An analysis of perspectives in interactive settings

Dong Nguyen
Carnegie Mellon University
Language Technologies Institute
Pittsburgh, PA 15213
dongn@cs.cmu.edu

Elijah Mayfield
Carnegie Mellon University
Language Technologies Institute
Pittsburgh, PA 15213
elijah@cmu.edu

Carolyn P. Rosé
Carnegie Mellon University
Language Technologies Institute
Pittsburgh, PA 15213
cprose@cs.cmu.edu

ABSTRACT
In this paper we investigate the effect of the context of interaction on the extent to which a contributor's perspective bias is displayed through their lexical choice. We present a series of experiments on political discussion data. Our experiments indicate that (i) when people quote contributors with an opposing view, they tend to quote the words that are less strongly associated with the opposing view. (ii) Nevertheless, in quoting their opponents, the displayed bias of their word distributions shifts towards that of their opponents. (iii) The personal bias of the speaker is displayed most clearly through the words that are not quoted, (iv) although characteristics of the quoted message do have a measurable effect on the words that are included in the contribution. And, finally, (v) posts are influenced by the displayed bias of previous posts in a thread.

Categories and Subject Descriptors
I.2.7 [**Artificial Intelligence**]: Natural Language Processing—*text analysis, discourse*

General Terms
Experimentation

Keywords
ideology, perspective, sentiment analysis, political discourse, social media

1. INTRODUCTION
This paper presents an investigation into the effect of interaction on the representation of perspective biases in text as measured through word distributions. Previous research has separately either investigated how word distributions are indicative of perspective biases or explored how patterns of interaction are influenced by the perspective biases of conversational participants. Our evaluation offers evidence that word distributions change depending on who the speaker is responding to, which thus argues in favor of taking an integrated approach.

The idea of investigating perspective biases of authors using a variety of word distribution modeling techniques is certainly not a new idea. State-of-the-art solutions to detection of perspective bias employ a variety of statistical techniques using the *bag of words* model. When trying to classify a person as politically left or right, for example, the association between each word and a perspective is taken as evidence for one side or the other, and the side with more evidence "*wins*".

Much evidence that both text and talk are heavily influenced by the perspective of the author or speaker can be found in prior work not only in the language technologies community (e.g. [12]), but also in the social sciences (e.g. [10]). Understanding and modeling this phenomenon has many practical applications. One example is to predict based on measured perspective bias whether a speaker will vote for or against a bill. Another example is monitoring discussion forum posts in order to track changes in popular opinion about political issues over time. Prior research has formulated this task as either a discrete categorization task (e.g. [12]) or as a rating problem, where the goal is to indicate through a point-based system the extent to which a perspective bias is displayed in text (e.g. [10]). These approaches tend to treat texts as representing the bias of the author only and do not consider how characteristics of the context, including the audience, or other interlocutors, exert a separate influence on the formulation of the text.

In this paper we will use data from an online political forum. The war in Iraq will be used as a running example throughout our paper. Consider the following example interaction, beginning with a user with a left affiliation:

> **Left person:** *does anybody really think that the election will stop the violence? [...] and if the election does not stop the violence, how in the hell will we ever get out of iraq?*
>
> **Right person:** *no, the election will begin the democratic process. there is no way to stop violence, we have violence here, every country has violence, and to have some polyannish idea that the violence is going to magically stop, is stupid. nearly 80% of the country is ready to have elections and move forward,[...]*

1st Workshop on Social Media Analytics (SOMA '10), July 25, 2010, Washington, DC, USA.

> **Left person:** *whoever thought that an election held at the point of a gun would ever wipe away the decades of simmering hate between sunnis, shi'ites and kurds and that, [...] once the "election" was held that they would all just start to work together for the creation of a vibrant, pro-western, american style democracy?*

We see here first evidence of the affiliation of each poster in their choice of words, but also evidence of the influence of posts on responding posts. For example, we see the initial left affiliated poster mentioning "*violence*" twice, which can be identified within this data as being left associated, using the model discussed later in the paper. However, the right affiliated user who responds uses this term several times. His own affiliation comes out in the reference to "*democracy*", and to a lesser extent "*elections*", which is also right associated but less strongly. The many references to specific people groups in the next post is very characteristic of the left style of posts in our data. However, we also see evidence of adopting the terms used in earlier posts in order to continue the discussion in a cohesive manner. In order to properly model the relationship between word distributions and perspective bias, we must account not only for the contribution of the speaker's own perspective bias, but also characteristics of the interaction that provide the context for the contribution.

Our contributions are twofold. We first propose a method that estimates the political orientation of text (section 3). We then present a series of experiments using this method to explore the effect of speaker affiliation and characteristics of the context on the measured bias (section 4). As part of this analysis, we will examine quoting behavior as well as the influence of contextual factors at multiple levels of abstraction (i.e., thread and quoting level).

2. RELATED WORK

Van Dijk [20] and other researchers from the critical discourse analysis tradition discuss how ideologies influence language and discourse, which therefore also influences how people acquire, learn or change ideologies [21]. An example of ideological language given by Van Dijk [20] is calling a group of people *"terrorists"* rather than *"freedom fighters"*. He furthermore states which information in a text is foregrounded or backgrounded depends on importance and relevance and thus is influenced by the ideology of the speaker. For example, on a topic such as amnesty for illegal immigrants, anti-amnesty proponents will focus more on the negative impacts of the immigrants, while the pro-amnesty side would focus on positive aspects such as the contribution of these immigrants to society.

Nevertheless, within this same discourse analysis tradition we can find reference to the impact of the context of an interaction on the formulation of a contribution to that interaction. For example, Kristeva [9] coined the term intertextuality, which refers to the way text and talk refer and build on other texts. Momani et al. [6] studied intertextual borrowings from opposing ideological text in political discourse. Through this process, speakers may quote other people, which might make their word distributions resemble those of their interlocutors. Nevertheless, Momani and others argue that when people borrow text from ideologically opposing views, they often select quotations that serve their own purposes, and thus do not represent the same thing that they do when they are employed by their interlocutors. This thus suggests that conversational text should not be treated in isolation when modeling perspective bias.

Analysis of the manner in which political ideology influences presentation of self through text is a major research topic in the political sciences. Estimating the policy positions of political persons has been a widely addressed research topic. Early approaches hand coded texts, which is highly labor intensive. This led to research exploring automatic methods for this problem. Developments in computational linguistics and machine learning led to the exploration of statistical approaches. Laver et al. ([10]) proposed the so called word scores approach, where the political orientation of a text is calculated by scoring every word depending on the probability of word in training documents and the political orientation of these documents. The statistical approach has shown to be effective. Lin et al. [11] showed that document collections representing different perspectives can be successfully distinguished from other types of collections based on word distribution divergences.

Classical machine learning techniques such as SVM and Naive Bayes have been applied to classify the political leanings of blogs (such as [5]). These techniques however do not model the generative process displaying perspective bias through text. Yu et al. [22] tried to classify whether a person would vote for or against a bill. However their performance decreased dramatically when they trained and tested across multiple debates representing a variety of topics rather than a specific debate covering only one topic. They point out that SVM classifiers seek to choose discriminative words with broad coverage. Unfortunately it is possible that such generalizable features are very rare. Some words can convey strong political opinions in only particular debates, and therefore are not picked up by classifiers such as SVM when a general model is trained that cuts across a variety of topics. Thus, the contribution of topic as an influencing factor on word choice must also be considered.

Topic modeling approaches have become very popular for modeling a variety of characteristics of unlabeled data. A well known approach is Latent Dirichlet Allocation (LDA) (Blei et al. [3]), which is a generative model and is effective for uncovering the thematic structure of a document collection. Two models that are specifically tailored to the problem of modeling different perspectives are the cross-collection Latent Dirichlet Allocation (ccLDA) model (Paul and Girju [16]) and the joint topic and perspective model for ideological discourse (Lin et al. [12]). Both assume that the frequency of a word depends on the relevance in the topic and on the perspective of the speaker or author. ccLDA [16] builds on the standard LDA model [3] and the cross-collection mixture model (ccMix) by Zhai et al. [23]. ccLDA discovers the topics across multiple text collection and estimates for each topic a shared distribution and collection specific distributions. The model of Lin et al. [12] assigns every word a topical weight indicating how often it was chosen depending on the topic, and an ideological weight which depends on the perspective of the speaker or author. However, Lin's model does not distinguish between different topics, but assumes all text is about the same topic. Neither approach takes the influence of context on lexical choice into account.

Instead of looking at the textual content, research has also been done that exploits behavioral patterns of users responding to users with different viewpoints. Malouf and Mullen [13] and Agrawal [1] both pointed out that perspective classification in social media is very hard, because of the informal nature of text present in media such as online forums and newsgroups and the fact that people are talking about the same topics. They showed that classification performance improved, when they made use of the observation that people tend to react to people they disagree with. The drawback of these approaches is that often information such as who the user is responding to is not available. But more importantly their use of the textual content was very limited.

Thus two lines of research involving perspectives have been observed. The first uses textual analysis, the second exploits interaction information (such as quoting patterns). However, we are not aware of research that combines these two aspects to analyze perspectives in text as we do in this paper.

3. ESTIMATION OF POLITICAL ORIENTATION

In order to be able to investigate changes in bias in text due to interaction factors, we need to have a method to measure the bias in text. In this section we present and evaluate various methods to measure the bias in text. In Section 4, we will introduce an online political forum that we will use for our analysis of perspectives in interactive settings. However, here we are using a different dataset for practical reasons that we will outline next.

3.1 U.S. Floor debates dataset

The political debates dataset consists of transcripts of U.S. floor debates from the year 2005 and is provided by Thomas et al. [18]. Each debate consists of a series of speech segments. Every segment is annotated with a speaker id and the party of the speaker. The provided data is split into a training, development and test set. We use the *stage three* version of the data as provided by Thomas et al. [18]. This version contains less noise, because single sentence speech segments containing the word *"yield"* are removed. The test set contains 860 speech segments. Speech segments by speakers marked as Independent will not be used in our experiments.

An advantage of this dataset is that the text is very clean, because the language is very formal and the transcripts are clean (e.g. no misspellings). In addition the amount of data per political affiliation is balanced and there are a lot of different speakers participating in the debates. Furthermore for our purpose to evaluate and compare different methods, this dataset also contains enough data to split into a reasonably sized train and test set.

All text is lowercased and terms occurring less than 10 times are discarded. Furthermore a stopword list is applied.

3.2 Method

This section outlines our general approach for estimating the bias in text. We take a topic modeling approach to estimate the word distributions for left and right text. We then estimate the bias of a text as a combination of the individual bias estimates of each word.

3.2.1 Topic modeling

Taking a topic modeling approach gives us the advantage that we can exploit the topic information when determining the political orientation of a word, because some words carry a stronger political orientation in connection with particular topics. Furthermore, this gives us an opportunity to perform an analysis on the extracted topics themselves. We apply LDA to build a topic model using a pre-specified number of topics. We then calculate post-hoc collection specific topic distributions, by calculating for each collection the topic distribution only over the documents from that collection. For example we can have a collection with documents written by left people, and a collection with documents written by right people. We build a standard LDA model over all documents, but then for each topic we also calculate a left distribution by calculating the distribution only over the documents from the left document collection and similarly for right. We use the Lingpipe toolkit [2] to train the LDA model.

3.2.2 Estimation of bias

We propose to estimate the bias of a text as a combination of the individual bias estimates of each word. The bias for a particular word depends on the topics sampled for the text and how distinguishing the word is for a particular affiliation. We first sample the text using Gibbs sampling to estimate the topics of the tokens in the text. The bias of the text is then calculated by looping over all tokens in the text and adding the bias estimation of each token, which indicates how left (or right) the token is for the sampled topic. We compute the average over 50 iterations for every text in order to get a more stable value.

Our general approach builds on the intuition that a word is more distinguishing for a particular affiliation if it has a high probability associated with that affiliation and a low probability for the other affiliation (e.g. a high probability for left and a low probability for right). This idea has already been used in previous research (e.g. [19], [16]) to extract distinguishing words when comparing distributions. For every affiliation we therefore order the words in descending order according to the probability of the word in the collection specific distribution of that topic. Each word is given a rank, with the first word having a rank 0. The resulting formula for a word w in topic t looks as follows:

$$bias(w,t) = \log(rank_{right}(w,t)+1) - \log(rank_{left}(w,t)+1)$$

In this formula a word gets a positive value if it is more distinguishing for left, while it gets a negative value if it is more right. For example a word that ranks 5^{th} in the left distribution and 100^{th} in the right distribution is more distinguishing of left than a word that ranks 10^{th} in the left distribution and 20^{th} in the right distribution. The use of a logarithm gives the difference in ranking more weight when the words have higher probability. We furthermore experiment with a variant that only takes into account a word if the word appears in the top r words of the particular topic for at least one collection. To investigate the effect of adding topical information, we also experiment without topical information (thus having only one left and right distribution pair). Thus we have the following methods:

Increase rank Our proposed method with topic information and no threshold

Increase rank, r=X Our proposed method with topic information and threshold X

Increase rank, no topics Our proposed method without topic information

3.3 Evaluation

In order to evaluate our bias estimation approaches, we build an LDA model with 10 topics over the debates dataset. We compare the classification accuracy of our bias estimation variants with two baselines, Naive Bayes and the majority baseline. For the variant where only highly occurring words are taken into account we set r=50, which we determined experimentally using a development set. We evaluate the accuracy on individual speech segments (Table 1) and speech segments aggregated per person (Table 2).

As can be seen in Table 1, both bias estimation methods perform better than Naive Bayes. The accuracy is relatively low however, which can be explained by the fact that we try to classify individual speech segments, which often do not convey a strong perspective or have a neutral nature (for example organizational speech segments). An example of such a speech segment is *"mr. speaker , i demand a recorded vote."*.

To see if the algorithms are better in classifying longer texts, we aggregate all texts of the same person in the test set. As can be seen in Table 2, this improves the performance. Note that we only use text in the test set, which therefore assigns for some persons still a small amount of text. The performance of increase rank increases less than the other methods (increase rank with r=50 and Naive Bayes). Increase rank with r=50 outperforms the other methods, and therefore is the preferred method for our next experiments.

We furthermore experiment with using the same bias estimation mechanism but without the topics. This method performs lower than the ranking mechanism that takes topics into account in both tests (increase rank without threshold). This suggests that adding topical information is effective, however the difference in between "increase rank" and "increase rank, no topics" is not statistical significant. Note further that when aggregating all text of persons (Table 2) topical information makes less of a difference since this aggregates text from multiple debates, making topic information less useful.

4. INTERACTIONAL DYNAMICS

We will first briefly outline our planned experiments. We then discuss our dataset, our experiments and results.

Table 1: Classification accuracy speech segments

Method	Test set accuracy (%)
Increase rank, no topics	61.88
Increase rank	63.52
Increase rank, r=50	65.38
Naive Bayes	62.70
Baseline (majority)	50.80

Table 2: Classification accuracy persons

Method	Test set accuracy (%)
Increase rank, no topics	66.80
Increase rank	67.84
Increase rank, r=50	72.96
Naive Bayes	69.87
Baseline (majority)	50.62

4.1 Overview experiments

We analyze the influence of interaction on perspectives with two different experiments.

Topic modeling experiment
With this experiment, we build a topic model, where text is now allocated to a collection depending on the interaction context (such as a left person responding to a right person or a left person responding to a left person). By analyzing the extracted topics and collection specific distributions we are able to analyze and compare word usage in different interactive contexts.

Bias estimation experiment
With this experiment we apply our method to estimate the bias in text to the online forum. We then perform statistical analysis to analyze the separate and joint effects of interaction factors on these bias estimates.

4.2 Online political forum dataset

Characteristics
The dataset is extracted from the forum of the website politics.com and is provided by Malouf and Mullen [13]. Users are able to indicate their political stance in their user profile. They modified these descriptions and assigned one of the following labels to the users: republican, conservative, r-fringe, democrat, liberal, l-fringe, centrist, independent, libertarian, green, unknown. These categories (excluding unknown) are grouped by Malouf and Mullen into 3 main categories: right, left and other. Table 3 presents some statistics about the dataset.

Preprocessing
First we filtered threads that were largely off topic. We

Table 3: Dataset forum politics.com

Total number of users	408
Left users	96
Right users	88
Number of threads	3861
Number of posts	77,854
Date	Dec. 2004 - June 2005

manually created a list with common political words (such as *"republican, democrat, bush, government, economy", etc.*). For each thread we counted how many times these words occurred in the thread. If the count was less than 4, we discarded the thread. Using only this crude heuristic we already removed 1446 of the 3861 threads.

Next we normalized the text by lowercasing all text and discarding terms occurring less than 20 times. Stop words are filtered using a stop list. Usernames, username variants and quotes are removed from text. We wanted to filter user references from training data in order to prevent our model from overfitting on these terms. However, only filtering the full usernames is not enough, since we observed that users often do not refer to other users with their full usernames, but often use variants such as abbreviations. We therefore automatically created for each username a list of variants. For example, we added variants such as *This* or acronyms (*TIAU*) when the username is *ThisIsAUsername*.

Extraction of interaction context
To each post we assign a label indicating the affiliation of the user and the affiliation the user is responding to. We view these labels as an indication of the interaction context the post was written in. Posts written by a user or responding to a user of which the affiliation was *unknown* or *other* are discarded. To each post we assign one of the following labels:

- **LR**: *left* responding to *right*
- **LL**: *left* responding to *right*
- **L**: The posts by *left* users for which the interaction context could not be determined
- **RL**: *right* responding to *left*
- **RR**: *right* responding to *right*
- **R**: The posts by *right* users for which the interaction context could not be determined

Malouf and Mullen [13] automatically extracted quoted data from their political dataset. Using these quotes, we can identify the post a user is reacting to. Unfortunately, most of the users do not explicitly quote a post they are reacting to. We apply the following heuristics if the post does not contain a quote to extract the affiliation the user is responding to. The first heuristic is formed by the observation that users often mention the poster he is reacting to. We match on these references using our automatically created list of username variants and substrings of usernames. If no username can be matched, we employ the following heuristics. The first post does not react to a particular affiliation, and the second post reacts to the first post. Also, when the previous post and the first post of a thread carry the same affiliation, we also mark the post as reacting to that affiliation. Furthermore, when all previous messages are from the same affiliation, this is also the response affiliation.

4.2.1 Comparison with the debates dataset

While the debates dataset is cleaner, we consider the political forum to be more appropriate for an analysis of the influence of interaction on bias representation.

Table 4: Representative words in a topic about war in Iraq with interactional collections

Global	Left replying to Left	Left replying to Right
iraq	wolfowitz	icbm
weapons	intelligence	dubya
saddam	qaeda	wmd
bush	iran	osama
war	defense	spin

Table 5: Representative words in a topic about war in Iraq (global, and separated by political leaning)

Global	Left	Right
iraq	shi'ite	whine
war	baghdad	defeat
military	country	democracy
iraqi	invasion	strategy
american	army	enemy
troops	sunnis	war
bush	men	mission
people	bush	freedom

Although both datasets are from an interactive setting, we observe that in the online forum people often address others directly when responding to each other. In contrary, speakers in the debates dataset do not directly speak to a particular person but address other persons indirectly, such as *"mr. speaker, i will be happy to respond to the gentleman"* or *"it is very clear that the gentleman from colorado"*. There is also a difference in the way users represent themselves. Speakers in the congress often have a complex, more hidden agenda. Not only the party they belong to, but also the district they are representing and their own beliefs influence the way they present themselves. In contrast, users in an online political forum are often anonymous and therefore can speak freely. There is no reason to assume they are not speaking primarily on their own behalf.

4.3 Topic modeling experiment

In section 4.2 we have described how we determine the interaction context of a post. We build a topic model with 6 collections, where the collection represents the interaction context of the text. We will have the following collections: left, left replying to right, left replying to left, right, right replying to left and right replying to right. We build a topic model with 10 collections and calculate post-hoc collection specific distributions as described in section 3.2.1.

We observed that the topics and the extracted distinguishing words were of less quality in general, because the amount of data per collection is less compared with only using two sides (left versus right). Due to the data sparsity we have not performed an extensive analysis. However, some topics did have a good quality. For example, Table 4 presents the most distinguishing words for two collections (left replying to left, and left replying to right) for a topic about the war. Especially the top words of left replying to right are striking. For example "*dubya*" is a nickname for president Bush often used by people who are criticizing him or the conservatives in general. Another example is the word *"spin"* referring to *"spinning the truth"*.

4.4 Bias estimation experiment

With this experiment we analyze the separate and joint effects of interaction factors on bias representation in text.

4.4.1 Methodology

We build a topic model from the online forum dataset with 15 topics and set r=60 (determined by varying parameter and validating using cross-validation). We use two collections, left and right, and aggregate posts by thread. Text of users for which the political affiliation is not known or marked as other are not taken into account. Using our proposed method we can estimate the bias in text. We then apply multivariate statistical analysis to measure the effects of the interactive setting on the bias. In particular we will perform the following analysis:

Quoting behavior analysis

We analyze the effect of quotes on the messages that are written in response to the quote. We will also focus on the question whether text is more influenced by the content of the quote the user is responding to or more by the bias estimation of the quote author.

Thread level analysis

This analysis will focus on the influence of the topic starter on the development of the thread.

4.4.2 Qualitative analysis

Here we present a qualitative analysis as evidence of the face validity of our model, i.e., that its rating of left and right affiliation match what we would intuitively expect. Throughout this section, we highlight, by example, how our model matches this analysis: words which our model classifies as right-leaning will be marked in bold and left-leaning words will be underlined. An example of these words can be found in Table 5. The clearest factor which emerges in analysis of political discourse is the motivating agenda from the party in power. The run-up to the war in Iraq, and the maintained message throughout the time that this data was collected, played the most significant role in the framing of the debate. This rhetoric over time has been analyzed in depth. This body of work has identified two key themes which run through right-leaning political discourse on this topic. The first can be described as "terror" language, and the second can be thought of as "imperialist" language.

The first strategy is described by Kellner [7] as "terror language" and by Simons [17] as "crisis rhetoric." This pattern of description includes messages which evoke emotional responses to the threat of attack. To be effectively persuasive, this language must define the target as evil or fundamentally opposed to those who are hearing the message. It must also argue that the target is a present and dangerous threat. This provokes the listener into a defensive posture, eliciting an emotional response, which means listeners will be more likely to support an extreme ideology.

The second strategy was explored in detail by Cloud [4] in a study of changes to airport security concerns after the terrorist attacks of September 11. Cultural and racial prejudices are expanded upon in this strategy, instilling in the listener an attitude of superiority. This is a long-held tradition exemplified historically by Kipling's *White Man's Burden* [8]. This message honoring the virtue of the listener's culture, instilling a sense of duty to bring that culture to others, is paralleled by the right-leaning rhetoric leading up to the war in Iraq.

These strategies have a common goal of instilling a patriotic sense of nationalism in the listener, though they accomplish this in different ways. We observe both strategies in our data. The impact of "terror language" can be seen through examples:

> **Right person:** *get real and learn the truth and not what your spoonfed confirming that operation iraqi freedom is an integral part of the* ***war*** *on* ***terror, soldiers*** *of the 7th marine regiment destroyed a suspected* ***terrorist*** *camp early sunday en route to baghdad [...] iraq has been listed by the state department for over 13 years through republican and democrat as a* ***terrorist*** *state*

The removed middle portion of this quote is an extensive list of military events in Iraq, most tied to Saddam Hussein and al Qaeda. The clear message is that the problem can be framed as a purely military issue, defining the problem in terms of "sides" which must be engaged and defeated in the interests of safety.

In the imperialist rhetoric, on the other hand, the war in Iraq is seen as a way to bring American ideals to a country that currently is viewed as inferior. This comes through the data we observe by framing the problem in terms of "democracy" and "freedom," such as in the excerpt below:

> **Right person:** *he said that we will stand with those who seek* ***freedom*** *in the world, and we will fight with them for* ***freedom*** *and* ***democracy****, because this is the answer to tyranny and oppression. how is it you can't see that?*

As political rhetoric is largely defined by the body currently in control of government, it is not surprising that the vast majority of research on discourse related to the Iraq war has concentrated on the tone taken by the administrations of then-President Bush and the governments of international allies. This rhetoric largely followed the two patterns described above. In fact, we can see this disparity in the biases our model displays, where the relative bias of words in our left-leaning topic are not nearly as discriminative as those in the right-leaning topic. However, it is also vital to understand the rhetoric of the opposition.

Researchers have engaged the problem of how activists engage in discourse where they must argue against the majority, focusing specifically on opposition to war. The challenge that these activists face is complicated by the strong emotional reaction brought on by terror language and imperialist language. To speak out against this rhetoric runs the risk of sounding weak and unpatriotic, both of which are viewed negatively in American discourse.

In our data we see two primary strategies, justified by prior research in political rhetoric, that are employed by anti-war speakers to avoid this response. The first is to connect speakers to the foreigners that they are discussing, building a "web of concern" which builds upon support for American troops and then relates the support for that group with emotional support for the non-Americans being discussed. This strategy has been detailed extensively by McCoy et al [15]. This allows the anti-war speaker to position themselves not as being unsupportive of American troops, but instead as being concerned with the well-being of all people, including both Americans and Iraqis. The posts make several explicit references to the "terror language" of the right, as terms such as "enemy" or "terrorist" are questioned and complicated. To illustrate this strategy, consider this excerpt from a post by a left-leaning poster:

> **Left person:** *i know this is hard for you to understand....but i have never "defended" the sunni rebels. i have only pointed out that they have, from their perspective, a legitimate reason for fighting us....and a reason that we would do well to acknowledge and adapt to. understanding that your* ***enemy*** *has a legitimate reason to hate you is not...i say again, not synonymous with "sympathizing" with them or "supporting" them. if* ***terrorists*** *and* ***insurgents*** *were identical, general casey would not have had to differentiate between them.*

The second strategy of activists that we explore is that of "harnessing" the dominant or establishment discourse, taking advantage of the strong emotional ties that have been instilled already. This strategy and several related strategies was described by Maney et. al [14] in another study of peace activists. In this case, the same arguments are used as in the dominant discourse, but the polarity is reversed, describing the same actions but in a more negative light.

> **Left person:** *please read the thing in context. every part that says "whereas" is justification for the resolution, not for invasion. not until you get to the part that says: "section 1. short title. this joint resolution may be cited as the 'authorization for use of military force against iraq resolution of 2002" do you even get to anything that justifies invasion. in fact, it clearly outlines what must first be done before invasion is an option. bush clearly ignored the criteria, and just beat the* ***war*** *drum.*

The strategies of the left are particularly interesting because they are by their nature dialogic. The "web of concern" strategy must be prompted by an assumed counter-argument, either because it is replying to a right-leaning post or because this response is expected. The dialogue that we presented at the beginning of this paper is an example of this "web of concern" response to the "imperialist" viewpoint from the right. The content of that post directly confronts the assumptions of the imperialist viewpoint, and the notion of a western-style democracy being a natural next step is complicated by the introduction of cultural factors such as the differences between Islamic sects.

The harnessing language of the left is also dialogic. In the example below, the language of nationalism that usually accompanies terror language from the right is attributed to the actions of terrorists. This forces the right-leaning speaker to hedge his statements and refine his viewpoint.

> **Left person:** *there are native iraqi sunnis who have every right to fight for their country like we would fight for ours. it was those individuals that i clearly spoke of.*
> **Right person:** *i didnt say that all* ***insurgents*** *who kill americans in iraq come from someplace else. i said that there are no "****freedom*** *fighters" which is what you compared those indiginous sunni young men in iraq. i'll simplify this for you. there is no one born in iraq who is fighting and killing americans for their* ***freedom*** *or the* ***freedom*** *of fellow iraqis.*

These strategies are not always followed closely, as they are not conscious efforts. In fact, as no central agreement is sought in most of the examples in our dataset, the text of participants in a discussion is often very polarizing. We see this in the example below:

> **Left person:** *by pointing out the inflation of* ***Saddam****'s body count by neocons in an effort to further vilify him and thus further justify our invasion we are not DEFENDING* ***saddam****....just pointing out how neocons rarely let facts get in the way of a good war.*
> **Right person:** *So wait, how many do you think* ***Saddam*** *killed or oppressed? You're trying to make him look better than he actually was. You're the one inflating the casualties we've caused! Seriously, what estimates (with a link) are there that we've killed over 100,000 civilians. [...]*

The left person uses subjective words representing his view such as *"neocons", "vilify"* and *"invasion"*. The right person responds, but does not repeat these subjective words. He only borrows words such as *"Saddam"* and *"inflating"* (variant of inflation) to keep the conversation flowing. People therefore tend not to borrow a lot of words from the participants with the opposing view, except the words that are functional for the conversation flow.

Despite this, we see that local context has an impact on the terminology and understanding of partisan speakers, and that their opinions may not be changed but they are communicated differently in response to their opposition. We will perform a quantitative analysis in the next section, but even at a surface level, our model appears to convey the intuitive sense that prior work has described when analyzing war rhetoric.

4.4.3 Quantitative analysis

Quoting behavior analysis

One of our goals is to investigate how the bias in text changes when people are quoting each other in the forum, and which factors influence these changes. For this analysis, consider that a responding message contains quoted text from the initiating message plus the responder's own contributed text. For each of these messages, we computed three word vectors, for which we then separately computed the bias score. The first vector contained the quoted text after the words that were repeated in the responder's text were removed (referred to as "words only in the quote"). The second contained the responder's contribution after the words that were included in the quote were removed from it ("words only in the post"). The final vector contained the words that were removed in computing the earlier two vectors ("words in both").

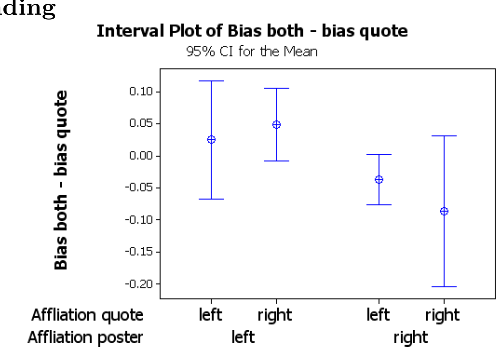
We first find a negative correlation between the bias of the words only in the quote and bias of the words only in the text ($r=-0.1$, $p < 0.05$). This can be explained by the fact that users respond more to others who hold opposing views (as observed by [13] who provided this data). Also the bias of the words only in the post are significantly affected by the poster's affiliation ($F(1,570) = 9.23$, $p < .0001$), but not by the affiliation of the person who is quoted. However, when we compare the bias estimation of the whole quote and that of the whole text, there is a positive correlation ($r=0.104$, $p < 0.02$). Thus the personal bias of a person is most clearly displayed through the words that are unique to his response (thus not occurring in the quote). However, the bias in their response does shift towards that of the text that is quoted.

Using an ANOVA model we investigate which aspect influences the response to a quote the most. We observe that the affiliation and the estimated bias of the user who is quoted does not have a significant effect, while the bias of a quote does have an effect on the text of the user ($r=0.104$, $p < 0.02$) . Thus it seems the content the user is responding to has a larger effect on a text than how the user who is quoted represents himself in general.

We are also interested in analyzing which words are repeated when someone is responding to a quote. We find that when the poster is right, the bias estimation of the words that are repeated are significantly more right ($F(1,469) = 4.2$, $p < 0.05$). They are also more right when the poster of the quoted material is right ($F(1,469) = 6.97$, $p < 0.01$). We calculate the difference between the bias of the words both in the quote and response and the bias of the whole quote. A positive value means that the words that are repeated are the words that are more left in the quote. The plot in Figure 1 displays the mean value and confidence interval of this value depending on the affiliation of the author and that of the author being responded to. We see that when people are responding to the opposing affiliation, they tend to repeat words that are more neutral or in line with their own affiliation. For example, when a left person is responding to a right person, he picks the words that are more left compared to the whole quote. Furthermore we observe more variance when people are responding to the same affiliation. This might indicate that posters are less choosy about what aspects of their interlocutor's message they quote when they see themselves as being more aligned overall.

Thus, it looks like words are repeated to keep the conversation flowing, but the person who is responding tries not to adjust his language too much by repeating words that are less strongly associated with his opponent. This observation is also in line with our previous analysis that showed that the unique words in a response clearly display the bias of the person and not that of the affiliation who is being quoted.

Figure 1: Relation words picked in quote when responding



Thread-level analysis
We are interested in the extent to which a post is influenced by the previous posts in the thread. For a post by user u in thread t on position j in the thread we aggregate the bias estimation values of all posts occurring before the post of interest in the thread:

$$BiasThread(j,t) : \sum_{i=0}^{j-1} bias(post_{i,t}) * \frac{length(post_{i,t})}{\sum_{i=0}^{j-1} length(post_{i,t})}$$

Note that we now take every post into account, even if the affiliation of the poster is not known or annotated as *other*. We furthermore calculate the difference between the bias estimation of the current post and that of the user without taking the current post into account. A positive value means that the current post is more left than the user normally is. We then find a small, but significant correlation between these two values ($r=0.133$, $p<0.01$), indicating that users talk more left than they usually do when the previous posts in that thread are very left, and more right when the previous posts are very right.

We are also interested in the effect of the first post on the further development of the thread. We calculate the correlation between the bias estimation of a thread (without taking the first post into account) and the bias estimation of the first post of a thread. We discard threads where the first post was short (less than 10 tokens) or the number of posts was small (less than 10). We find a correlation between the bias estimation of the whole thread and the estimation of the first post ($r=0.250$, $p<0.01$). However, this turned out to be a byproduct of the fact that more posts by right users are placed when the topic starter is right then when the topic starter is left ($F(2,1824) = 240.7$, $p < .0001$, $R^2=.2$). Once we took the proportion of left and right users into account, the bias estimation of the first post did not had a significant effect on the bias estimate of the thread anymore.

Limitations
Our dataset is very noisy, due to the informal text and sometimes incorrect quoting annotations. Furthermore the number of active users is quite small, and the left people clearly dominate the right people in the amount of text they write. It is therefore hard to generalize the results we have found.

5. CONCLUSIONS

We presented an investigation into the effect of interaction on the representation of perspective biases in text. We found evidence that word distribution changes depending on who the speaker is responding to. When quoting others, users try to keep the conversation flowing but most of the times only repeat words from the quote which are more neutral or in line with their own beliefs when they quote others. They brought their own viewpoint in the response through the words that were unique to the post (thus not occurring in the quote). Our analysis furthermore suggested that the content of a quote exerts more influence on a response than how the quote author represents himself in general. In addition, our analysis on the thread level revealed that people tend to talk more right when the previous posts in the thread are right and similarly for left.

We intend to improve our method to estimate the political affiliation of a text. We are also interested in experimenting with other, larger datasets. We would like to perform our analysis on a dataset where the language is less polarizing and users have more incentives to come to agreement or to understand each other.

6. ACKNOWLEDGMENTS

This paper is based on research funded by NSF grants DRL-0835426 and HCC-0803482.

7. REFERENCES

[1] R. Agrawal, S. Rajagopalan, R. Srikant, and Y. Xu. Mining newsgroups using networks arising from social behavior. In *WWW '03: Proceedings of the 12th international conference on World Wide Web*, pages 529–535, New York, NY, USA, 2003. ACM.

[2] Alias-i. lingpipe-3.9.0, http://alias-i.com/lingpipe.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, 2003.

[4] D. L. Cloud. Flying while arab: The clash of civilizations and the rhetoric of racial profiling in the american empire. *Terrorism: Communication and Rhetorical Perspectives*, 2008.

[5] M. Jiang and S. Argamon. Political leaning categorization by exploring subjectivities in political blogs. In *In Proceedings, 4th International Conference on Data Mining*, pages 647–653, 2008.

[6] F. M. Kawakib Momani, Muhammad A. Badarneh. Intertextual borrowings in ideologically competing discourses: The case of the middle east. *Journal of Intercultural Communication*, (22), 2010.

[7] D. Kellner. Bushspeak and the politics of lying: Presidential rhetoric in the "war on terror.". *Presidential Studies Quarterly*, 2007.

[8] R. Kipling. The white man's burden. McClure's Magazine, 1899.

[9] J. Kristeva. *The Kristeva reader*. Oxford: Blackwell, 1986.

[10] M. Laver, K. Benoit, and T. College. Extracting policy positions from political texts using words as data. *American Political Science Review*, 2003.

[11] W.-H. Lin and A. Hauptmann. Are these documents written from different perspectives?: a test of different perspectives based on statistical distribution divergence. In *ACL-44: Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 1057–1064, Morristown, NJ, USA, 2006. Association for Computational Linguistics.

[12] W.-H. Lin, E. Xing, and A. Hauptmann. A joint topic and perspective model for ideological discourse. In *ECML PKDD '08: Proceedings of the European conference on Machine Learning and Knowledge Discovery in Databases - Part II*, pages 17–32, Berlin, Heidelberg, 2008. Springer-Verlag.

[13] R. Malouf and T. Mullen. Graph-based user classification for informal online political discourse. In *Proceedings of the 1st Workshop on Information Credibility on the Web*, 2007.

[14] G. M. Maney, L. M. Woehrle, and P. G. McCoy. Ideological consistency and contextual adaptation: U.s. peace movement emotional work before and after 9/11. *American Behavioral Scientist*, 2009.

[15] G. M. M. Patrick G. McCoy, Lynne M. Woehrle. Discursive legacies: The u.s. peace movement and "support the troops". *Social Problems*, 2008.

[16] M. Paul and R. Girju. Cross-cultural analysis of blogs and forums with mixed-collection topic models. In *EMNLP '09: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 1408–1417, Morristown, NJ, USA, 2009. Association for Computational Linguistics.

[17] H. W. Simons. From post-9/11 melodrama to quagmire in iraq: A rhetorical history. *Rhetoric and Public Affairs*, 2007.

[18] M. Thomas, B. Pang, and L. Lee. Get out the vote: determining support or opposition from congressional floor-debate transcripts. In *EMNLP '06: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pages 327–335, Morristown, NJ, USA, 2006. Association for Computational Linguistics.

[19] P. Treeratpituk and J. Callan. Automatically labeling hierarchical clusters. In *dg.o '06: Proceedings of the 2006 international conference on Digital government research*, pages 167–176, New York, NY, USA, 2006. ACM.

[20] T. A. van Dijk. Discourse semantics and ideology. *Discourse and society*, 6(2):243–289, 1995.

[21] T. A. van Dijk. Ideology and discourse: a multidisciplinary introduction. 2003.

[22] B. Yu, S. Kaufmann, and D. Diermeier. Exploring the characteristics of opinion expressions for political opinion classification. In *dg.o '08: Proceedings of the 2008 international conference on Digital government research*, pages 82–91. Digital Government Society of North America, 2008.

[23] C. Zhai, A. Velivelli, and B. Yu. A cross-collection mixture model for comparative text mining. In *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 743–748, New York, NY, USA, 2004. ACM.