Politifact Language Audit

Dallas Card,* Lucy H. Lin,† and Noah A. Smith†
*School of Computer Science, Carnegie Mellon University
†Paul G. Allen School of Computer Science & Engineering, University of Washington

March 23, 2018

1 Introduction

We report on attempts to use currently available automated text analysis tools to identify possible biased treatment by Politifact of Democratic vs. Republican speakers, through language.

We begin by noting that there is no established method for detecting such differences, and indeed that "bias" is complicated and difficult to operationalize into a measurable quantity. This report includes several analyses that are representative of the tools available from natural language processing at this writing. In each case, we offer (i) what we would expect to see in the results if the method picked up on differential treatment between Democrats vs. Republicans, (ii) what we actually observe, and (iii) potential problems with the analysis; in some cases we also suggest (iv) future analyses that might be more revelatory.

2 Initial Corpus Exploration

When thinking about how language use may differ between two groups (in this case, the reporting on two different political parties), we are not interested merely in differences in *content*, as content will vary due to inherent differences between the groups. Rather, the goal is to isolate differences in tone or stance or style, after controlling for differences in content.

While it may not be possible to completely distinguish style from content, as any linguistic choice has the potential to make some contribution to the overall framing of the subject, the differences that are of interest here are those that might be perceived as representing a bias on behalf of the source (Politifact authors, editors, and supporters). Thus, it is important to understand how the overall coverage of the two parties differs, even before looking at the language. As such, we begin with a brief overview of the basic properties of the Politifact corpus, in terms of party, ruling, and subject, without yet looking at the language. Most figures for this section can be found in the Appendix.

The earliest articles in this corpus date from 2007, but the bulk of them are from 2010 to 2017 (Figure 20), thus mostly overlapping with the Obama presidency. If we divide them by party, we see that the majority of articles are about Republicans, with about 1.4 times as many about Republicans as Democrats (Figure 21).[1] There are approximately 10,000 articles about people from one of these two parties, and that is the subset we will focus on in this analysis.

If we look at the set of individual speakers covered by this corpus, we find that Barack Obama is the speaker who has been reported on most frequently, being the subject of approximately 600 articles. Hillary Clinton is the next most popular Democrat, with about half as many articles as Obama. The coverage of Republicans, by contrast, is more evenly distributed across several popular subjects, including Donald Trump, Mitt Romney, and John McCain (see Figures 22 and 23). Thus, we expect that we will need to account for speaker names, particularly for these high-profile individuals.

[1] This might be explainable by regional partnerships, which we understand are not necessarily balanced across Democratic- vs. Republican-controlled jurisdictions.

The rulings used by Politifact fall on a 6-point scale, ranging from "True" to "Pants on Fire!" (baldly false). There is a fairly even distribution of articles across the first five categories, with a smaller number (approximately 700) being rated as "Pants on Fire!" (see Figure 24).

If we divide the articles by party, and plot the total number of articles per party that have been given each ruling, re-centering to reflect the imbalance between the number of articles about each party (see Figure 1), we find that Democrats are over-represented in all three "true" categories, whereas Republicans are over-represented in all three "false" categories. Thus, our analysis will need to account for words related to the truthfulness of claims.[2]

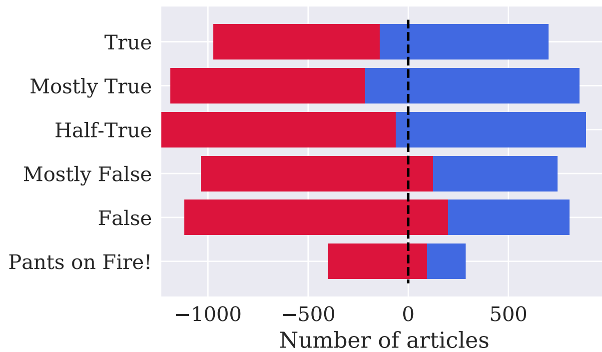


Figure 1: Number of articles per ruling, colored by party, and centered to reflect the overall balance of articles between parties.

As a further illustration of this, Figures 25 and 26 show the distribution of ratings for Barack Obama and Donald Trump, the two most covered individuals in the corpus. The most commonly used ruling about claims made by Obama is "Mostly true", where as for Trump, the dominant ruling, unsurprisingly, is "False".

Finally, we also look at the "subject" tags that have been provided for each article by Politifact. The articles in this corpus cover a large number of subjects, including the *economy*, *health care*, *education*, *elections*, *immigration*, and so on (see Figure 27).

Figure 2 show the number of articles per subject, divided by party in the same way as the rulings in Figure 1. This figure reveals that there is surprisingly even coverage between the two parties for most subjects. There are still a few subjects where the different priorities of the two parties is reflected in the number of articles per party on those subjects, such as *terrorism*, *immigration*, and *education*, but in most cases, the within-subject split is similar to the overall division between parties. As such, we don't expect that any differences in language we find between the two parties will be strongly based in the content of the articles (at least in terms of these subjects).

In summary, there are obvious differences between the parties in terms of the overall coverage by Politifact, including the total number of articles (with more about Republicans), the correlation between party and ruling (with Republicans claims more likely to be ruled as False), and the distribution of individuals covered (with coverage of Democrats disproportionately focused on Barack Obama and Hilary Clinton). However, the coverage per party is relatively balanced for most subjects (with immigration being a relatively extreme exception). As a result, we will need to deal with differences in language that arise from these underlying differences in coverage, including party affiliation, terms associated with truthfulness, and individual names. In the following investigation, we will seek to differentiate the language used in articles about statements made by members of the two parties.

3 Internal Analysis

In this section, we present analyses based entirely on Politifact data (text and rulings). Later analyses in Section 4 will also use text drawn from other sources and other linguistic resources.

[2]Note that our analysis cannot determine whether there are partisan biases in Politifact's judgments about truthfulness nor about Politifact's selection of which statements to examine.

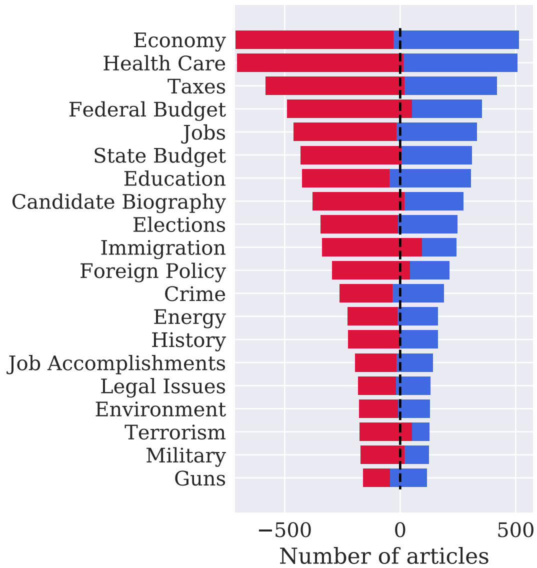


Figure 2: Number of articles per subject, split by party.

3.1 Predicting Party from an Article

In order to analyze differences in language between the coverage of Republicans and Democrats, one approach is to build predictive models that aim to guess the party affiliation of an article's subject, given the text. This is essentially a **text classification** problem, where the party affiliation is a label and the features are derived entirely from the text. To the extent that a classifier can accurately discern whether an article is about a Democrat vs. a Republican, it might be a sign of differential treatment, but only if the textual cues it relies on are stylistic rather than contentful.

To begin, we remove all quoted text from each article, leaving only what has been written by Politifact.[3] We also do basic preprocessing: splitting text into tokens, replacing urls, email addresses and Twitter @mentions with special chararacters, and dropping most punctuation. This gives us a vocabulary of approximately 66,000 types, with a total of approximately 7.1 million tokens. Because bigrams are often more evocative than unigrams, giving more context for interpretation, we also extract bigram phrases from the text, giving us an additional 1.5 million types.

As a strong, interpretable baseline, we use l_1-regularized logistic regression, which is capable of selecting a relatively small number of features (unigrams and bigrams predictive of party affiliation) from a high-dimensional vocabulary. Applying this model to the unigram and bigram features is able to achieve approximately 80% accuracy.[4] Because of the imbalance of articles between Republicans and Democrats, always predicting the most common party would obtain approximately 58% accuracy.

The real-valued weights returned by logistic regression for each feature give an indication of the relative importance of each feature (word or bigram). The conventional way to describe the meaning of a weight value β for word w is that, in a text with all other words held exactly the same, the addition of one instance of w will change the classifier's probability assignment by a factor of e^{β}. However, the most highly-weighted terms are often rare and idiosyncratic (in this case, we find terms such as "economy_data" and "always_passed"). Because of this, we multiply the weights by the average frequency of each type in held-out data, to get a more meaningful measure of the impact of each term on model predictions.[5] Figure 3 shows the highest impact terms for both parties.

[3] There are slightly more quoted lines in articles about Republicans than there are in articles about Democrats, on average 11.0 vs. 9.9, a small but significant difference. The same holds for the total number of characters (letters) in quotes.

[4] Because there is no formal "test set" in this corpus, to estimate model performance, we first randomly split the articles into 64% training data, 16% development data, and 20% test data. We then train a model on the training data, using the development data to estimate performance for tuning the strength of l_1 regularization, and measure final performance on the test data. We repeat this three times using different random splits; we report the average, though all three tests obtained the same accuracy to within 1%.

[5] In order to get the most robust estimate of the importance of each term, we report values for models trained without first dividing the corpus

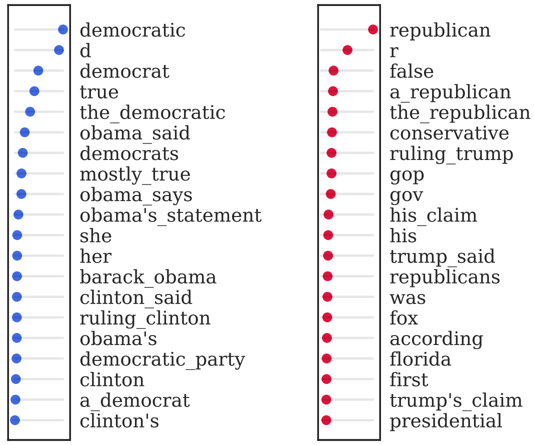


Figure 3: Most impactful terms for predicting party affiliation in a logistic regression classification model

As can be seen in Figure 3, this approach does discover differences between the terms used in articles about each party, but mostly for the content-based reasons we expected from our initial exploration of the corpus. Inspecting the top terms, most of them can be attributed to an explicit mention of the party or a substitute (such as “d”), the name of a speaker that has received a lot of coverage (“Obama” or “Trump”), or a ruling as to the truthfulness of the claim, which naturally matches the correlation noted above in Figure 1, with more Republican claims being ruled False, and vice versa. These terms also capture some terms related to the different positions held by the various parties during this time period, such as “gov”, which occurs as an abbreviation of Governor.

In order to look past these surface features, we repeat the preprocessing, but this time we filter out more terms. We choose to filter the terms used in rulings (“true”, “false”, “mostly”, “half”, “pants”, and “fire”); the names of the parties, and all obvious synonyms, including “rnc” and “dccc”; terms with obvious potential correlation (“liberal”, “conservative”); terms reflecting the positions held differentially by members of the parties (“house”, “senate”, “presdient”, “gov”, “governor”, “speaker”, “rep”, etc.); the names of all states; the first and last names of all speakers in the corpus; and a standard list of stopwords from the Snowball stemmer.[6]

Figure 4 shows the corresponding results for this approach. As we would expect, the performance is degraded without these features, such that the model only obtains 65% accuracy.

Some of the features that remain have a fairly obvious provenance in content, such as “fox_news”, but most are fairly ambiguous, and seem unlikely to be polarizing. There are a few terms in each list which do hint at a subtle difference in coverage, but might relate primarily to the relative ratio of true/false rulings between the two parties. Specifically, some of the highest impact Republican terms have a subtle suggestion of falseness; the word “claim” itself might connote that a proposition is questionable (compare to “statement”), as do “however” and “actually”. On the Democratic side, by contrast, there seems to be a higher likelihood of pointing to direct evidence, as in “source”, “reported”, and “statistics”.

In order to get a better sense of the relevance of any particular terms, it is helpful to look at its appearances in context (O’Connor, 2014). For example, Table 1 shows randomly chosen examples in context of the term “least” (the most influential term for predicting a speaker to be a Democrat). This reveals that it is almost always used in the construction “at least”, but does not provide a great deal of insight into why this would be more strongly associated with Democrats. We provide a browsable interface for further exploration of this model, which may lead to greater

into training and test sets. Instead, we randomly divide the data into five subsets; for each one, we train a model on the remaining four-fifths of articles, using the held-out one-fifth as a development set to choose the strength of l_1 regularization. We also use frequency of terms in this held out one-fifth of articles to compute the impact of terms from the model weights. Finally, we then average the weights and impact values for each of the five models. It is impossible to simultaneously estimate a model from all of the data *and* estimate model performance, hence the combination of approaches used here.

[6] http://snowball.tartarus.org/algorithms/english/stop.txt

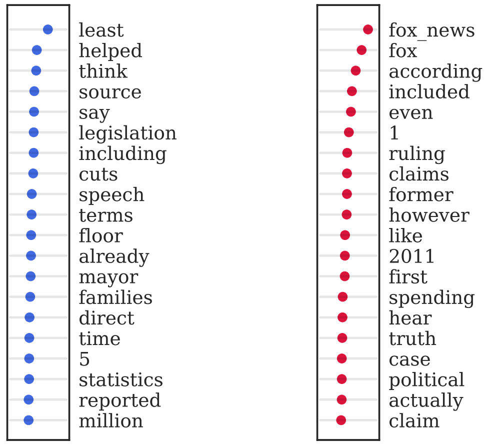


Figure 4: Most impactful terms for predicting party affiliation in a classification model, after filtering out obvious terms.

insight.[7]

Party	Example		
D	took place in those early months. At	least	two media outlets interviewed clients who
D	of the 2004 election led to at	least	18 lawsuits that landed in Brunner's lap.
D	U.S. Labor Secretary Robert Reich gives at	least	partial credit to economic pain orchestrated by
D	data that shows there has been at	least	one period in American history in which
D	finding ways to completely avoid or at	least	reduce the number of affected children by
R	with Rolling Stone that at the very	least	put Obama in a difficult position. Petraeus'
R	Obamacare, for example, companies with at	least	50 full-time employees face fines for not
R	tax or fee could be worth at	least	$3 million annually. In 2006, Indiana
R	they would pay a fine of at	least	$5,000. In the course of our research,
R	age 16, have lived here for at	least	five years, and are in school, are
R	he wrote. Carson reiterated that argument at	least	twice – in an Oct. 8, 2015,

Table 1: Examples of the term "least" in context.

In summary, this part of our analysis finds no obvious differences in the language that is used to describe individuals of each party in a way that shows any indication of bias or differential treatment. If follow-on study were to be conducted, we might recommend annotating samples of documents according to the frequencies of different kinds of evidence used for the ruling, then comparing Democratic vs. Republican articles

3.2 Document Modeling

The above section made an attempt to isolate the effect of party on the language is used to describe each party, but does not account for many correlations that may nevertheless be present, such as names of cities, and terms associated with subjects that are more heavily reported for one side as opposed to another.

In this section, we make use of a new tool for modeling documents with metadata, called Scholar,[8] which is similar to a topic model, but with the ability to introduce additional prior knowledge. We use this to explore themes in the corpus, and how they relate to party.

[7] `http://ark.cs.washington.edu/demo/politifact`

[8] This work is currently under double-blind peer review. We will attach a draft of the paper to the final report and ask that it not be circulated outside the team.

As suggested by the overrepresentation of one party in articles about certain subjects (e.g., immigration), there is an inherent connection between party and subject that cannot be easily disambiguated from how Politifact discusses a party in relation to a subject. Nevertheless, by not restricting ourselves to the predefined subjects, additional salient groupings of terms may appear. The aim of this analysis is to discover such groupings so that they may be inspected by humans.

As a starting point, we use Scholar to jointly learn a set of topics and a classifier which predicts party from those topics. Following the same pre-processing as above, we additionally restrict the vocabulary to the most common 5,000 words (excluding numbers), and arbitrarily chose the number of topics to be 20. Figure 5 shows the top words for these discovered topics, along with the corresponding probability the model assigns to predicting that an article entirely about that topic would be focused on a Democrat.

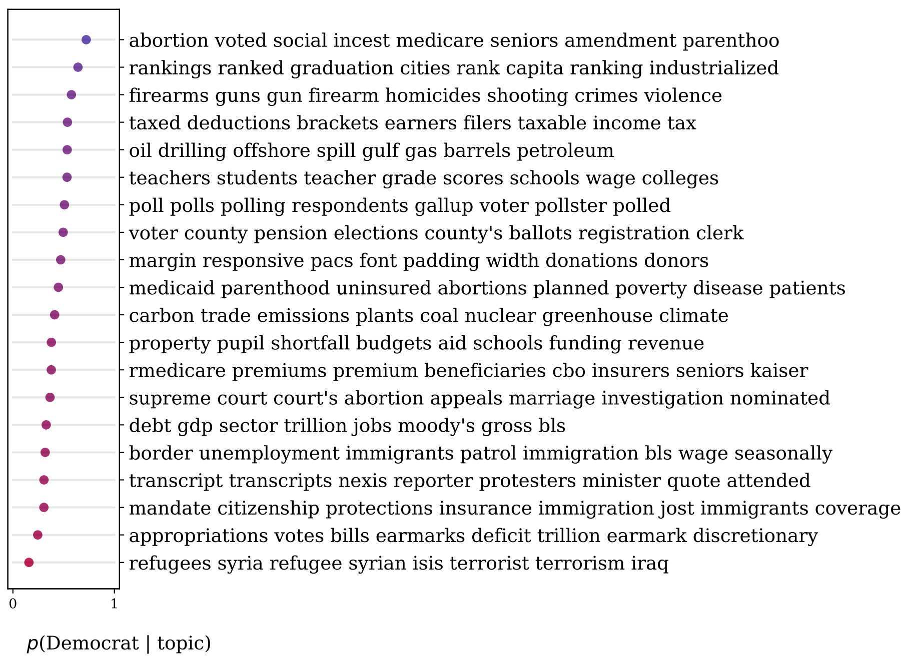


Figure 5: Topics discovered in the Politifact corpus, and the corresponding probability of an article being predicted to be about a Democrat.

This analysis does not provide much insight beyond Figure 2, but reveals that we do not need the explicitly identify subject tags in order to find topics which relate to party; here for example, we find that articles about Democrats focused more on issues like maternal health, education, and guns, whereas articles about Republicans are more likely to be focused on immigration, budgets, and terrorism. These tendencies seem to reflect the underlying priorities of these parties.[9]

An alternative analysis builds in an assumption of explicit deviations in word frequencies for party membership, along with interactions between topics and party. This analysis provides some insight into the different ways in which the parties talk about different issues, differences which seem to be reflected in the writing by Politifact. Note that because of the different way of incorporating party information (as a conditioning variable), we end up with different topics in this iteration. Despite using a much smaller vocabulary than the logistic regression model above, this model obtains approximately the same performance as the model from Figure 4, achieving 65% accuracy on average.

Table 2 shows a subset of topics and the Democratic and Republican variations (full list in Table 5 in the Appendix). Once again, the topics themselves are interpretable, as are the party-specific deviations. This result also demonstrates

[9] A follow-on analysis might test for these tendencies in language produced by members of the parties themselves, such as speeches, press releases, or congressional floor debates.

syria nato troops afghanistan iraq iraqi qaida missiles
D: withdrawal afghanistan commander vietnam troops iraq trans bin
R: iran iran's refugees iranian bomb guantanamo russian syrian
petroleum pipeline drilling plants coal oil electricity gas
D: solar renewable electricity crude oil leases petroleum mine
R: epa carbon emissions greenhouse pipeline gallon eia dioxide
abortion abortions gun parenthood incest guns firearms planned
D: incest abortion firearms dealers rape gun firearm abortions
R: abortion payer arms abortions core treaty gun bargaining
jobs losses sector job creation bls manufacturing employment
D: jobs sector counting outsourcing employment growth creation job
R: stimulus advisers bls jobs seasonally reinvestment gains cbo
crimes convicted police offenses enforcement criminal fbi's homicides
D: fatal weapon firearms shootings homicides guns gun shooting
R: refugees ice refugee deported customs deportation syrian sanctuary
deficit discretionary surplus deficits entitlement entitlements spending budget
D: deficit surpluses security sequester surplus social deficits discretionary
R: debt medicare trillion obligations sequestration outlays wars budgetary
cdc disease medicaid uninsured poverty insured cancer stamps
D: contraception marijuana opioid veterans reproductive uninsured deaths
R: premium patients insured poverty premiums medicaid marketplace parent

Table 2: A subset of topics and their Democratic and Republican variations discovered by including topic-party interaction terms.

how certain subjects intersect. For example, for the topic about crime ("crimes", "convicted", "police"), the Democrat's variation is primarily about gun violence ("fatal", "weapon", "firearms"), whereas the Republican variation is primarily about immigration ("refugees", "ICE", "deported"). Again, we believe this analysis reveals more about the differences between the parties than it does in the ways that Politifact is presenting them, but it does reinforce the notion that for many issues, the two parties focus on different elements, or frame them in different ways, and this difference is reflected in Politifact's coverage.

4 External Analysis

The analysis in the previous section focused on how the language used in Politifact articles differs by party. In this part, we extend the analysis by relating the language present in Politifact documents to the language used by the media, in politics more broadly, and to other characterizations of language considered interesting in other realms of inquiry. We consider these analyses *external* because they are also based on language drawn from outside the Politifact corpus.

4.1 Sentiment Analysis

A commonly used dimension of variation for natural language data is *sentiment*, the extent to which a piece of text expresses a negative or positive attitude towards its subject. Sentiment analysis can be framed as a text classification task, and as such would best be done using a model trained on in-domain labeled data.

Annotating Politifact articles for sentiment is out of the scope of this project, in part due to cost, but also due to the challenge of asking human annotators to judge polarity separate from their own personal reactions to articles and also separate from Politifact's rulings. We also ruled out the use of classifiers trained on out-of-domain datasets (e.g., restaurant or product reviews) because we believe the linguistic cues used in such reviews to overtly signal the author's attitude will be quite different from linguistic cues of a Politifact reporter who aims to present a ruling objectively.

We turn instead to simple sentiment classification based on predefined lists of positive and negative words. We view this way of operationalizing an author's attitude as problematic in general, since the meaning of a word depends heavily on the context of its use. Nonetheless, this method is easily replicable and, if a difference between positive and negative word usage was in evidence, it might warrant further investigation.

We make use of the NRC Word-Emotion Association Lexicon (Mohammad and Turney, 2013), which includes word lists for positive and negative sentiment, as well as a set of eight emotions (which we do not use here).

Figure 6 shows the usage of positive and negative sentiment words in articles about Republicans and Democrats (stripped of quotes, as before). As can be seen, there are slightly more negative *and* positive sentiment words in articles about Republicans on average; however those articles are also slightly longer on average. Comparing the average proportion of negative and positive words per article finds no significant difference between parties using a basic test for difference in means.[10]

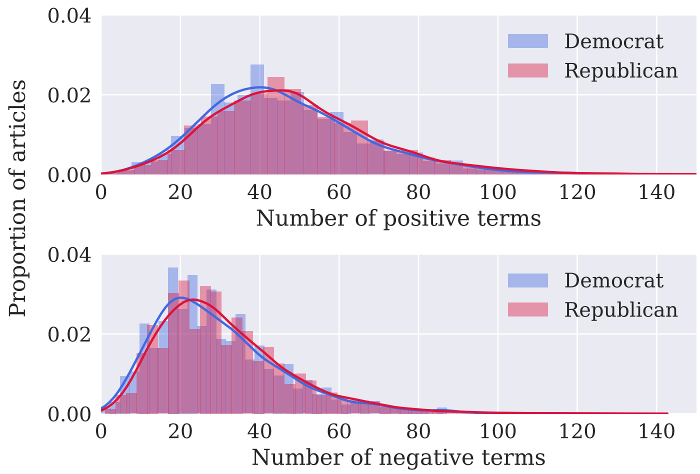


Figure 6: Usage of positive (top) and negative (bottom) words by party, where the x-axis is the term counts per article and the y-axis is the proportion of Republican/Democratic articles with those counts.

From this simple analysis, we conclude that, if there is a difference in sentiment conveyed between Politifact analysis of Democrats vs. Republicans, it is expressed subtly and or in domain-specific ways.

4.2 Hedging

We next examine the use of *hedging*, or expressions of uncertainty. Examples of such language include "in our judgement" and "suggests." If Politifact articles hedge more about one party over another, then it might be taken as a sign of differential treatment.

As a basic indicator of hedging use, we use a list of phrases collected by Tan and Lee (2016).[11] As in the case of sentiment, a simple word list cannot encompass all ways that one might use hedging in language, and use of the phrases in the list might not always be a hedge. Nonetheless, a non-trivial difference between the parties could be of further interest.

Fig. 7 compares the prevalence of hedging phrases in articles about Republicans and Democrats. The top hedge phrases used, regardless of which party is being written about, are "may," "might," "likely," "seem," and "possible." We find that the distribution of how often hedge phrases are used in an article is very similar between Republicans and Democrats. Any differences in hedging are likely to be more subtle. Note, for example, that hedging phrases could be used to hedge against the ruling or against someone else's words (e.g., the subject of the article). At this writing, we do not know of high-accuracy fine-grained hedge analysis methods.

[10]Using "LIWC words" (Linguistic Inquiry and Word Count) returns similarly minute differences, with the most extreme difference being the "female referents" category (used more frequently in articles about Democrats), a difference which is partially, though not completely, explained by the large number of articles about Hillary Clinton.

[11]The list of hedge phrases can be seen at `https://chenhaot.com/data/hedges.txt`.

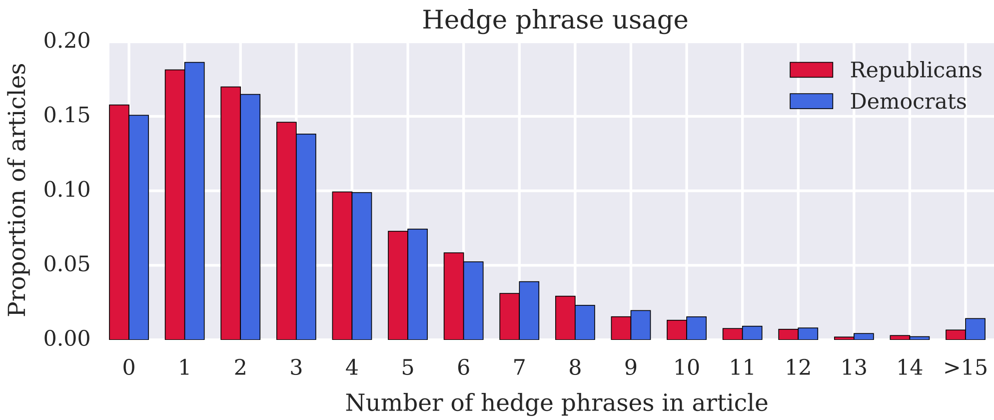


Figure 7: Hedge phrase usage by party, where the x-axis is the count of hedge phrases in an article and the y-axis is the proportion of Republican/Democratic articles with those counts.

4.3 Tone Classification

While the idea of sentiment seeks to capture the overall polarity of an article (e.g., "great" vs. "sad!"), perhaps a more relevant analysis would be how a Politifact article about a particular subject, such as immigration, reads to a person who holds a strong political position on that issue. This is a more difficult linguistic analysis task, but we can approximate it using different external resources.

We are unaware of any existing word lists which would capture the various political positions on a range of issues, so we instead make use of a corpus of news articles that have been annotated in terms of *tone* (Card et al., 2015). For each of several issues in this corpus, tone corresponds, roughly, to whether an article is supportive, neutral, or opposed to a particular political position on the issue, such as pro-immigration vs. anti-immigration, or pro-gun control vs. anti-gun control. In this analysis, we train classifiers to predict tone (pro, neutral, or anti) for each of these issues, and then apply them to the articles in the Politifact corpus. These results must be interpreted with some caution, as the news articles on which these classifiers were trained may differ in important ways from the Politifact articles, most notably in length, but also in style and the goal of the author (largely to report on events rather than to analyze statements for truthfulness).

The two issues for which we have reasonable overlap between the annotated data and the Politifact subjects are immigration and gun control.[12] We are interested in whether Politifact articles (stripped of quotes) appear to our classifiers as similar to text that (implicitly or explicitly) adopts a particular position. This is a proxy for how Politifact articles might appear to those who are have strong opinions about these issues.

Figures 8 and 9 show the total number of articles that are classified into each of the three tones for the subjects of *immigration* and *guns*, respectively. The counts have been colored by party and re-centered to reflect the overall balance of articles between parties within the subset of Politifact articles about each subject. The highest impact terms in these classifiers are shown in Figure 28 and 29 in the Appendix.

First, these two issues are relatively partisan, with most articles about *immigration* being about a Republican speaker, and the opposite for *guns*. Second, for immigration (Figure 8) the Politifact coverage appears relatively balanced to our classifiers, with approximately the same number of total articles being classified as pro-immigration and anti-immigration. Third, there is a slight correlation between tone and party, with a larger proportion of articles that are classified as anti-immigration being about a Republican, when compared to the proportion for pro-immigration.

For gun control, by contrast (Figure 9), the results are both less balanced and more extreme, with far more articles being classified as pro-gun control, and almost none being classified as neutral. Again, there is a correlation between party and tone, with most articles classified as anti-gun control being about Republicans, even though most articles

[12] We use the *guns* subject provided by Politifact, which may have imperfect alignment with "gun control" – the issue around which our external corpus was built.

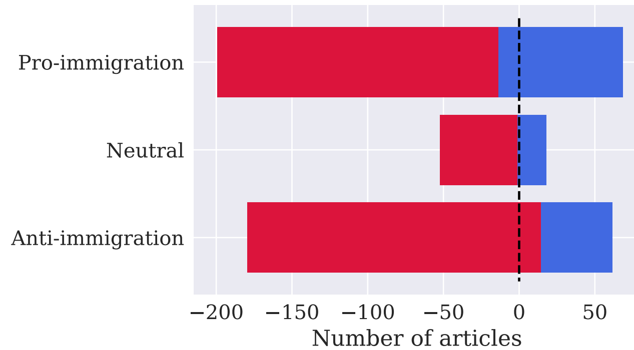


Figure 8: Number of Politifact articles about *immigration* classified according to the tone of coverage, based on classifiers trained on annotated news articles about immigration. Overall, there is a relatively even balance between articles classified as pro-immigration and anti-immigration, with a slight correlation between party and tone.

with the subject *guns* are about Democrats overall.

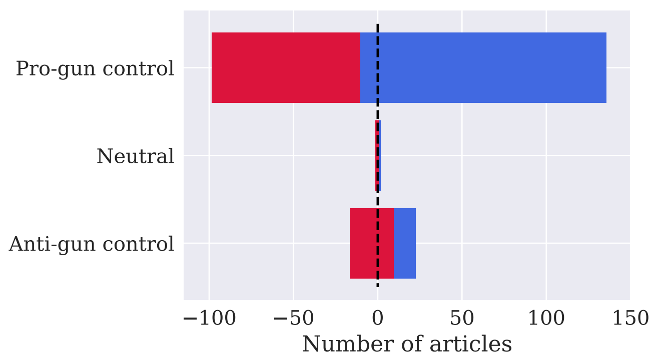


Figure 9: Numbers of Politifact articles about *guns* classified according to the tone of coverage, based on classifiers trained on annotated news articles about gun control. The results here are highly imbalanced, with most articles being classified as pro-gun control, and showing a slight correlation between party and tone.

These findings may partly reflect biases in the annotated news articles, of which 61% were annotated as pro-gun control, and 27% as anti-gun control, but the imbalance in the Politifact data is more extreme. More likely it is the case that news coverage annotated as pro-gun control tends to use language that is in fact relatively neutral (highest impact terms in the model include "gun", "guns", "weapons", "handgun", "gun_violence", "handguns", "children", and "law"), whereas those that take an anti-gun control position focus on a narrow subset of elements, as illustrated by the highest impact terms: "firearms", "law_abiding", "amendment", "anti_gun", as well as "criminal", "carry", and "away" (see Figure 29 in the Appendix for more details).

Given the data that the classifier was trained on, it seems unlikely that a Politifact article will appear to our classifier as anti-gun control, unless it makes heavy use of the lexicon particular to people who hold that position.

As an example, Figure 10 shows the beginning of a Politifact article that is classified as having anti-gun control tone. While it is obvious why some terms carry the weight that they do (e.g., "gun rights"), others are less obvious (e.g., "sought") but some of these terms nevertheless hold a clue to what might stand out to someone who holds a strong position on gun control.

In summary, this analysis tried to analyze the *tone* of a subset of Politifact articles using a corpus of tone-annotated news articles. Only a small subset of Politifact articles overlapping in issues with that corpus could be analyzed. Those discussing immigration were found to be evenly balanced in tone between pro- and anti-immigration, suggesting that

Marco Rubio says Second **Amendment** is unique in **speech** to **NRA**
Sen. Marco Rubio, R-Fla., went all in for **gun_owners** in a **speech** to the **NRA** as the **potential presidential** candidate **sought** to squelch any doubts about his commitment to **gun_rights**.
Rubio was one of a few 2016 Republican **presidential** hopefuls who **spoke** at the **NRA's annual convention** in Indianapolis on April 25. He described the right to own a **gun** as part of the American Dream, highlighted the _quote_removed_ of existing **gun** laws and chastised the **media** for stigmatizing **gun_owners**.
At one point, Rubio described the Constitution of the United **States** as unique. _quote_removed_ he said. _quote_removed_
Wayne LaPierre, the **NRA's executive_vice** president and CEO, made a similar **statement** during his own **speech** to the **annual meeting**, saying, _quote_removed_
Is it **correct** that the United **States** is the only country that gives residents a constitutional right to bear arms?

Figure 10: The start of an article that is classified as having an anti-gun control tone, with **pro-** and **anti-** gun control terms identified by color.

Politifact is making relatively even use of terms that statistically associated with each of these positions. The articles on guns, by contrast, were found to be less balanced, but this seems to be at least partially an artifact of the annotated news articles, in which very common and relatively neutral terms (e.g., "gun" and "gun control") are statistically correlated with a pro-gun control tone; in the Politifact corpus, these terms apparently outweigh the less frequent but more highly-charged anti-gun control terms, such as "gun rights." Nevertheless, it might be worth further exploring through a user study the extent to which Politifact articles about guns signal a pro- or anti-gun control tone.

4.4 Ideological Proportions

In this section, we apply a different kind of probabilistic model to match partisan phrases in Politifact articles. This model, the cue-lag ideological proportions model (CLIP; Sim et al., 2013), is applied to identify proportions of ideologies expressed in the Politifact rulings, in aggregate, while managing uncertainty about those terms' associations with different ideologies.

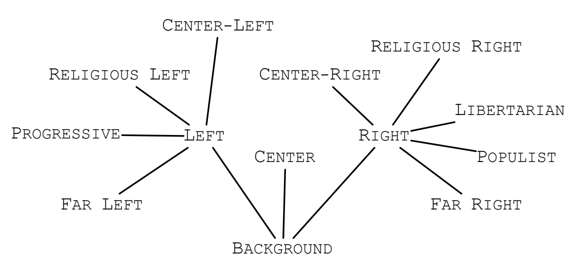


Figure 11: Ideology labels from Sim et al. (2013). The BACKGROUND label is for text without ideological lean.

4.4.1 Model Description

We briefly describe the CLIP model and lexicon, referring the reader to the original paper (Sim et al., 2013) for further details.

CLIP was originally designed and used to measure political candidates' ideological positioning during their speeches, building on the intuition that a candidate will expend more words in a speech to the part of the audience from whom they seek approval. For example, in a primary election, presidential candidates are expected to more heavily use the

Original sentence	Ever since the **Affordable Care Act** was enacted, **Congressional Republicans** have been consistent in their criticism of former **President Barack Obama**'s **health care** law.
Cue-lag representation	$\ldots \xrightarrow{3}$ afford_care_act $\ldots \xrightarrow{2}$ congression_republican $\ldots \xrightarrow{8}$ presid_barack_obama $\xrightarrow{1}$ health_care ...

Table 3: Example of cue-lag representation.

language of their party's "base" members (though not exclusively), and in the general election they will move toward more balanced language that can appeal to members of the other party as well. The model was used to empirically test this hypothesis, which is widely accepted based on theoretical arguments.

Central to this model is the use of a lexicon of automatically discovered terms, called "cues," associated with different ideologies. In this work, "ideology" refers to a more fine-grained category than party affiliation; see Fig. 11 for a visualization of the set of ideologies and their hierarchical relationships. The lexicon consists of approximately 8500 cue terms (bigrams and trigrams). Some examples are "class_struggle" for FAR LEFT, "border_secur" for POPULIST, and "medic_marijuana" for LIBERTARIAN.[13] These cues were extracted from a corpus of 112 political books and 765 magazine issues chosen because the authors of these texts are, by their own self-identification, representative of particular ideologies. To guide the cue extraction process, an expert manually provided ideology labels (and in the case of book chapters, topics like RELIGION and ECONOMY). Here we use the existing CLIP lexicon and ideology tree, with the caveat that using more recent writings from different ideologies might lead to changes in the cue lexicon.

CLIP is only concerned with ideological cues; all other words in the text are treated as filler, interesting only for the number of non-cue words that occur between consecutive cues. The cue-lag representation captures this by representing a text as a sequence of alternating cues (from the lexicon) and integer lags (count of non-cue terms between consecutive cues). An example is shown in Table 3.

CLIP is essentially a hidden Markov model, where the hidden states correspond to ideologies and the output at each time step is a cue term and lag value. The intuition is that a speaker will generate a cue term based on his/her hidden state. A speaker can move between any two hidden states (ideologies) across the course of a text; however, nearby cues are more likely to be from similar ideologies (e.g., LEFT and CENTER-LEFT), whereas longer lags may come from a change across ideologies (e.g., LEFT and RIGHT). Ideological proportions are determined via approximate posterior inference, managing uncertainty about each cue's association with different ideologies.

4.4.2 Findings

We apply the CLIP model to Politifact rulings in the same way that it was previously applied to political speeches. To reflect language use by Politifact specifically, we remove all quotations (speakers and cited experts). If the Politifact ruling text is biased toward the left, then we would expect that the model would find higher ideological proportions across the LEFT ideologies (and vice versa).

Across all rulings. We first look at model output across all Politifact ruling explanations. Figure 12 shows that overall, the CLIP model finds that the rulings use more language that the model associates with RIGHT ideologies. Though it could be possible that this is simply due to the larger number of rulings about Republicans, we will see below that this is unlikely as the RIGHT-leaning trend holds true for both Republicans and Democrats separately.

We can also see that the proportions of RIGHT-leaning ideologies and LEFT-leaning ideologies stay relatively constant across time (though the fine-grained ideologies within the LEFT and RIGHT fluctuate more).

By speaker party. We next consider rulings of statements by the speaker's party, discarding third-party and pundit comments. Figures 13a and 13b show output for Republican and Democratic politicians respectively. What CLIP finds is that there is still slightly more use of RIGHT-leaning language, even when providing judgments of Democratic speakers. That is, it does not find consistent evidence of either (a) the language heavily mirroring that of the speaker, in which case we would expect the corresponding ideology to have high proportion; or (b) the language pushing back against the speaker, in which case we would expect the *opposing* ideology to have higher proportion.

[13]Lists of the top cue terms for each ideology are in Table 1 of the supplementary material to the original paper, found at `http://www.cs.cmu.edu/~nasmith/papers/sim+acree+gross+smith.emnlp13-supp.pdf`.

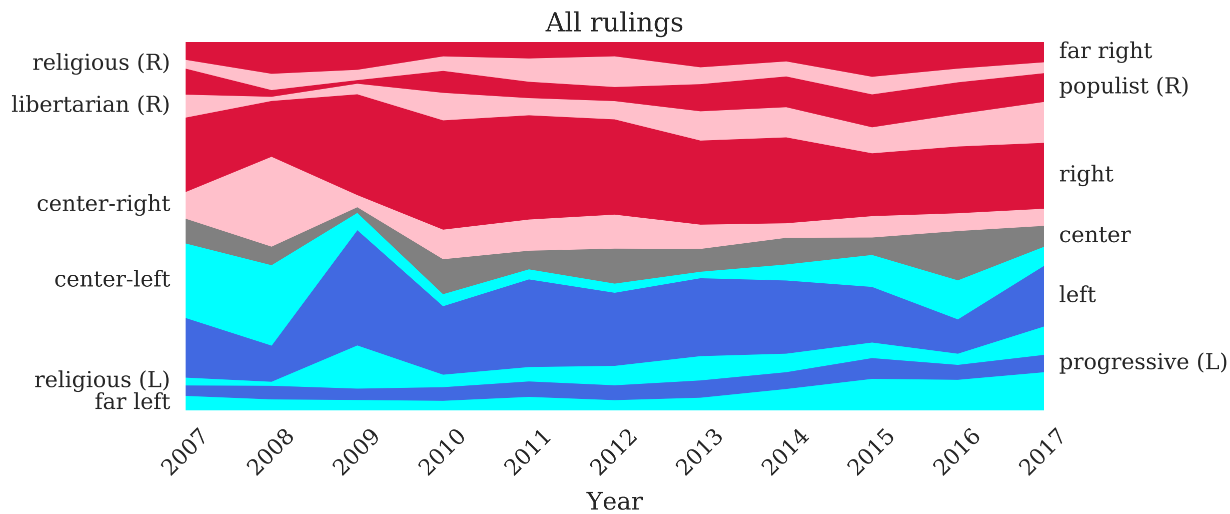


Figure 12: Output of the CLIP model across all Politifact rulings, where the shaded-in areas denote expected proportions for the labeled ideologies across time.

In Table 4, we list the 25 most frequent cues in rulings about Democrats and Republicans. There is a high degree of overlap between the two (particularly concerning mentions of money, taxes, and President Obama). As mentioned previously, cues are not associated with a single political ideology; almost all of the most frequent Democratic cues are listed in both LEFT and RIGHT ideologies; the association of a particular instance of a cue with one party or another is probabilistic, depending on other cues in the context.

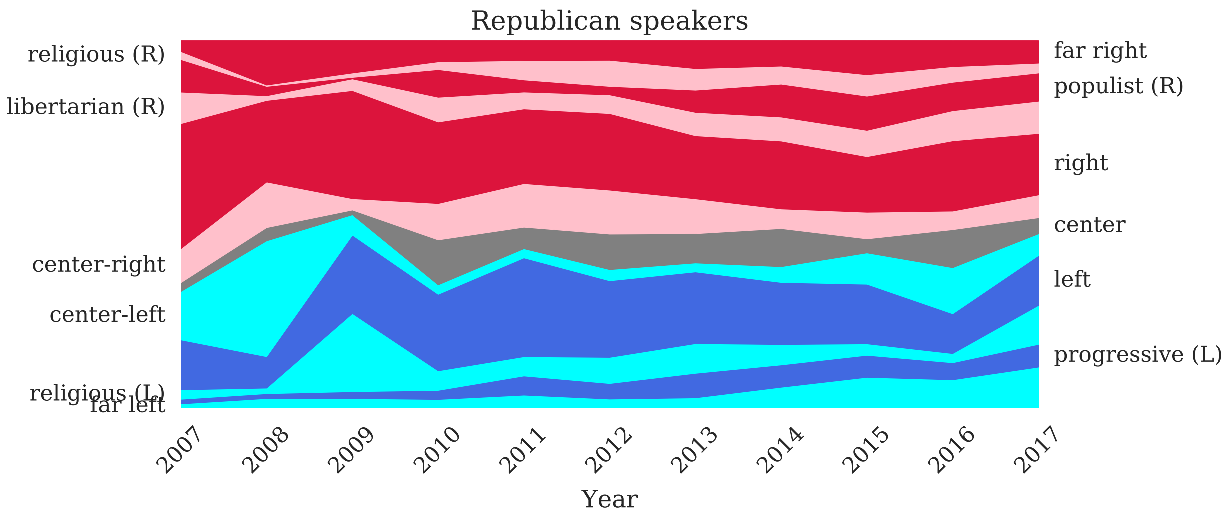
Top cues that do not overlap between parties include the more frequent use of "illeg_immigr" in Republican rulings, as well as the mentions of Donald Trump and Hillary Clinton in their respective parties. While the CLIP model takes more into account than just cue frequency, these lists suggest that it is not just the presence of a small number of dominant cues that are pushing the language in an ideological direction. (We look further at the usage of specific phrases with partisan leanings – for example, "illegal immigrant" compared to "undocumented immigrant" or "illegal alien" – in section 4.5.)

By ruling type. Fig. 14 shows the expected proportions across all RIGHT-leaning vs. LEFT-leaning ideologies by ruling from "Pants on Fire" to "True." Again, CLIP finds that language use tends to skew right across the different rulings. However, there is not a clear pattern in how proportions shift going from false to true; while more of the false rulings are about Republicans (and vice versa), the proportion of LEFT language does not increase with veracity.

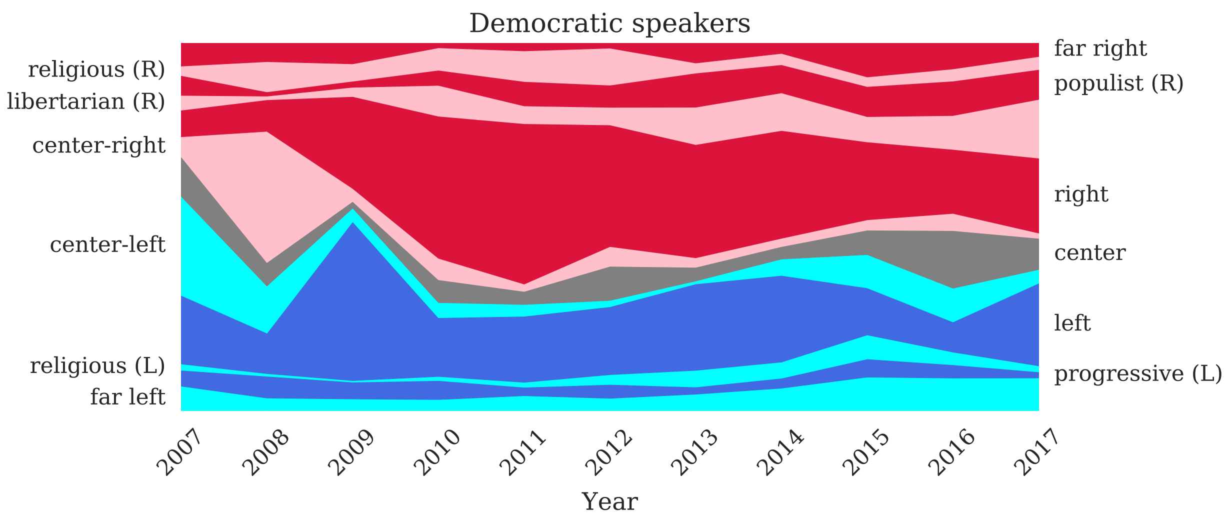
Specific speakers. Finally, we run CLIP on the two most frequent speakers in the Politifact rulings: Presidents Barack Obama and Donald Trump (Figure 15a and 15b). We also plot the LEFT vs. RIGHT proportions comparing Obama and Trump to their respective parties and the overall proportions in Fig. 16.

The broad LEFT/RIGHT ideological proportions for these rulings are similar for Obama and the Democratic rulings as a whole; Trump's proportions skew more rightward. However in both cases, the differences between the presidents and their parties seem to be topical in nature (that is, in response to claims that Obama/Trump made). For Obama, these include terms related to health care and the Affordable Care Act, his 2008 and 2012 presidential candidacies (e.g., his campaign, John McCain, Mitt Romney), and tax and business policies. For Trump, the most prominent cue terms are related to his 2016 presidential candidacy (e.g., Hillary Clinton, tax returns, voter fraud), immigration (note that both "illegal" and "undocumented" immigration terms appear), monetary numbers, and comments on the Obama administration. We note that, because the cue lexicon predates the Trump candidacy/election, it may not fully reflect more recent partisan issues.

Classification using proportions. To determine if there is any predictive value to the ideological proportions found (which might indicate differential language), we set up a classification task where the goal is to predict the speaker's party (Republican or Democrat) given the proportions of the article. Applying l_2-regularized logistic regres-



(a)



(b)

Figure 13: Output of the CLIP model for Republican (top) and Democratic (bottom) speakers.

Republican	**Democrat**
__MONEY___billion	__MONEY___million
unit_state	__MONEY___billion
__MONEY___million	unit_state
health_care	health_care
__NUM___year	__NUM___year
__NUM___million	__NUM___million
feder_govern	social_secur
presid_barack_obama	presid_barack_obama
health_insur	tax_cut
__MONEY___trillion	health_insur
white_hous	white_hous
illeg_immigr	rhode_island
social_secur	feder_govern
tax_increas	tax_rate
donald_trump	hillari_clinton
tax_cut	democrat_parti
tax_rate	__NUM___state
properti_tax	incom_tax
incom_tax	__MONEY___trillion
afford_care_act	__NUM___peopl
obama_administr	properti_tax
__NUM___job	__NUM___job
__NUM___state	tax_break
unemploy_rate	attornei_gener
suprem_court	minimum_wage

Table 4: Top 25 most frequent ideological cues in rulings of Republican (left) and Democrat (right) speakers.

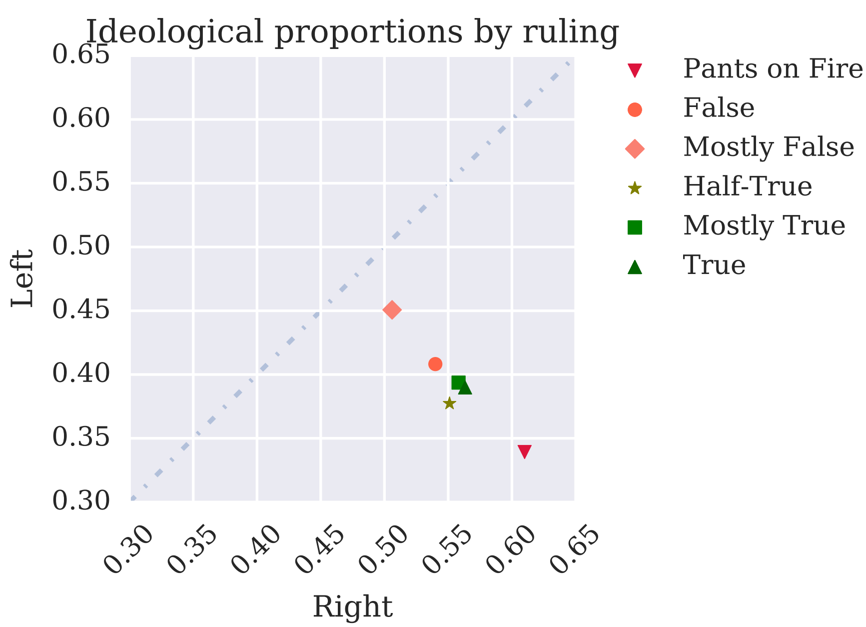
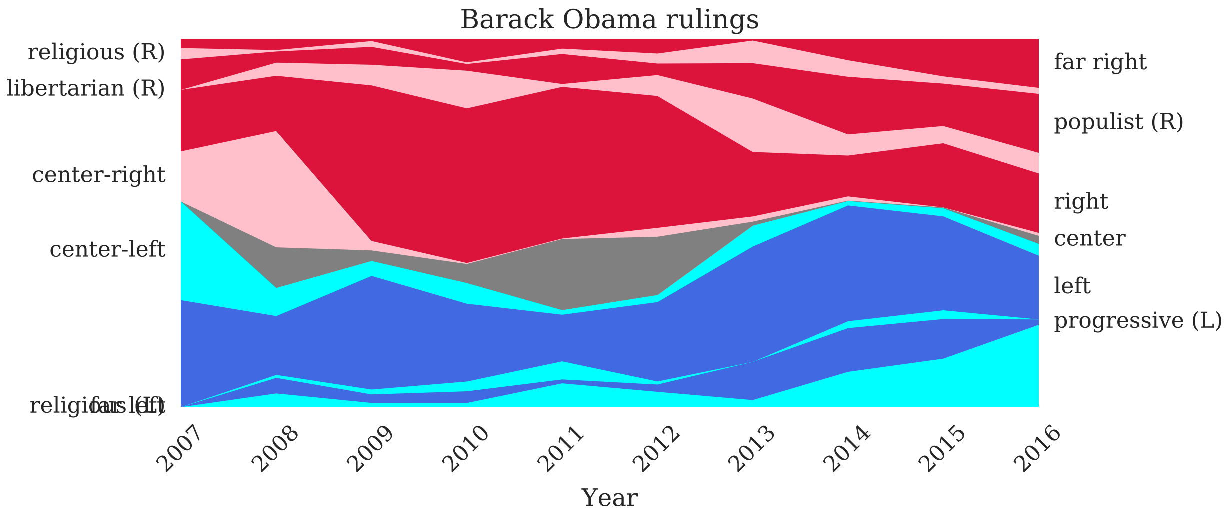
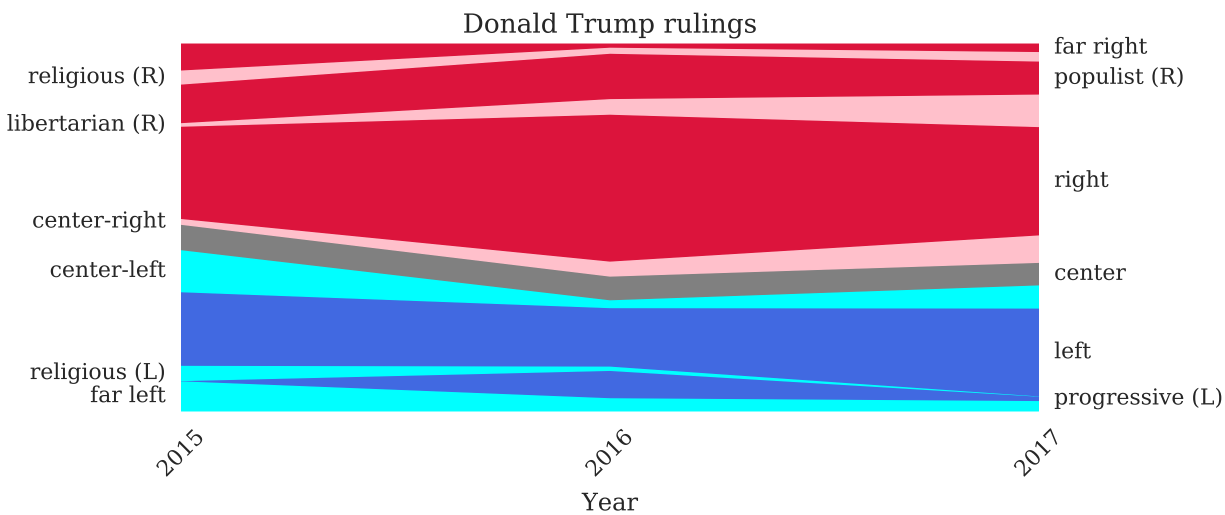


Figure 14: Output of the CLIP model by ruling type (from “Pants on Fire” to “True”), where the x and y axes are the expected proportions across left and right-leaning ideologies, respectively.



(a)



(b)

Figure 15: CLIP output for rulings on statements by Presidents Barack Obama (top) and Donald Trump (bottom). The range of years was clipped to contiguous years with five or more rulings.

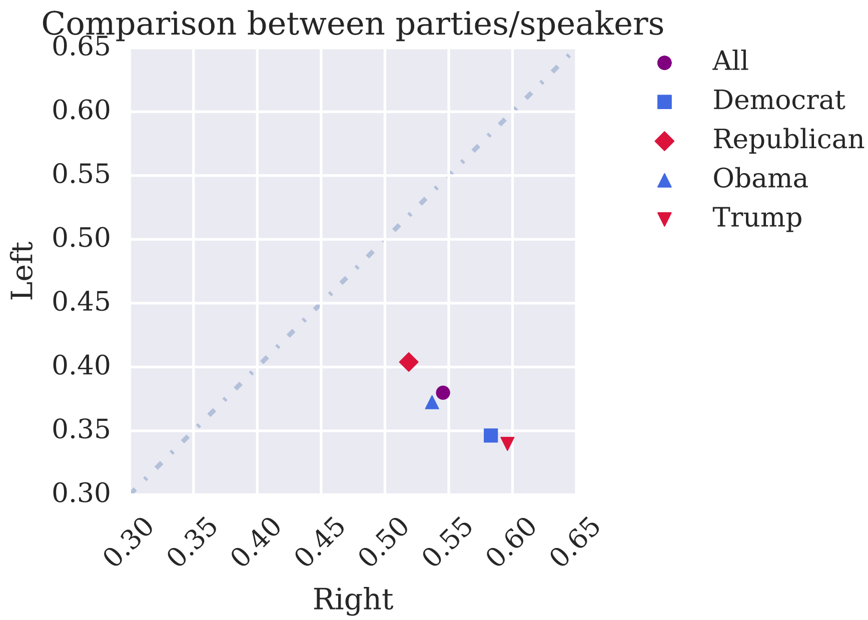


Figure 16: Output of the CLIP model comparing Obama and Trump to their parties (and all rulings). The x and y axes are the expected proportions across left and right-leaning ideologies, respectively.

sion results in approximately 61% accuracy.[14] Since predicting the most common party (Republican) would result in approximately 58% accuracy, the ideological proportions found by CLIP do not seem to have much predictive value.

We conclude this section by reiterating that the automatically induced lexicon used in this analysis does not make any attempt to disentangle topical and stylistic cues from other linguistic cues. Statistical association between a word and an ideology can arise for different reasons, and does not imply that a person using a word holds or is sympathetic to a particular ideology. Although the CLIP model explicitly aims to account for the ambiguity of these associations, style and topic effects remain difficult to remove. Further, the texts from which the lexicon was induced may be too distant in genre and time (they all predate 2014) to enable identification of bias in language of more recent Politifact articles.

4.5 Partisan Phrases

In the previous section, we looked at the Politifact corpus through a model which manages the uncertainty of cue terms' associations with different ideologies. Here, we focus on specific phrases from more recent literature which are established to be highly partisan, at least in the context of Congressional debate.

One way of looking at bias in language is the extent to which an author uses terms and phrases that "sound like" a particular person, group, or other source. As observed above, Democrats and Republicans tend to talk about different things, and in some cases have different language for discussing the same issue. In many cases, this may be the result of explicit political calculation about how best to frame an issue (see, for example, Lakoff et al., 2008). In describing political issues, an article may unintentionally end up sounding like something that was written by a particular party, depending on both the subject being discussed, and the specific language that is used.

Outside of well-known examples, however, it may not be obvious what language is associated with each party. In a National Bureau of Economic Research working paper, Gentzkow et al. (2016) investigate whether partisanship has increased over time by studying the partisanship of political language. Specifically, they analyze at the Congressional Record to find phrases (bigrams) that are much more frequently used by one party as opposed to the other. In doing

[14] Similar to §3.1, we split the data into 64% training data, 16% development data, and 20% test data, and use the development data to tune the regularizer strength during training. We compute the accuracy on the test portion for three different random splits of the data and report the average.

so, they are able to discover sets of "partisan phrases" for each session of Congress, after accounting for terms which occur differentially as a function of which party represents each state in each house, and who controls each branch of government.

Although this provides a starting point, there are several difficulties with this approach. First, this method does not necessarily distinguish between issues that are primarily discussed by one party from ways of discussing those issues that are highly partisan. For example, "Al Qaeda" appears as a highly partisan Republican phrase in the 114th Congress, presumably not because the Democrats use a different name to refer to this group, but because they rarely discuss it, relative to Republicans. Second, certain phrases may change polarity over time. For example, "American people" was a phrase strongly associated with the Democrats in 2007–08, and the Republicans in 2014–15. Third, some phrases may be quite generic, but register as partisan because they are associated with a specific piece of legislation, such as "mental health", which was strongly partisan in the 2014-15 because of the Republican-sponsored "Helping Families in Mental Health Crisis Act of 2016" (Gentzkow et al., 2016).

Because of the above difficulties, we do not attempt a full analysis of partisan phrases here. Rather, we consider a few examples based on the work of Gentzkow et al., to see how certain sets of partisan phrases are represented in the Politifact corpus. If, for example, Politifact tended to only use the terminology associated with one party, this might be a sign of differential treatment.

As an example of partisan phrases, consider variations on the concept of people who reside in the United States without authorization. There are two competing narratives about these "illegal immigrants." One is that they are "illegal aliens," who "steal jobs" and deny opportunities to legal immigrants; the other is that they are "undocumented immigrants," who are already part of the fabric of society and deserve a path to citizenship (Merolla et al., 2013). Figure 17 compares the usage of three competing terms in the Politifact corpus, re-centered to reflect the overall distribution of these terms in articles about Democrats and Republicans. [15]

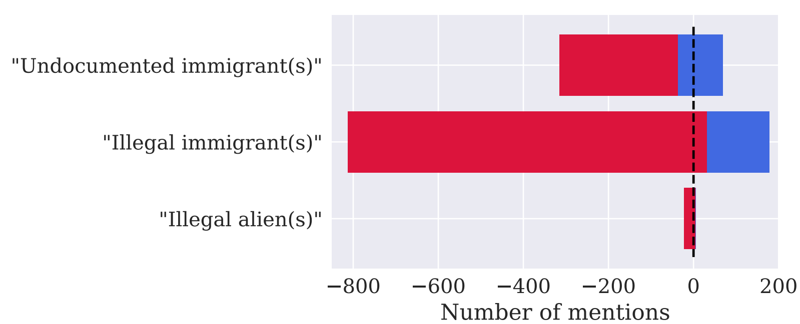


Figure 17: Comparison of the usage of immigration terms in articles about Democrats and Republicans.

We already know that more articles about immigration are focused on Republican speakers; what Figure 17 shows, however, is that Politifact primarily uses the relatively neutral "illegal immigrants," but also frequently uses "undocumented immigrants," regardless of which party is being discussed. However, the phrase "illegal aliens" is almost never used. As usual, this is based on Politifact articles from which quotes have been removed. Although the analysis above suggested that Politifact's coverage is relatively balanced in terms of pro- and anti- immigration language, a reader who strongly prefers one of these terms might perceive this imbalance as a liberal bias.

Another related pair of terms has to do with energy, which is primarily a Republican issue. In particular, Gentzkow et al., found that "natural gas" is a strongly Republican term, whereas Democrats are more likely to talk about "oil companies." Figure 18 compares the use of these terms, along with the usage of "solar power" / "solar energy", and "nuclear power" / "nuclear energy" for comparison, terms which seem likely to have their own partisan associations. The results are largely in agreement with the work of Gentzkow et al., and our own intuitions, although "natural gas" is actually relatively balanced in its usage by Politifact.

Finally, Figure 19 shows the usage of four phrases that are among the most partisan for the 2014–15 Congress, all of which represent a kind of threat. This seems to be a fairly clear case of the parties talking about different topics,

[15] "Illegal immigrants" was one of the most strongly-partisan Republican terms in the 110th Congress, according to Gentzkow et al. (2016).

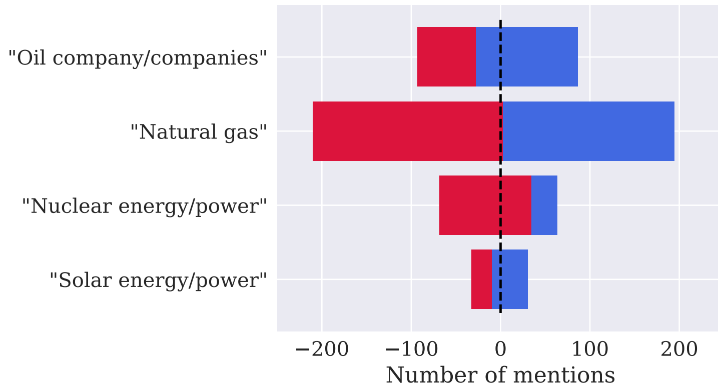


Figure 18: Comparison of the usage of energy terms in articles about Democrats and Republicans.

rather than using different terms, but it is still interesting to see how frequently these phrases appear in the Politifact corpus.[16] Unsurprisingly, the usage of these terms is correlated with party. More importantly, the vast difference in total mentions suggest that someone who does not believe in climate change, or who thinks "radical Islamic terrorism" is a major threat to America, might conclude that Politifact has a different set of priorities. We do not take this as a sign of differential treatment (style), but rather as an issue of content.

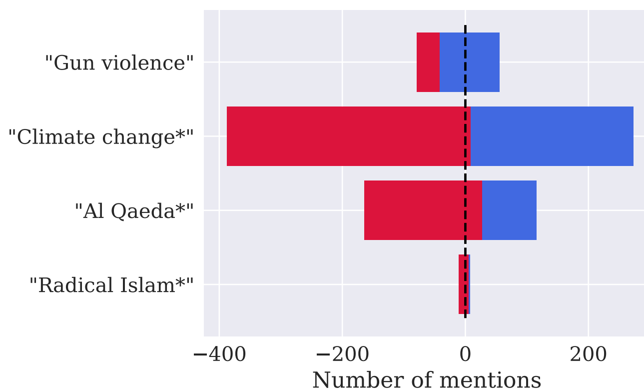


Figure 19: Comparison of the frequency of usage of partisan threat terms in articles about Democrats and Republicans

Given that political parties actively engage in attempts to reframe issues to emphasize particular aspects or interpretations, what counts as partisan speech is a constantly moving target. Moreover, it is difficult to discuss certain subjects, such as immigration, without using terms that can sound partisan to some audiences. The brief analysis in this section highlighted a few examples of highly partisan phrases (as determined by the Congressional Record) and the frequencies with which they are used by Politifact, noting that certain phrases strongly associated with the Republican party – "illegal aliens" and "radical Islam" – are rarely used in Politifact's fact-checking. While we do not advocate for trying to achieve parity across all ways of framing an issue, we include this analysis to emphasize how the choice of language may have an impact on how Politifact articles are perceived by partisan readers.

For a more complete analysis, it might be beneficial to find more examples of relatively partisan synonyms (as in the case of immigration), and compare their usage by Politifact to that of left- and right-leaning news sources. This might provide greater insight into how Politifact articles appear to readers that strongly identify with one or the other political party.

[16]The full set of terms queried for this graph are "gun violence", "climate change", "climate change's", "climate changes", "al qaeda", "al qaeda's", "al qaida", "al qaida's", "radical islam", "radical islamic", "radical islamism", "radical islamist", and "radical islamists".

5 Conclusion

Our analyses were not able to detect any systematic differences in the treatment of Democrats and Republicans in articles by Politifact. We offer three suggestions for follow-up study or tracking in future. First, are the types and amount of evidence offered in articles similar for members of both parties? This is beyond the scope of text analysis tools, but it should be relatively straightforward for experts to code a sample and perform a comparison. Second, language used in stories about guns may signal a pro-gun control stance. The tools used for this analysis are not a perfect fit for the data; we therefore suggest a small user study in which users' perceptions of slant are compared to their own positions on the gun control issue. Finally, we suggest comparing the rates of usage of known-partisan phrases in Politifact articles to rates in other text collections whose partisan status (or neutrality) is established.

Acknowledgments

This research was supported in part by funding from the Knight Foundation, in cooperation with Politifact. LHL is supported in part by an NSF Graduate Research Fellowship. The findings reported here do not necessarily represent the views of anyone other than the authors.

References

D. Card, A. E. Boydstun, J. H. Gross, P. Resnik, and N. A. Smith. The media frames corpus: Annotations of frames across issues. In *Proceedings of ACL*, 2015.

M. Gentzkow, J. M. Shapiro, and M. Taddy. Measuring polarization in high-dimensional data: Method and application to congressional speech. July 2016.

G. Lakoff, H. Dean, and D. Hazen. *Don't Think of an Elephant!: Know Your Values and Frame the Debate*. Chelsea Green Publishing, 2008.

J. Merolla, S. Karthick-Ramakrishnan, and C. Haynes. "illegal," "undocumented," or "unauthorized": Equivalency frames, issue frames, and public opinion on immigration. 11(03):789 –807, 2013.

S. M. Mohammad and P. D. Turney. Crowdsourcing a word-emotion association lexicon. 29(3):436–465, 2013.

B. O'Connor. MITEXTEXPLORER: Linked brushing and mutual information for exploratory text data analysis. In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pages 1–13, 2014.

Y. Sim, B. D. L. Acree, J. H. Gross, and N. A. Smith. Measuring ideological proportions in political speeches. In *Proceedings of EMNLP*, 2013.

C. Tan and L. Lee. Talk it up or play it down? (un)expected correlations between (de-)emphasis and recurrence of discussion points in consequential u.s. economic policy meetings. Presented in Text as Data, 2016.

Appendix: Additional Tables and Figures

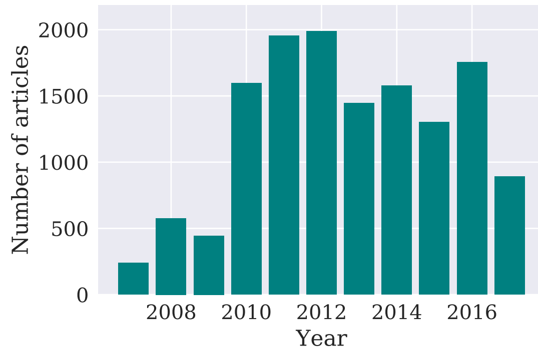


Figure 20: Articles per year

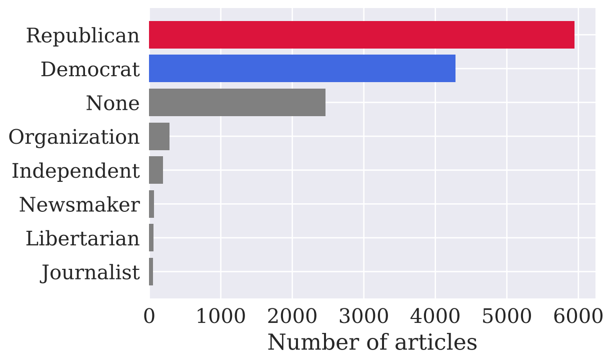


Figure 21: Articles per party

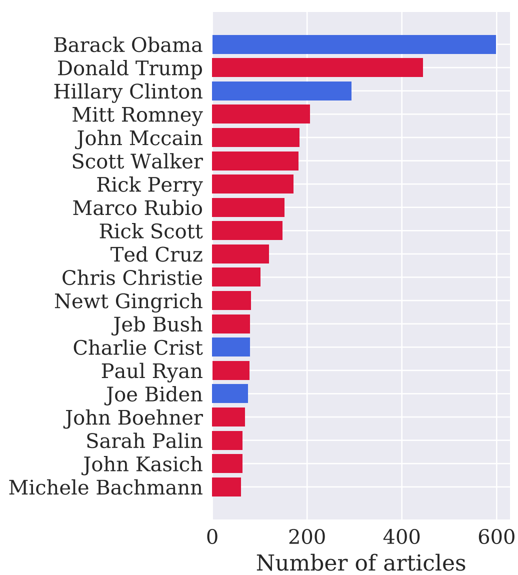


Figure 22: Articles per speaker (top 20).

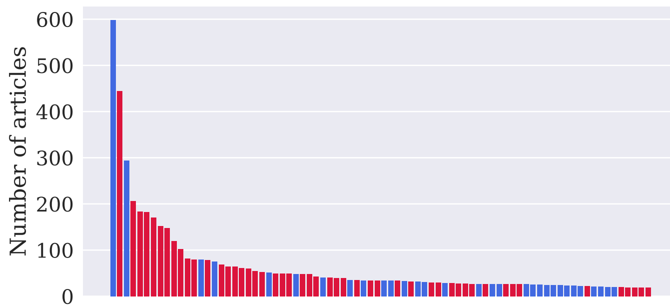


Figure 23: Articles per speaker (top 80).

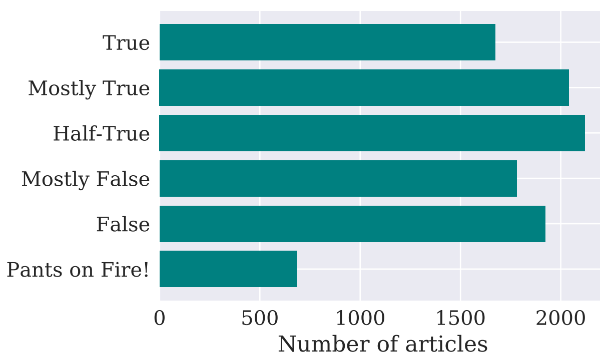


Figure 24: Articles per rating

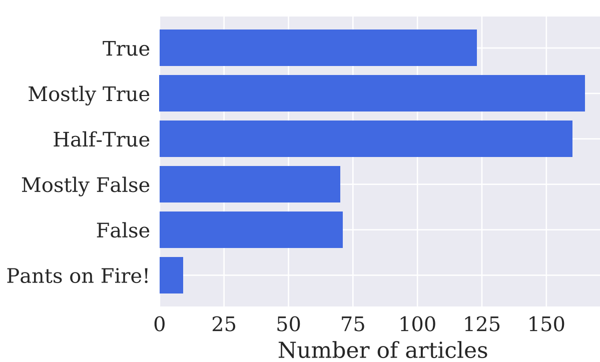


Figure 25: Articles about Obama

Figure 26: Articles about Trump

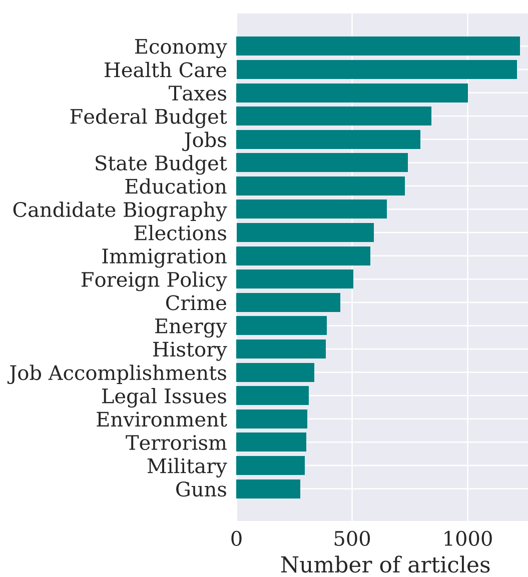


Figure 27: Number of articles per subject.

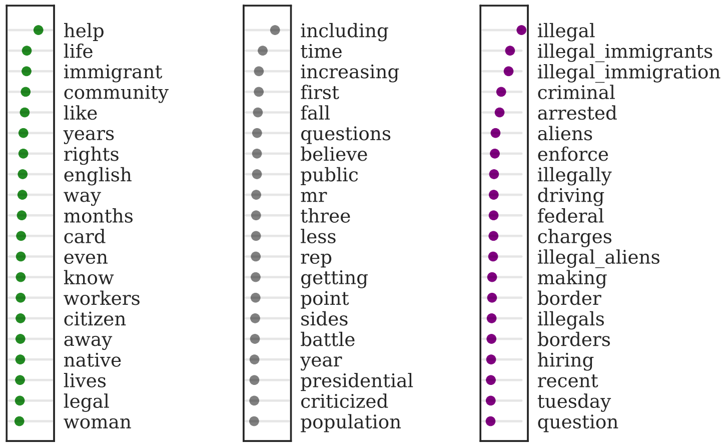


Figure 28: Highest impact terms in immigration tone classifier for **pro-immigration** (left), neutral (center), and **anti-immigration** (right).

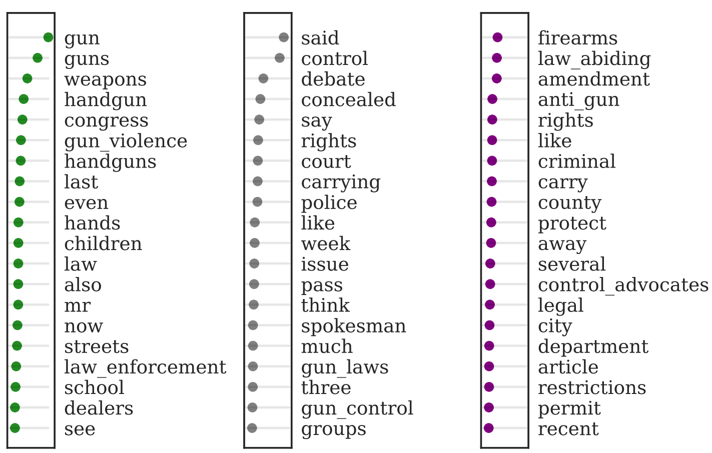


Figure 29: Highest impact terms in gun control tone classifier for both **pro-gun control** (left), neutral (center), and **anti-gun control** (right).

expire tax hikes brackets cuts taxation excise deductions
D: millionaires deductions breaks deduction loopholes class wealthy middle
R: excise hikes cigarette hike fee increases resolutions franchise

taxed earnings richest earners industrialized income median investments
D: hedge occupations kingdom men earnings gap earn wages
R: industrialized corporate oecd deductions burden statutory capita filers

poll polling respondents polls polled gallup quinnipiac pollster
D: polled gun checks respondents poll popularity quinnipiac nra
R: pollsters poll polling polls rasmussen scientists pollster gallup

marriage supreme court's gay justices court constitutional unconstitutional
D: discrimination incest personhood marriage gay rape nominees abortion
R: violate court's scholars ginsburg religious roosevelt clause constitutional

syria nato troops afghanistan iraq iraqi qaida missiles
D: withdrawal afghanistan commander vietnam troops iraq trans bin
R: iran iran's refugees iranian bomb guantanamo russian syrian

petroleum pipeline drilling plants coal oil electricity gas
D: solar renewable electricity crude oil leases petroleum mine
R: epa carbon emissions greenhouse pipeline gallon eia dioxide

scores graduation graders grade reading graduating graduates dropout
D: dropout graduation graduating students scores colleges graduates rankings
R: graders temperature exam core scientists scientific curriculum earth

abortion abortions gun parenthood incest guns firearms planned
D: incest abortion firearms dealers rape gun firearm abortions
R: abortion payer arms abortions core treaty gun bargaining

transcript video reporter audio quote events nexis footage
D: screen narrator allegation clip social privatization airing ad
R: protesters benghazi protests father minister attended biography traveled

indicating statesman click demographer san dallas abortions property
D: cities city's demographer residents metropolitan paso san metro
R: patrol border abortions apprehensions apprehended rio grande comptroller

jobs losses sector job creation bls manufacturing employment
D: jobs sector counting outsourcing employment growth creation job
R: stimulus advisers bls jobs seasonally reinvestment gains cbo

ballots elections voter clerk voting supervisors ballot absentee
D: ballots provisional dekalb id voter boards board's supervisors
R: pension elections salary appointed municipal voter primaries complaint

crimes convicted police offenses enforcement criminal fbi's homicides
D: fatal weapon firearms shootings homicides guns gun shooting
R: refugees ice refugee deported customs deportation syrian sanctuary

deficit discretionary surplus deficits entitlement entitlements spending budget
D: deficit surpluses security sequester surplus social deficits discretionary
R: debt medicare trillion obligations sequestration outlays wars budgetary

wage minimum teachers teacher hourly tuition classroom schools
D: wage minimum discrimination earnings hourly teacher occupations teachers
R: passel font decoration width rgba padding webkit float

donors responsive donations donation pacs donated parenthood loans
D: responsive donors pacs fundraising loans donations student donor
R: donated parenthood donations investigation arabia securities font saudi

insurers medicare beneficiaries marketplaces wilensky seniors marketplace premium
D: voucher traditional pocket beneficiaries medicare beneficiary seniors immigration
R: advantage illegal immigrants amnesty fencing enrollees fence marketplaces

earmark earmarks bridge pork tarp bailout projects stimulus
D: asset bailout auto banks derivatives automakers tarp mortgage
R: pork freddie earmarks earmark bridge stimulus fannie wasteful

gdp presidents unemployment presidency decline fallen risen burtless
D: wage minimum adjusted seasonally surpluses barrels peaked inflation
R: debt burtless gdp missed presidents trillion unemployment recessions

cdc disease medicaid uninsured poverty insured cancer stamps
D: contraception marijuana opioid veterans reproductive uninsured deaths medicaid
R: premium patients insured poverty premiums medicaid marketplace parent

Table 5: Full list of topics and topic-party interactions.