

# Staying on Topic: An Indicator of Power in Political Debates

**Vinodkumar Prabhakaran**

Dept. of Computer Science  
Columbia University  
New York, NY, USA

vinod@cs.columbia.edu

**Ashima Arora**

Dept. of Computer Science  
Columbia University  
New York, NY, USA

aa3470@columbia.edu

**Owen Rambow**

CCLS  
Columbia University  
New York, NY, USA

rambow@ccls.columbia.edu

## Abstract

We study the topic dynamics of interactions in political debates using the 2012 Republican presidential primary debates as data. We show that the tendency of candidates to shift topics changes over the course of the election campaign, and that it is correlated with their relative power. We also show that our topic shift features help predict candidates' relative rankings.

## 1 Introduction

The field of computational social sciences has created many interesting applications for natural language processing in recent years. One of the areas where NLP techniques have shown great promise is in the analysis of political speech. For example, researchers have applied NLP techniques to political texts for a variety of tasks such as predicting voting patterns (Thomas et al., 2006), identifying markers of persuasion (Guerini et al., 2008), capturing cues that signal charisma (Rosenberg and Hirschberg, 2009), and detecting ideological positions (Sim et al., 2013). Our work also analyzes political speech, more specifically, presidential debates. The contribution of this paper is to show that the topic shifting tendency of a presidential candidate changes over the course of the election campaign, and that it is correlated with his or her relative power. We also show that this insight can help computational systems that predict the candidates' relative rankings based on their interactions in the debates.

## 2 Motivation

The motivation for this paper stems from prior work done by the first author in collaboration with other researchers (Prabhakaran et al., 2013a; Prabhakaran et al., 2013b). Prabhakaran et al.

(2013a) introduced the notion of power in the domain of presidential debates, and Prabhakaran et al. (2013b) followed it up with an automatic power ranker system based on interactions within the debates. The power that a candidate had at a certain point in the election campaign was modeled based on his or her recent poll standings: in elections, popularity is power. Those studies analyzed the 2012 Republican presidential primary debates and found that a candidate's power at the time of a debate correlates with the structure of interactions within the debate (e.g., turn frequency and interruption patterns). Another finding was that the candidates' power correlates with the distribution of topics they speak about in the debates: candidates with more power spoke significantly more about certain topics (e.g., economy) and less about certain other topics (e.g., energy). However, these findings relate to the specific election cycle that was analyzed and will not carry over to political debates in general.

A further dimension with relevance beyond a specific election campaign is how topics evolve during the course of an interaction (e.g., who attempts to shift topics). In (Prabhakaran et al., 2014), we explored this dimension and found that candidates with higher power introduce significantly more topics in the debates, but attempt to shift topics significantly less often while responding to a moderator. We used the basic LDA topic modeling method (with a filter for substantivity of turns) to assign topics to turns, which were then used to detect shifts in topics. However, segmenting interactions into coherent topic segments is an active area of research and a variety of topic modeling approaches have been proposed for that purpose. In this paper, we explore the utility of one such topic modeling approach to tackle this problem.

While most of the early approaches for topic segmenting in interactions have focused on the

content of the contribution, Nguyen et al. (2012) introduced a system called Speaker Identity for Topic Segmentation (SITS) which also takes into account the topic shifting tendencies of the participants of the conversation. In later work, Nguyen et al. (2013) demonstrated the SITS system’s utility in detecting influencers in Crossfire debates and Wikipedia discussions. They also applied the SITS system to the domain of political debates. However they were able to perform only a qualitative analysis of its utility in the debates domain since the debates data did not have influence annotations. In this paper, we use the SITS system to assign topics to turns and perform a quantitative analysis of how the topic shift features calculated using the SITS system relate to the notion of power as captured by (Prabhakaran et al., 2013a).

The SITS system associates each debate participant with a constant scalar value that captures his or her tendency to shift topics. However, since we want to investigate how each candidate’s topic shifting tendency relates to his or her changing power over the course of the campaign, we introduce a variation of the SITS analysis in which we represent a different “persona” for each candidate in each debate. Once equipped with this notion of “persona”, we find that the topic shifting tendency of a candidate does indeed show a great deal of fluctuation during the election campaign period. We also find that this fluctuation in topic shifting tendencies is significantly correlated with the candidates’ power.

As an additional contribution of this paper, we demonstrate the utility of our topic shift features extracted using both types of SITS-based analyses in improving the performance of the automatic power ranker system presented in (Prabhakaran et al., 2013b). We also investigated the utility of topic shifting features described in (Prabhakaran et al., 2014) extracted using LDA based topic modeling. However, they did not improve the performance of the ranker, and hence we do not discuss them in detail in this paper.

### 3 Data

We use the presidential debates corpus released by Prabhakaran et al. (2013a), which contains manual transcripts of 20 debates held between May 2011 and February 2012 as part of the 2012 Republican presidential primaries. The corpus also captures each candidate’s power at the time of each debate,

computed based on their relative standing in recent public polls. The poll numbers capture how successful candidates are in convincing the electorate of their candidature, which in turn affects their confidence within the debates. These debates serve as a rich domain to explore manifestations of power since they are a medium through which candidates pursue and maintain power over other candidates. Prabhakaran et al. (2013b) offers a detailed description of how the relative standings in national and state-level polls from various sources are aggregated to obtain candidates’ power.

The transcripts are originally obtained from The American Presidency Project, where each turn of the conversation is manually demarcated and their speakers identified. The turns in the corpus are preprocessed using the Stanford CoreNLP package to perform basic NLP steps such as tokenization, sentence segmentation, parts-of-speech tagging and lemmatization.

## 4 Modeling Topic Shifts

Topic segmentation, the task of segmenting interactions into coherent topic segments, is an important step in analyzing interactions. In addition to its primary purpose, topic segmentation also identifies the speaker turn where the conversation changed from one topic to another, i.e., where the topic shifted, which may shed light on the characteristics of the speaker who changed the topic. We use the SITS approach proposed by (Nguyen et al., 2012) to detect topic shifts. We also propose a different way of using SITS to obtain an analysis of our corpus, which we call SITS<sup>var</sup>. We discuss both in turn, and then provide a discussion.

### 4.1 Segmentation using SITS

Most computational approaches towards automatic topic segmentation have focused mainly on the content of the contribution without taking into account the social aspects or speaker characteristics. Different discourse participants may have different tendencies to introduce or shift topics in interactions. In order to address this shortcoming, Nguyen et al. (2012) proposed a new topic segmentation model called Speaker Identity for Topic Segmentation (SITS), in which they explicitly model the individual’s tendency to introduce new topics.

Like traditional topic modeling approaches, the SITS system also considers each turn to be a

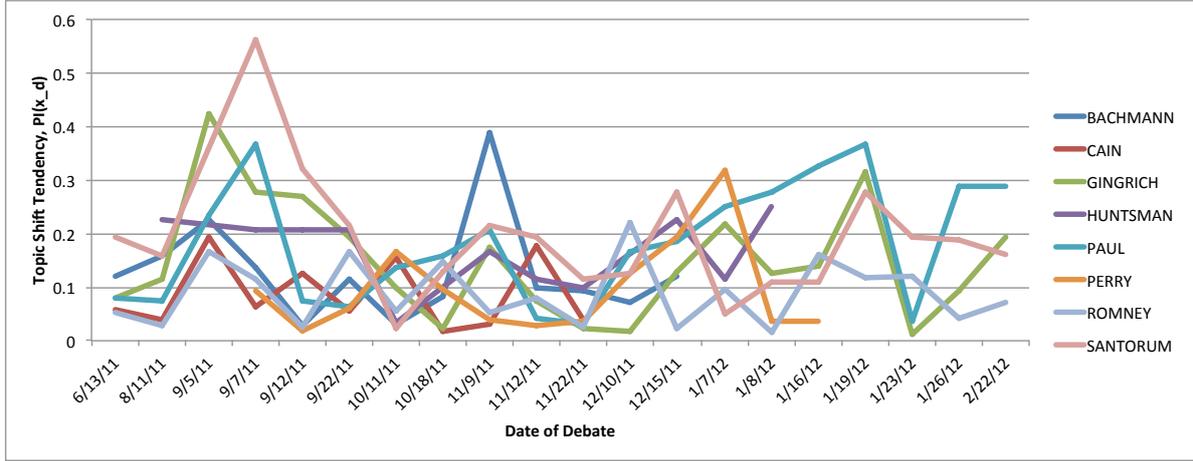


Figure 1: SITS<sup>var</sup> Topic shift tendency values across debates

bag of words generated from a mixture of topics. These topics themselves are multinomial distributions over terms. In order to account for the topic shifts that happen during the course of an interaction, they introduce a binary latent variable  $l_{d;t}$  called the topic shift to indicate whether the speaker changed the topic or not in conversation  $d$  at turn  $t$ . To capture the individual speaker’s topic shifting tendency, they introduced another latent variable called topic shift tendency ( $\pi_x$ ) of speaker  $x$ . The  $\pi_x$  value represents the propensity of speaker  $x$  to perform a topic shift.

## 4.2 Segmentation using SITS<sup>var</sup>

Within the SITS formulation, the topic shifting tendency of an individual ( $\pi_x$ ) is considered a constant across conversations. While an individual may have an inherent propensity to shift topics or not, we argue that the topic shifting tendency he or she displays can vary based on the social settings in which he or she interacts and his or her status within those settings. In other words, the same discourse participant may behave differently in different social situations and at different points in time. This is especially relevant in the context of our dataset, where the debates happen over a period of 10 months, and the power and status of each candidate in the election campaign vary greatly within that time period.

We propose a variant of SITS which takes this issue into account. We consider each candidate to have a different “persona” in each debate. To accomplish this, we create new identities for each candidate  $x$  for each debate  $d$ , denoted by  $x_d$ . For example, ‘ROMNEY\_08-11-2011’ de-

notes the persona of the candidate ROMNEY in the debate held on 08-11-2011. Running the SITS system using this formulation, we obtain different  $\pi_{x_d}$  values for candidate  $x$  for different debates, capturing different topic shift tendencies of  $x$ .

## 4.3 Execution

We perform both the SITS and SITS<sup>var</sup> analyses on the 20 debates in our corpus. We used the non-parametric version of SITS for both runs, since it systemically estimates the number of topics in the data. We set the maximum number of iterations at 5000, sample lag at 100 and initial number of topics at 25. We refer the reader to (Nguyen et al., 2013) for details on these parameters.

For each candidate, we calculate the mean and standard deviation of the topic shift tendency ( $\pi_{x_d}$ ) of his or her personas across all debates he or she participated in. We then average these means and standard deviations, and obtain an average mean of 0.14 and an average standard deviation of 0.09. This shows that the topic shift tendencies of candidates vary by a considerable amount across debates. Figure 1 shows the  $\pi_{x_d}$  value fluctuating across different debates.

## 5 Analysis of Topic Shift Features

Nguyen et al. (2013) used the SITS analysis as a means to model influence in multi party conversations. They propose two features to detect influencers: Total Topic Shifts (TTS) and Weighted Topic Shifts (WTS).  $TTS(x, d)$  captures the expected number of topic shifts the individual  $x$  makes in conversation  $d$ . This expectation is calculated through the empirical average of samples

Feature Set	Feature	Correlation
TopSh	Total Topic Shifts (TTS)	0.12
	Weighted Topic Shifts (WTS)	0.16
TopSh <sup>var</sup>	Total Topic Shifts (TTS <sup>var</sup> )	0.12
	Weighted Topic Shifts (WTS <sup>var</sup> )	0.15
	Topic Shift Tendency (PI <sup>var</sup> )	<b>-0.27</b>

Table 1: Pearson Correlations for Topical Features  
**boldface** denotes statistical significance ( $p < 0.05$ )

from the Gibbs sampler, after a burn-in period. We refer the reader to (Nguyen et al., 2013) for more details on how this value is computed.  $WTS(x, d)$  is the value of  $TTS(x, d)$  weighted by  $1 - \pi_x$ . The intuition here is that a topic shift by a speaker with low topic shift tendency must be weighted higher than that by a speaker with a high topic shift tendency. We use these two features as well, and denote the set of these two features as TopSh.

We also extract the TTS and WTS features using our SITS<sup>var</sup> variation of topic segmentation analysis and denote them as TTS<sup>var</sup> and WTS<sup>var</sup> respectively. In addition, we also use a feature  $PI^{var}(x, d)$  which is the  $\pi_{x,d}$  value obtained by the SITS<sup>var</sup> for candidate  $x$  in debate  $d$ . It captures the topic shifting tendency of candidate  $x$  in debate  $d$ . (We do not include the SITS  $\pi_x$  value in our correlation analysis since it is constant across debates.) We denote the set of these three features obtained from the SITS<sup>var</sup> run as TopSh<sup>var</sup>.

Table 1 shows the Pearson’s product correlation between each topical feature and candidate’s power. We obtain a highly significant ( $p = 0.002$ ) negative correlation between topic shift tendency of a candidate (PI) and his/her power. In other words, the variation in the topic shifting tendencies is significantly correlated with the candidates’ recent poll standings. Candidates who are higher up in the polls tend to stay on topic while the candidates with less power attempt to shift topics more often. This is in line with our previous findings from (Prabhakaran et al., 2014) that candidates with higher power attempt to shift topics less often than others when responding to moderators. It is also in line with the findings by Prabhakaran et al. (2013a) that candidates with higher power tend not to interrupt others. On the other hand, we did not obtain any significant correlation for the features proposed by Nguyen et al. (2013).

## 6 Topic Shift Features in Power Ranker

In this section, we investigate the utility of the SITS and SITS<sup>var</sup> based topic shift features described above in the problem of automatically ranking the participants of debates based on their power. Prabhakaran et al. (2013b) define the problem as follows: given a debate  $d$  with a set of participants  $C_d = \{x_1, x_2, \dots, x_n\}$  and corresponding power indices  $P(x_i)$  for  $1 < i < n$ , find a ranking function  $r : C_d \rightarrow \{1 \dots n\}$  such that for all  $1 < i, j < n$ ,  $r(x_i) > r(x_j) \iff P(x_i) > P(x_j)$ . For our experiments, we use the SVM<sup>rank</sup> based supervised learned power ranker presented in that work to estimate this ranking function.

As we do in (Prabhakaran et al., 2013b), we here report Kendall’s Tau and Normalized Discounted Cumulative Gain values (NDCG and NDCG@3) on 5-fold cross validation (at the debate level). All three metrics are based on the number of rank inversions between original and predicted ranking. While Tau treats all rank inversions equal, NDCG and NDCG@3 penalize the inversions happening in the top of the ranked list more than those happening in the bottom. NDCG@3 focuses only on the top 3 positions in the ranked list.

We use the best performing feature set of (Prabhakaran et al., 2013b) as the baseline (BL), which contains three features: Words Deviation (WD), Question Deviation (QD) and Mention Percentage (MP). WD and QD capture the deviation of percentage of words spoken by the candidate and questions addressed to the candidate from the expected fair share of those measures in the particular debate. The fair share for debate  $d$  is  $1/|C_d|$  — the percentage each candidate would have gotten for each feature if it was equally distributed. This deviation measure is used instead of the raw per-

	Kendall’s Tau	NDCG	NDCG@3
BL	0.55	0.962	0.932
TopSh	0.36	0.907	0.830
TopSh <sup>var</sup>	0.39	0.919	0.847
BL + TopSh	0.59	0.967	0.929
BL + TopSh <sup>var</sup>	<b>0.60</b>	<b>0.970</b>	<b>0.937</b>
BL + TopSh + TopSh <sup>var</sup>	0.59	0.968	0.934

Table 2: Power Ranker results using topic shift features on 5-fold cross validation  
 BL: Baseline system (Prabhakaran et al., 2013b)  
 NDCG: Normalized Discounted Cumulative Gain

centage in order to handle the fact that the percentage values are dependent on the number of participants in a debate, which varied from 9 to 4. MP captures the percentage of mentions of the candidate within a debate.

Table 2 shows the results obtained using the baseline features (BL) as well as combinations of TopSh and TopSh<sup>var</sup> features. The baseline system obtained a Kendall Tau of 0.55, NDCG of 0.962 and NDCG@3 of 0.932. The topic shift features by themselves performed much worse, with TopSh<sup>var</sup> posting marginally better results than TopSh. Combining the topic shift and baseline features increases performance considerably. TopSh<sup>var</sup> obtained better performance than TopSh across the board. BL + TopSh<sup>var</sup> posted the overall best system obtaining a Tau of 0.60, NDCG of 0.970, and NDCG@3 of 0.937. These results demonstrates the utility of topic shift features in the power ranking problem, especially using the SITS<sup>var</sup> formulation. We also experimented with all subsets of TopSh and TopSh<sup>var</sup>; the best results were obtained using all features in each set.

## 7 Related Work

Studies in sociolinguistics (e.g., (Ng et al., 1993; Ng et al., 1995; Reid and Ng, 2000)) have long established that dialog structure in interactions relates to power and influence. Researchers in the NLP community have studied power and influence in various genres of interactions, such as organizational email threads (Bramsen et al., 2011; Gilbert, 2012; Prabhakaran and Rambow, 2013; Prabhakaran and Rambow, 2014), online discussion forums (Danescu-Niculescu-Mizil et al., 2012; Biran et al., 2012) and online chat dialogs (Strzakowski et al., 2012). The correlates analyzed in these studies range from word and phrase patterns,

to derivatives of such patterns such as linguistic coordination, to deeper dialogic features such as argumentation and dialog acts. Our work differs from these studies in that we study the correlates of power in topic dynamics. Furthermore, we analyze spoken interactions.

## 8 Conclusion

In this paper, we studied how topic shift patterns in the 2012 Republican presidential debates correlate with the power of candidates. We proposed an alternate formulation of the SITS topic segmentation system that captures fluctuations in each candidate’s topic shifting tendencies, which we found to be correlated with their power. We also showed that features based on topic shift improve the prediction of the relative rankings of candidates. In future work, we will explore a model that captures individuals’ inherent topic shift propensities, while also capturing their fluctuations due to social factors.

## Acknowledgments

This paper is based upon work supported by the DARPA DEFT Program. The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. We also thank the anonymous reviewers for their constructive feedback.

## References

- Or Biran, Sara Rosenthal, Jacob Andreas, Kathleen McKeown, and Owen Rambow. 2012. Detecting influencers in written online conversations. In *Proceedings of the Second Workshop on Language in Social Media*, pages 37–45, Montréal, Canada, June. Association for Computational Linguistics.
- Philip Bramsen, Martha Escobar-Molano, Ami Patel, and Rafael Alonso. 2011. Extracting social power relationships from natural language. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 773–782, Portland, Oregon, USA, June. Association for Computational Linguistics.
- Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. Echoes of power: language effects and power differences in social interaction. In *Proceedings of the 21st international conference on World Wide Web*, WWW '12, New York, NY, USA. ACM.
- Eric Gilbert. 2012. Phrases that signal workplace hierarchy. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, CSCW '12*, pages 1037–1046, New York, NY, USA. ACM.
- Marco Guerini, Carlo Strapparava, and Oliviero Stock. 2008. Corps: A corpus of tagged political speeches for persuasive communication processing. *Journal of Information Technology & Politics*, 5(1):19–32.
- Sik Hung Ng, Dean Bell, and Mark Brooke. 1993. Gaining turns and achieving high in influence ranking in small conversational groups. *British Journal of Social Psychology*, pages 32, 265–275.
- Sik Hung Ng, Mark Brooke, and Michael Dunne. 1995. Interruption and in influence in discussion groups. *Journal of Language and Social Psychology*, pages 14(4), 369–381.
- Viet-An Nguyen, Jordan Boyd-Graber, and Philip Resnik. 2012. Sits: A hierarchical nonparametric model using speaker identity for topic segmentation in multiparty conversations. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 78–87, Jeju Island, Korea, July. Association for Computational Linguistics.
- Viet-An Nguyen, Jordan Boyd-Graber, Philip Resnik, Deborah A. Cai, Jennifer E. Midberry, and Yuanxin Wang. 2013. Modeling topic control to detect influence in conversations using nonparametric topic models. *Machine Learning*, pages 1–41.
- Vinodkumar Prabhakaran and Owen Rambow. 2013. Written dialog and social power: Manifestations of different types of power in dialog behavior. In *Proceedings of the IJCNLP*, pages 216–224, Nagoya, Japan, October. Asian Federation of Natural Language Processing.
- Vinodkumar Prabhakaran and Owen Rambow. 2014. Predicting power relations between participants in written dialog from a single thread. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 339–344, Baltimore, Maryland, June. Association for Computational Linguistics.
- Vinodkumar Prabhakaran, Ajita John, and Dorée D. Seligmann. 2013a. Power dynamics in spoken interactions: a case study on 2012 republican primary debates. In *Proceedings of the 22nd international conference on World Wide Web companion*, pages 99–100. International World Wide Web Conferences Steering Committee.
- Vinodkumar Prabhakaran, Ajita John, and Dorée D. Seligmann. 2013b. Who had the upper hand? ranking participants of interactions based on their relative power. In *Proceedings of the IJCNLP*, pages 365–373, Nagoya, Japan, October. Asian Federation of Natural Language Processing.
- Vinodkumar Prabhakaran, Ashima Arora, and Owen Rambow. 2014. Power of confidence: How poll scores impact topic dynamics in political debates. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, page 49, Baltimore, MD, USA, June. Association for Computational Linguistics.
- Scott A. Reid and Sik Hung Ng. 2000. Conversation as a resource for in influence: evidence for prototypical arguments and social identification processes. *European Journal of Social Psych.*, pages 30, 83–100.
- Andrew Rosenberg and Julia Hirschberg. 2009. Charisma perception from text and speech. *Speech Communication*, 51(7):640–655.
- Yanchuan Sim, Brice D. L. Acree, Justin H. Gross, and Noah A. Smith. 2013. Measuring ideological proportions in political speeches. In *Proceedings of the 2013 Conference on EMNLP*, pages 91–101, Seattle, Washington, USA, October. Association for Computational Linguistics.
- Tomek Strzalkowski, Samira Shaikh, Ting Liu, George Aaron Broadwell, Jenny Stromer-Galley, Sarah Taylor, Umit Boz, Veena Ravishankar, and Xiaoi Ren. 2012. Modeling leadership and influence in multi-party online discourse. In *Proceedings of COLING*, pages 2535–2552, Mumbai, India, December. The COLING 2012 Organizing Committee.
- Matt Thomas, Bo Pang, and Lillian Lee. 2006. Get out the vote: Determining support or opposition from congressional floor-debate transcripts. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pages 327–335, Sydney, Australia, July. Association for Computational Linguistics.