant-supervision. SentProp constructs a lexical graph from unlabeled corpora (using a small seed set) and then propagates it (using random walk method). This approach is particularly appropriate when dealing with relatively small samples of unlabeled data. Another interesting feature of the method is that, using bootstrapping, it provides a measure of uncertainty, making it more robust than other approaches. For these reasons, SentProp seems a more adequate method for my task.

To extract what frames each side uses, I turn to topic models. Topic models is broadly defined as a method to find the underlying topics in a corpus of documents. Its most common implementation is the Latent Dirichlet Allocation (Blei et al., 2003), which, given a set number of topics, computes topic probabilities for each document. Although initially used as an exploratory method, social scientists have been using topic models combined with some covariates to measure what is understood as a latent linguistic, political and/or psychological variable. Thus, the combination of the sentiment polarity score with the topic models is a potentially adequate method to get a sense of the underlying frames that people use to talk about DACA. Structural Topic Models was created precisely as a solution to this problem, that is combining topic models with some covariate (Roberts et al., 2013). STM has been used in political science in open-ended survey experiment questions. Gadarian et al. (2014) found that after being exposed to concerns about immigration, respondents would talk about immigration in terms of security and welfare concerns, whereas those in the control group would emphasize legality and citizenship challenges. In short, when paired with the treatment covariate, STM could more accurately predict the topics for each condition.

3 Analysis

3.1 Data

My data are tweets of people who mention the word DACA or Dreamer (or its hashtag formulations). Twitters API free access allows users to scrape data in two different ways. The first option is to do a limited historical search (up to one week before the day); the second is a live streaming filtered by keyword of the twitter traffic. I used this second option and I streamed tweets during the days before the DACA expiration deadline (March 5th) and after. Twitter documentation states that only a percentage of the actual data traffic can be streamed using this method, however they do not specify what this percentage is. By some accounts, it is relative to the volume of tweets at any given moment, meaning that it can be anything from the total amount (for very restricted stream queries) to a small percentage such as 1% (if your stream query is very broad). Although I do not expect to have collected all the data that was being tweeted at the time of my streaming, I am confident that because of the randomization of the API, it is not a systematically biased sample.

Another question that comes from these data is how representative the Twitter population is of the entire United States population. It has been studied and we know that not all ages are represented in Twitter (with the biases privileging younger and middle-age population), that there is a strong male bias, even if it seems to have been reducing over the last few years. Similarly, when it comes to gender, African-American seem to be oversampled and Hispanic tend to be underpresented (Mislove et al., 2011). None of these are major concerns to the goal of this project, since I am looking specifically at the Twitter population, rather than the American population in general. Another concern in using Twitter text as data is that you miss other information that twitter users deploy on the online forum (Tufekci, 2014). For instance, the use of images or video to complement one's argument, as well as tweets where the main message is contained in the image (as a way to bypass the 140 characters or to bring conversations form other media, i.e. text messages, etc.).

From the raw streaming, I compiled about 3 millions tweets. However, this included a great deal of repeated tweets. This can be due to several reasons. For instance, repeated tweets can simply be retweets (either as such or as a new tweet following the old-school Twitter etiquette of ”RT ”). It can also be that short messages might be written exactly in the same way by different people. Finally, it can also be that they come from bots that mass-tweet the same message in order to influence the twittersphere. Although there are approaches to detecting what accounts are bots or not, for this project I assume that bots are also participating on the Twitter discussion and as such I do not see a reason to eliminate them (Davis et al., 2016). However, to avoid giving more weight to those tweets that have been captured on my stream in disproportionate ways, I eliminate all the repeated tweets. After this, I end up with approximately 290,000 unique tweets in English.

3.2 Categorizing polarity

In order to give a measure of ”polarity” (defined as for or against) to my data to be used as a covariate for the the STM, I use SocialSent (Hamilton et al.,

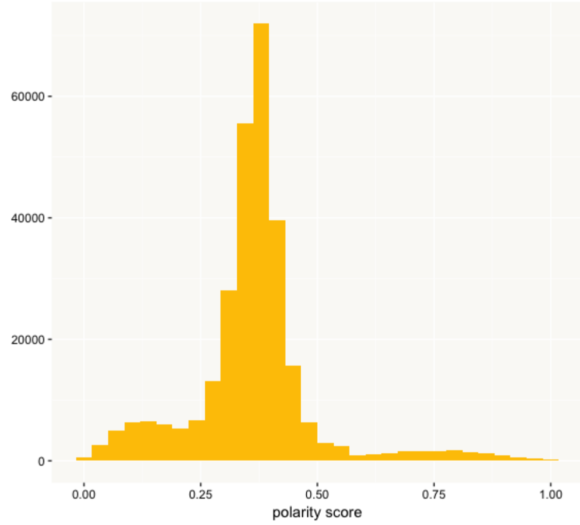

Figure 1: Distribution of polarity scores

2016). Even though SocialSent includes twitter-specific seed words, the assumption that those in favor of DACA will use more positive words and those against will use negative (or vice versa) fails at first look of the data. In fact, people on both sides use overall more negative words, since people on both sides were very angry at the situation (those for DACA because it was ending with no legal solution in place putting the lives of many at risk; those against because they never liked the policy to begin with and have strong antipathy for the recipients). Thus, I developed my own set of seed words of for/against sentiment. Using Nelson's (2017) Computational Grounded Theory approach, I both looked at my data quantitatively (looking at word distributions) and qualitatively as well as informing my decisions using theory. So, for instance, it is well known that the words ”illegal” and ”undocumented” have implied political positions in them (Merolla et al., 2013); ”illegal” is used more by detractors and ”undocumented” by those more in favor (or at least neutral). Similarly, my qualitative assessment yield that some hashtags were clearly attached to one side or the other (e.g. #NoDaca was for those against and #ProtectDreamers, for those in favor).

The SentProp algorithm is very sensitive to how it is initialized. The authors recommend using their default SVD-based embeddings on a restricted vocabulary (¡50000 words), as is the case in mine. For comparison purposes I also tried word2vec, and their SVD embeddings seem to give more stable results, so I stayed with their rec-

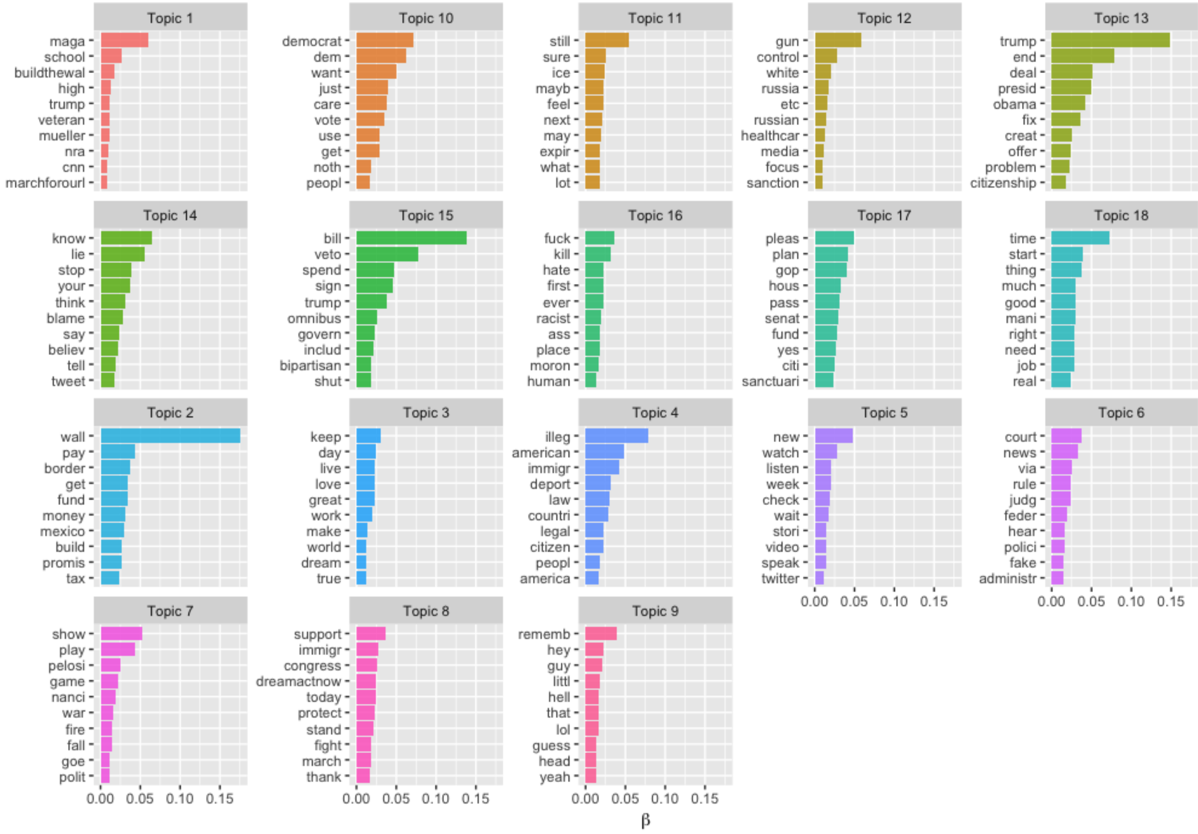

Figure 2: Highest word probability for each topic

ommendation. After obtaining the polarity of each word in my dictionary, I propagate the polarity from word to tweet by averaging the polarity of every word in a tweet multiplied by the inversed standard deviation of the word polarity. This is to ensure that words with a more certain polarity score get more weight in defining the polarity of a tweet. Thus, for instance, if a tweet was using one of the polarity hashtags, they would get a high polarity score in either direction. In figure 1 we can see the distribution of polarity. This distribution is biased towards the ”against” position. This could be because the dataset is truly biased or because the algorithm is not working as expected. A qualitative assessment of my data shows that for extreme polarities (¿0.6 and ¡0.4) the tweets are classified at about 85% accuracy. Moreover, most tweets seem to have middle-of-the-road opinion. That is probably due to the averaging of polarities by tweet, longer tweets get pulled towards the center.

3.3 Extracting frames

Topic models assume that there is a latent topic (o topics) that people draw from when constructing their texts. For the purposes of this research, I understand this latent topics as discursive frames. Using Structural Topic Models (Roberts et al., 2013) and the polarity score as a topical prevalence covariate we can further analyze the data and get the main discursive frames. After removing common words (including the seed words and Daca and Dreamers from SentProp to avoid confounders), I optimized for heldout likelihood and semantic coherence and settled for 18 topics. Figure 2 shows the main words associated with each topic. Topic 8 seems to present the traditional vocabulary associated with those arguing against DACA using the legalistic discourse. They use the vocabulary of the law, drawing distinctions between ”legal” and ”illegal” persons, citizens versus ”others” to make arguments as to why these immigrants are not and should not be American citizens. Looking at figure 4, we can see that, indeed, topic 4 is highly associated with negative opinions. At the same time, this analytical approach allow us to separate this legalistic framework from topic 1, which is a more racialized framework, bring up Trump’s ”Make America Great Again” approach and the infamous wall

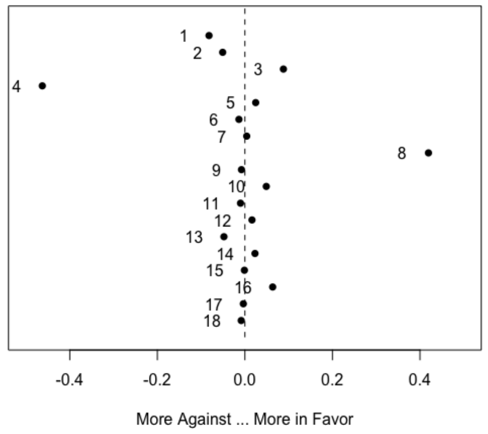

Figure 3

to the south. This topic is also associated with negative opinion but it can be distinguished from topic 4. Looking at the positive polarity, topic 8 ranks the highest. Looking at the words in figure 2 for topic 8, it seems that it is mostly people tweeting about the march against the March 5th deadline on Congress.

4 Conclusion and future work

Hall argued that people make use of different sets of frameworks, that work as the scaffoldings of interpretation and understanding, and that these are not directly accessible or conscious by the speakers, but rather operate at a deeper level and are different depending on what our position is. This paper has shown that automated text analysis methods allow us to systematically extract meaning and build the discourse frameworks that people rely on to make sense of the DACA debate. Even when people use very similar words to talk about the same thing, they are coming from different frameworks which imbues them with very different meanings and implications. A good example is the word ”citizen.” Although the word itself can be used by both proponents and detractors, when people use it consistently with topic 3 (the ”legality discourse” topic), they are most likely defending a position against. It has been studied how legality discourse is used to hide prejudiced ideologies. In the case of the DACA debate, appealing to the ”illegality” of the Dreamers is a way to mask racial animus towards Hispanics (Brown, 2013).

Future work on this topic should include labeling the data in order to achieve higher accuracy and a better evaluation of the results. With labeled data, the classification task could make use of more sophisticated approaches such as Iyyer et al.'s 2014 recursive neural networks approach for classifying political ideology. Further work should also investigate the role of retweets and social networks in the dissemination of opinions and arguments. A future iteration of this project could rebuilt the Twitter networks of followers and following of the tweets' authors in order to see how different political positions map out in the social network. This could also open it to better understanding of political polarization on the Internet.

References

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*. volume 10, pages 2200–2204.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3(Jan):993–1022.

Eduardo Bonilla-Silva. 2017. *Racism without racists: Color-blind racism and the persistence of racial inequality in America*. Rowman & Littlefield.

Hana E Brown. 2013. Race, legality, and the social policy consequences of anti-immigration mobilization. *American Sociological Review* 78(2):290–314.

Fabio Celli, Evgeny Stepanov, Massimo Poesio, and Giuseppe Riccardi. 2016. Predicting brexit: Classifying agreement is better than sentiment and pollsters. In *Proceedings of the Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (PEOPLES)*. pages 110–118.

Clayton Allen Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. Botornot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, pages 273–274.

Christiane Fellbaum. 1998. *WordNet*. Wiley Online Library.

Shana Kushner Gadarian and Bethany Albertson. 2014. Anxiety, immigration, and the search for information. *Political Psychology* 35(2):133–164.

Stephan Greene and Philip Resnik. 2009. More than words: Syntactic packaging and implicit sentiment. In *Proceedings of human language technologies:*

The 2009 annual conference of the north american chapter of the association for computational linguistics. Association for Computational Linguistics, pages 503–511.

Stuart Hall. 2001. Encoding/decoding. *Media and cultural studies: Keyworks* 2.

William L Hamilton, Kevin Clark, Jure Leskovec, and Dan Jurafsky. 2016. Inducing domain-specific sentiment lexicons from unlabeled corpora. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*. NIH Public Access, volume 2016, page 595.

Mohit Iyyer, Peter Enns, Jordan Boyd-Graber, and Philip Resnik. 2014. Political ideology detection using recursive neural networks. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. volume 1, pages 1113–1122.

Ian Haney López. 2015. *Dog whistle politics: How coded racial appeals have reinvented racism and wrecked the middle class*. Oxford University Press.

Jennifer Merolla, S Karthick Ramakrishnan, and Chris Haynes. 2013. illegal,undocumented, or unauthorized: Equivalency frames, issue frames, and public opinion on immigration. *Perspectives on Politics* 11(3):789–807.

Alan Mislove, Sune Lehmann, Yong-Yeol Ahn, Jukka-Pekka Onnela, and J Niels Rosenquist. 2011. Understanding the demographics of twitter users. *ICWSM* 11(5th):25.

Saif M Mohammad and Peter D Turney. 2010. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*. Association for Computational Linguistics, pages 26–34.

Laura K Nelson. 2017. Computational grounded theory: A methodological framework. *Sociological Methods & Research* page 0049124117729703.

Margaret E Roberts, Brandon M Stewart, Dustin Tingley, Edoardo M Airoldi, et al. 2013. The structural topic model and applied social science. In *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*.

Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. In *Proceedings of the 25th international conference on world wide web*. International World Wide Web Conferences Steering Committee, pages 613–624.

Zeynep Tufekci. 2014. Big questions for social media big data: Representativeness, validity and other methodological pitfalls. *ICWSM* 14:505–514.