

Bias in Hard News Articles from Fox News and MSNBC: An Empirical Assessment Using the Gramulator

Brock Terwilleger, Philip M. McCarthy, and Travis Lamkin

Department of English, University of Memphis
Memphis, TN, USA
{btrwllgr, pmmccrth, talamkin}@memphis.edu

Abstract

Hard news articles, just like op-ed articles, can reflect a media organization’s bias. This study assesses bias in the hard news articles published by Fox News and MSNBC. Indicative linguistic features identified by the Gramulator reveal biases in corpora from the two networks.

Introduction

This study examines bias in articles found on the websites of cable news networks. More specifically, this study focuses on the linguistic features that distinguish the coverage of news topics on Fox News and MSNBC. The research question is, “Do the different linguistic features in online hard news articles reflect common views of Fox and MSNBC’s biases?” If so, where and how do the corpora refer to subjects associated with the networks’ ideologies? Our hypothesis is that the corpora will reflect conservative and liberal biases respectively, because Fox News and MSNBC take these positions in their televised content.

Previous research has identified bias in the televised content from Fox News and MSNBC, but their electronic content remains relatively unstudied. Studies on evening newscasts include the work of Fico et al. (2008), who found election coverage unbalanced. Studies on political talk shows include the work of Holtzman et al. (in press), who identified positive associations with Republicans and Democrats on Fox and MSNBC respectively, using the Contrast Analysis of Semantic Similarity (CASS) tool. However, few researchers have studied hard news articles online. Weatherly et al. (2007) found that participants perceived CNN’s headlines as more liberal than Fox’s, but the content from Fox and MSNBC is largely unexamined.

The goal of this study is to identify the different lexical choices and frequently mentioned subjects that manifest a

network’s bias in its hard news online. Finding bias in electronic hard news will provide another angle from which to assess the biases of these networks, strengthening the case for institutional bias at Fox News and MSNBC.

The Corpus

The corpus includes 2,529 texts collected from the websites of Fox News and MSNBC. The two main corpora are sub-divided into five sub-corpora (see Table 1).

Table 1: Corpus Composition

	Fox	MSNBC	Beginning date
BP oil spill	547	131	4/1/2010
Economy	394	166	11/1/2009
Elections	261	65	2/1/2010
Health care	251	182	12/1/2010
Immigration	435	97	8/1/2008
<i>Total</i>	1,888	641	

The *sister sub-corpora* (Fox and MSNBC) contain an unequal number of texts because they cover set timeframes. This difference will reflect disparities in coverage influenced by the networks’ ideologies. Thus, the corpora will reveal what each network deemed newsworthy during a given news cycle. The timeframes vary because the Fox News archives limit access. The date of the oldest article available determined the beginning date for both sister corpora. The end date for all timeframes is June 30, 2010.

The Tool

This study uses the Applied Natural Language Processing (ANLP) tool the Gramulator (McCarthy, Watanabe, and Lamkin in press). This tool identifies the indicative linguistic features (e.g. lexical choices, syntactic structures, etc.) that differentiate corpora. Therefore, it is more appropriate for contrastive corpus analysis than other

ANLP tools such as LIWC (Pennebaker et al. 2007) and CASS (Holtzman et al. in press) that draw semantic connections among words within texts in a single corpus. Also, the tool lets users see linguistic features in context.

The Gramulator identifies *differentials*, or collocations that appear significantly more in one corpus than another. First, the tool finds the characteristic collocations of a corpus, or *typicals*. Then, it eliminates collocations that are common between the two corpora. The remaining collocations, the differentials, are indicative of that corpus relative to its sister corpus. Thus, the Gramulator highlights the linguistic differences between contrastive corpora.

Results

Themes among the differentials suggest conservative and liberal bias in the hard news articles from Fox News and MSNBC respectively. Fisher's exact tests indicated the significance of the differentials.

Pan-corpora Differentials. Two differentials appeared across the Fox sub-corpora. *President Obama* ($p < .001$) was one of the three most frequent differentials in all the sub-corpora. And, *fox news* reaches between $p = .004$ and $p < .001$ in every sub-corpora except for health care. No differentials reoccurred across the MSNBC sub-corpora.

BP Oil Spill. Given the uneven size of the Fox sub-corpora, it was randomly divided into two sets. Differentials were taken from a training set of 131 texts, equivalent to the number in the MSNBC sub-corpus. A within *t*-test suggested that the Fox differentials were more present in the Fox test set than the MSNBC differentials [$t(1,415) = 5.375, p < 0.001, d = 0.264$]. Therefore they are likely indicative of the whole sub-corpus.

The results suggest that Fox displayed a pro-business slant, as opposed to MSNBC's pro-environment slant. On Fox, *the moratorium* often coincided with discussions of the economic impact of the government's response. In contrast, MSNBC emphasized the spill's environmental impact through references to oil, e.g., *oil slick* ($p = .005$) and *oil on* ($p < .001$). Also, MSNBC chose a more connotative reference for the rig, *the gusher* ($p = .034$) than Fox's *undersea well* ($p < .006$).

Economy. Fox appears to have focused on policy and MSNBC on positive indicators in the market. Several Fox differentials refer to stimulus and reform measures, notably *the bill* ($p < .001$) and *the legislation* ($p < .001$). Concerns about the national debt appeared alongside *the budget* ($p = .001$). In contrast, MSNBC often mentioned that an *index rose* ($p < .001$), suggesting a *recovery* ($p < .001$).

Elections. The findings suggest that Fox News depicted races as a referendum on the administration, whereas MSNBC explained Democrats' reelection strategies. Fox emphasized the president's ownership of contentious policies, e.g., *President Obama's* ($p = .009$). MSNBC discussed *the incumbent* ($p = .005$) and *the GOP's* ($p < .01$) candidates. MSNBC also outlined political

calculations, shown by *the district* ($p = .001$), *seat and* ($p = .001$), *the backing* ($p < .015$), and *money and* ($p = .003$).

Health Care. The results may indicate that Fox emphasized the legislation of health care policy and MSNBC the health care system. Fox frequently mentioned that Democrats needed a *simple majority* ($p = .014$) to *pass health* ($p = .014$) legislation. MSNBC discussed the *nation's health care* ($p = .001$) in terms of *medical care* ($p < .001$) and *insurance coverage* ($p = .005$), noting the lack of a *government-run insurance* ($p < .001$) plan in the bill.

Immigration. Fox seemed to stress border security, where as MSNBC seemed to emphasize the immigrant experience. Fox frequently discussed the use of *customs and* ($p = .017$) other law *enforcement to* ($p = .012$) secure the border. In contrast, MSNBC covered the debate from a more personal angle, evidenced by *her husband* ($p < .001$), *their home* ($p = .047$), and *family members* ($p = .005$).

Conclusion

The themes present among the Fox News and MSNBC corpora seem to reflect the networks' biases. Fox News' differentials displayed conservative concerns over business interests, the national debt, President Obama's policies, reconciliation, and border security. In contrast, MSNBC's differentials reflected liberal concerns over the environment, the success of stimulus spending, Democratic control of Congress, a national health insurance program, and the perspective of immigrants. The presence of bias in hard news articles from Fox News and MSNBC further suggests institutional bias at these networks.

References

- Fico, F.; Zeldes, G. A.; Carpenter, S.; and Diddi, A. 2008. Broadcast and Cable Network News Coverage of the 2004 Presidential Election: An Assessment of Partisan and Structural Imbalance. *Mass Communication & Society*, 11(3): 319-339.
- Holtzman, N. S.; Schott, J. P.; Jones, M. N.; Balota, D. A.; and Yarkoni, T. in press. Exploring Media Bias with Semantic Analysis Tools: Validation of the Contrast Analysis of Semantic Similarity (CASS). *Behavior Research Methods*.
- McCarthy, P.M.; Watanabe, S.; and Lamkin, T. in press. The Gramulator: A Tool for the Identification of Indicative Linguistic Features. In P. M. McCarthy and C. Boonthum eds., *Applied Natural Language Processing and Content Analysis: Identification, Investigation, and Resolution*. Hershey, PA: IGI Global.
- Pennebaker, J. W.; Booth, R. J.; and Francis, M. E. 2007. *Linguistic Inquiry and Word Count (LIWC2007): A Text Analysis Program*. Austin, TX: LIWC.net.
- Weatherly, J. N.; Petros, T. V.; Christopherson, K. M.; and Haugen, E. N. 2007. Perceptions of Political Bias in the Headlines of Two Major News Organizations. *Harvard International Journal of Press/Politics*. 12(2): 91-104.