

gov_51_final_project

Owen Bernstein and Lindsey Greenhill

11/15/2020

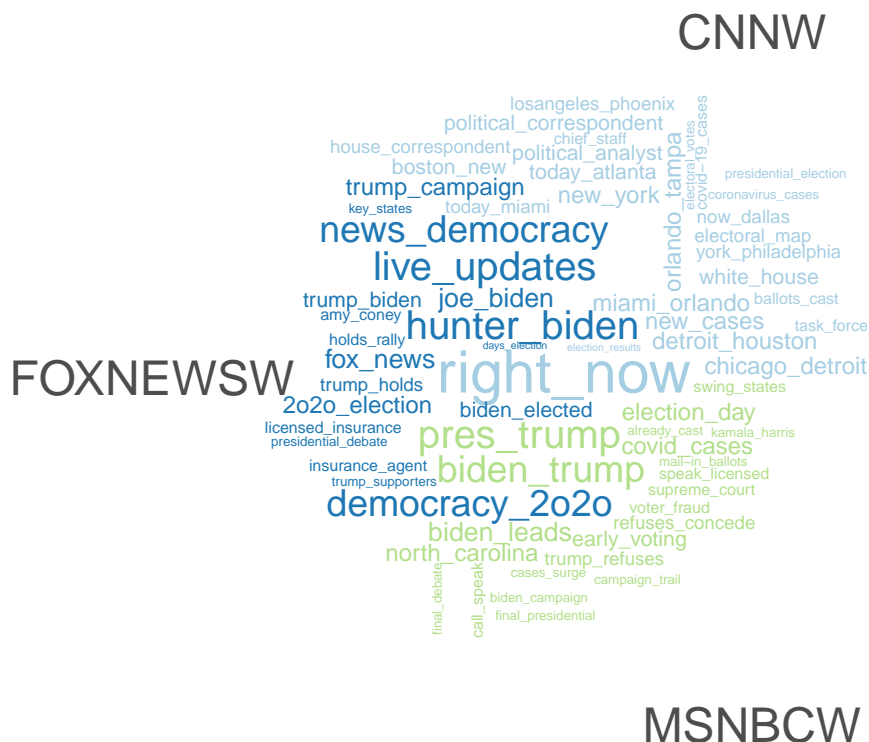
Preliminary Analysis

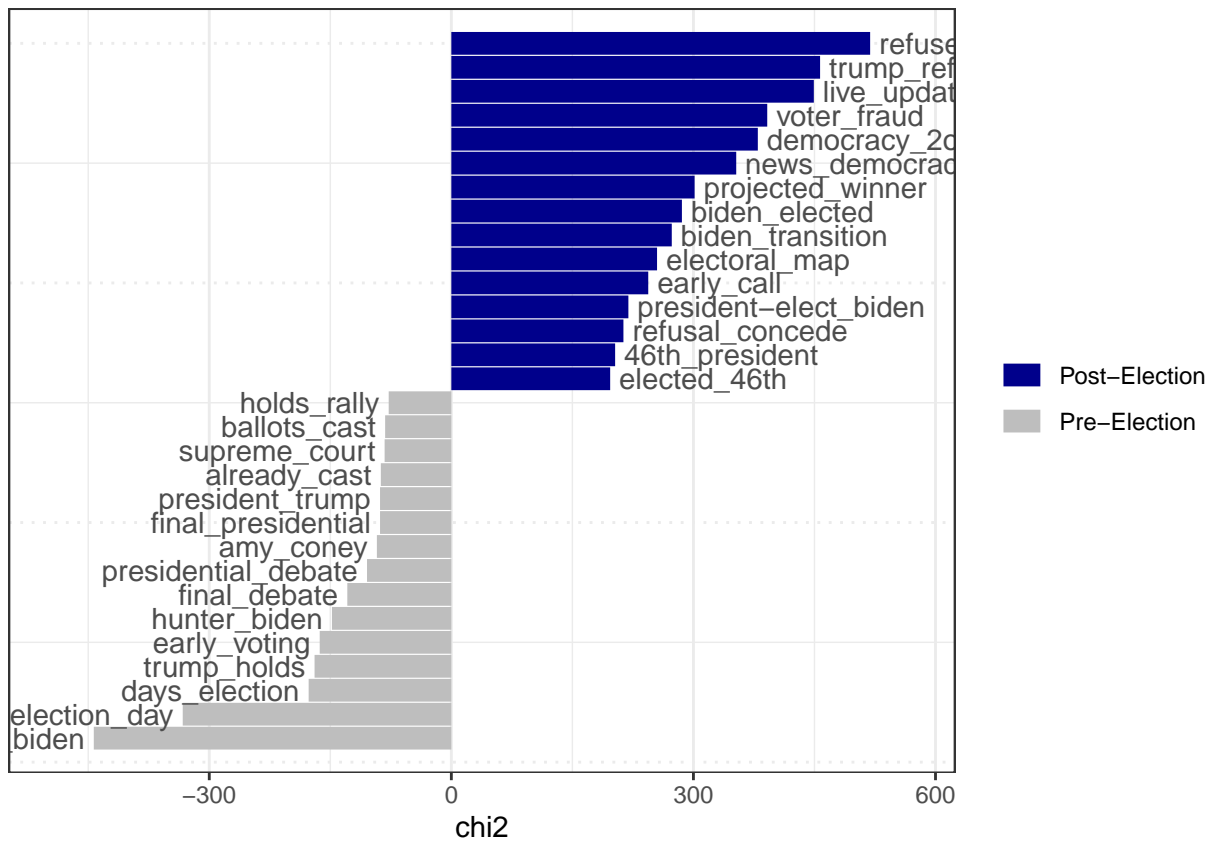
In this analysis, we gathered, cleaned, and wrangled news data from October 17 to November 14, 2020. We chose to include Fox News, CNN, and MSNB in our analysis and assigned those channels ideological rankings of 1, 2, and 3, respectively. The plots below show our preliminary findings.

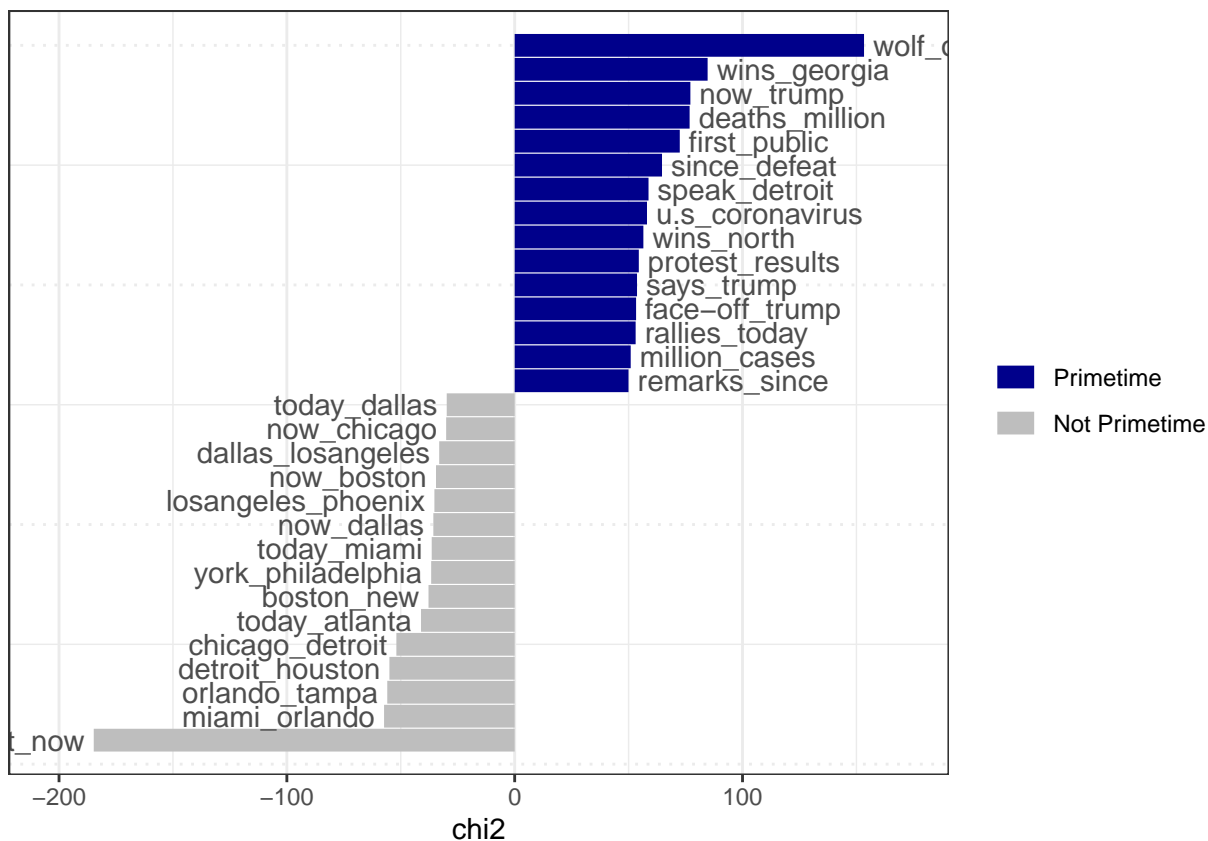
How much specific language is each channel using?

Visualizing the language each channel uses by a variety of variables.

```
## Warning in wordcloud_comparison(x, min_size, max_size, min_count, max_words, :  
## president_trump could not be fit on page. It will not be plotted.
```







We classified the below categories using word baskets built by the study referenced in our project proposal. The table below shows the raw counts of how many times each channel used a word associated with these categories.

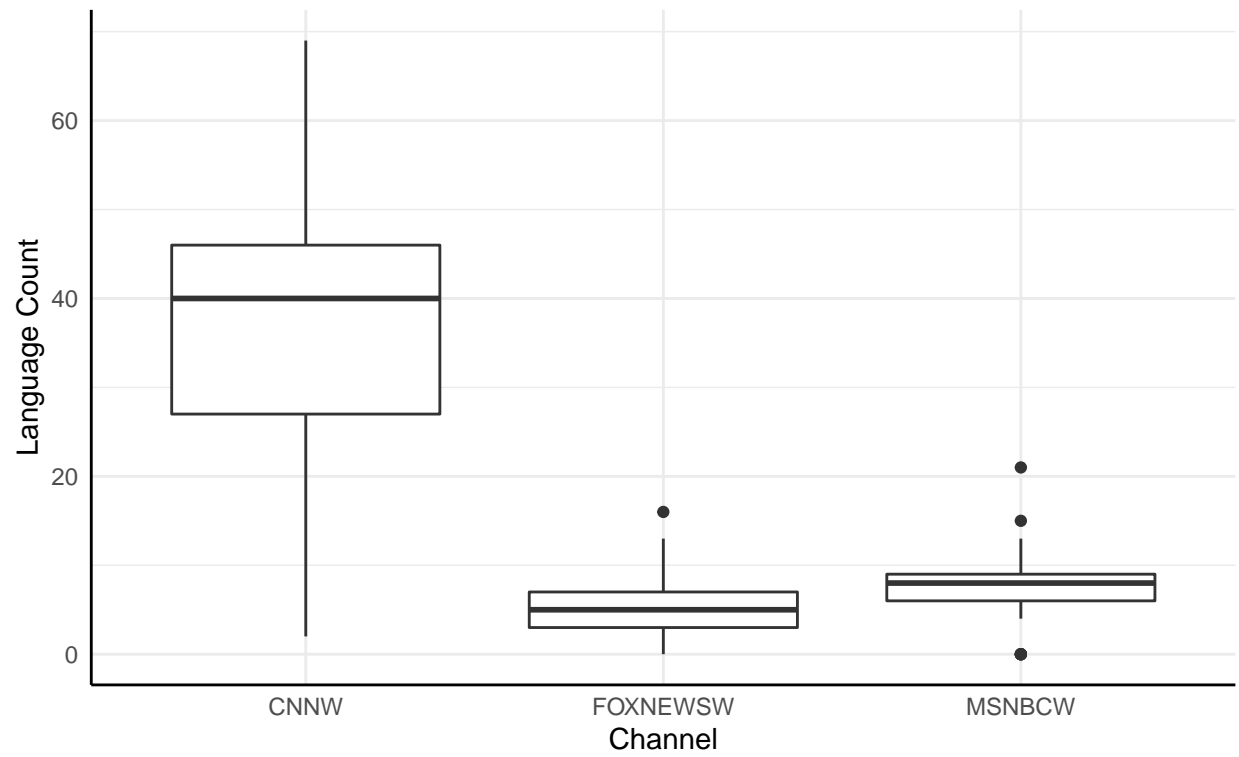
Language Sentiment Across Channels					
October 17 to November 14					
Channel	Word Count for Each Content Category				
	Populism	Environment	Immigration	Progressive	Conservative
CNNW	1022	29	98	940	35
FOXNEWSW	163	15	74	107	10
MSNBCW	219	24	77	80	35

How much specific language is each channel using each day?

The histograms below visualize how much channels use the categorical language on a daily basis.

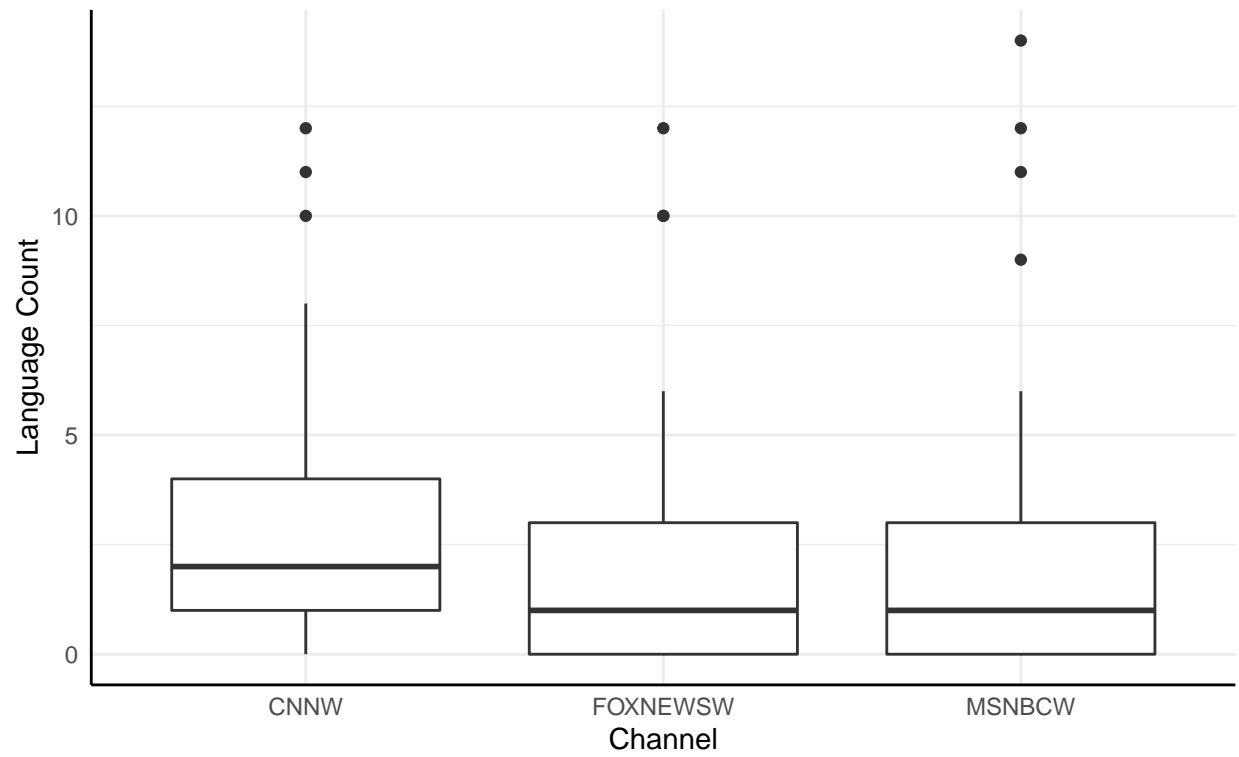
Populist Language in Chyrons by Channel

October 17 to November 14

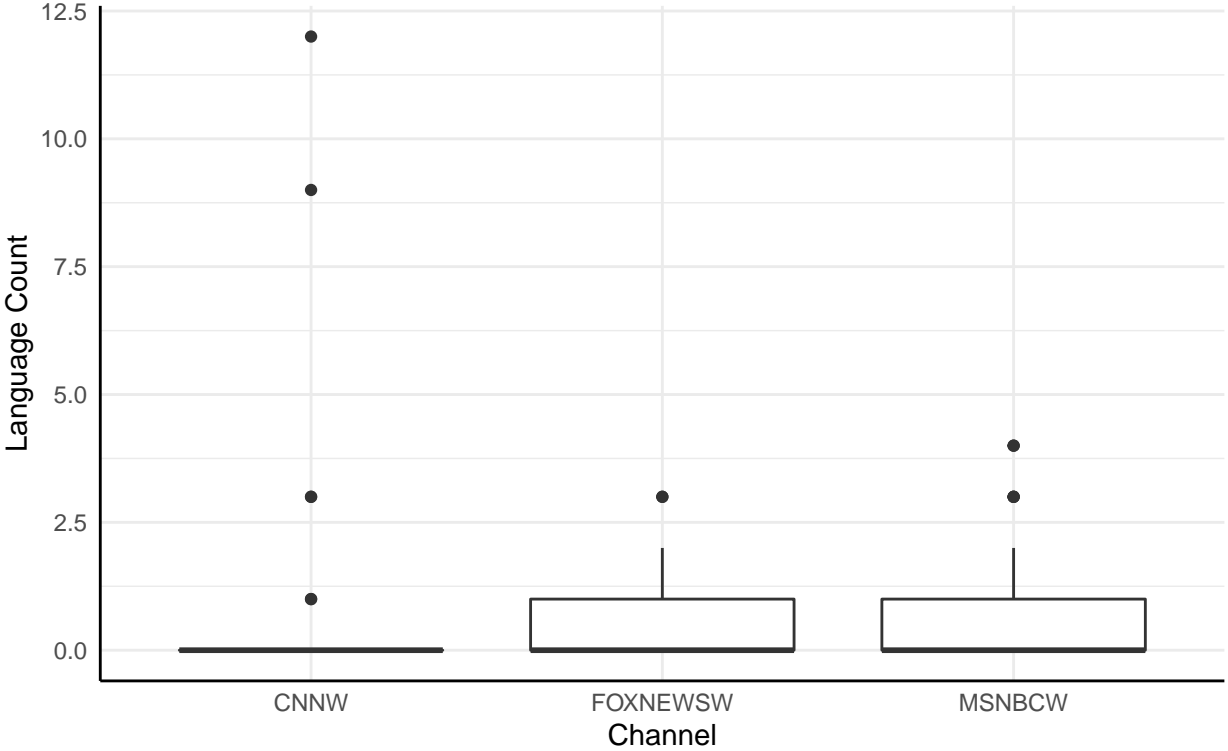


Immigration Language in Chyrons by Channel

October 17 to November 14

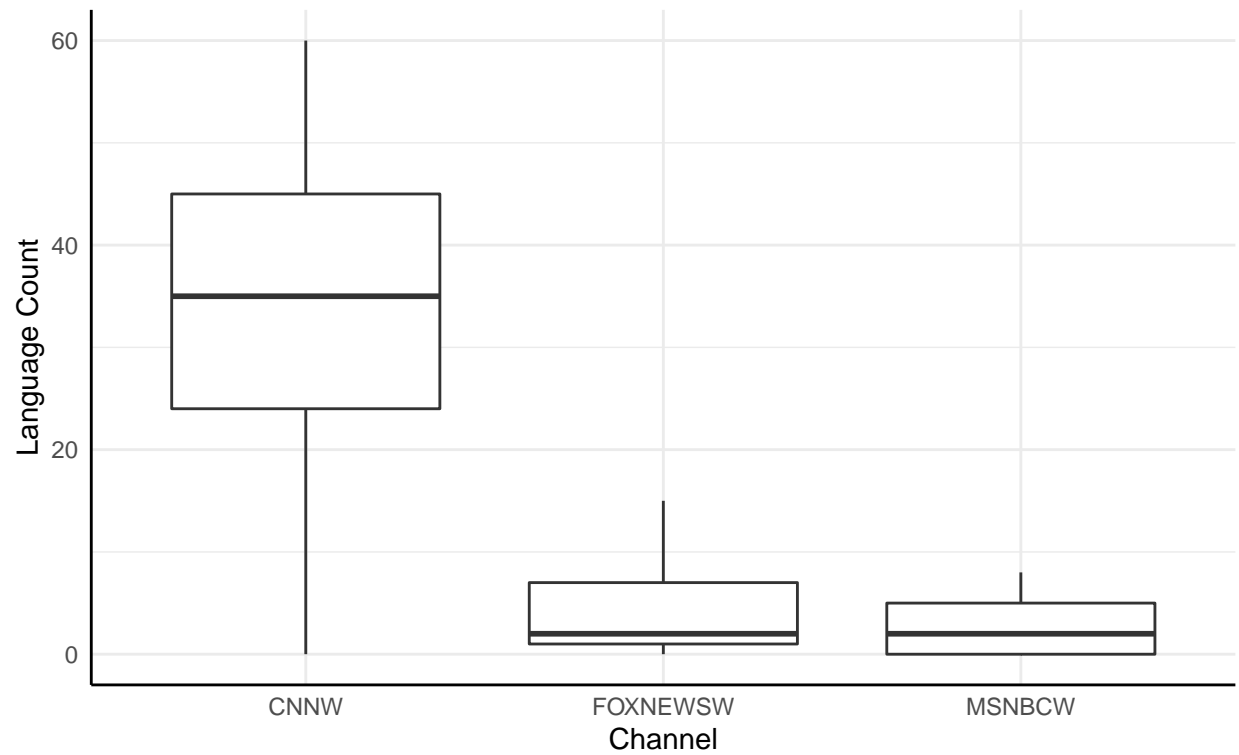


Environmental Language in Chyrons by Channel
October 17 to November 14



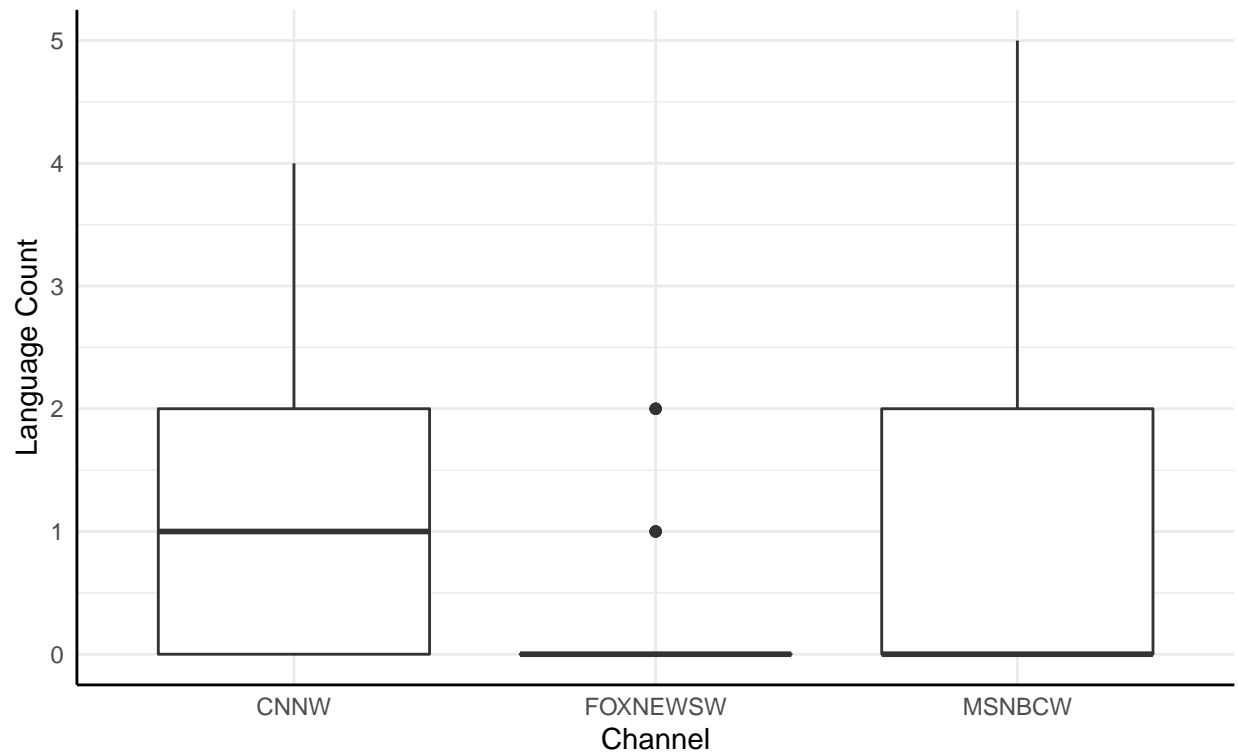
Progressive Language in Chyrons by Channel

October 17 to November 14



Conservative Language in Chyrons by Channel

October 17 to November 14



How does usage vary across time?

The plots below visualize the change in daily language use for each channel.

Regression: Does Ideology have an affect on language usage?

The table below shows the results of five different regressions and the plot below shows the line of best fit for each regression.

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu

% Date and time: Sat, Dec 12, 2020 - 7:21:34 PM

There does not appear to be a strong relationship between ideology and language usage.

	<i>Dependent variable:</i>				
	populism	immigration	environment	progressive	conservative
	(1)	(2)	(3)	(4)	(5)
as.factor(channel)FOXNEWSW	−14.810*** (1.300)	−0.414 (0.438)	−0.241 (0.242)	−14.362*** (1.779)	−0.431*** (0.155)
as.factor(channel)MSNBCW	−13.835*** (1.305)	−0.350 (0.440)	−0.084 (0.243)	−14.888*** (1.787)	0.006 (0.156)
as.factor(election)Pre-Election	3.998*** (1.096)	1.018*** (0.369)	0.070 (0.204)	2.469 (1.499)	0.085 (0.131)
as.factor(primetime)Primetime	−9.442*** (1.064)	−2.096*** (0.359)	−0.596*** (0.198)	−11.477*** (1.457)	−0.640*** (0.127)
Constant	19.860*** (1.262)	2.106*** (0.425)	0.755*** (0.235)	20.413*** (1.728)	0.871*** (0.151)
Observations	173	173	173	173	173
R ²	0.602	0.204	0.057	0.479	0.177
Adjusted R ²	0.592	0.185	0.035	0.466	0.158
Residual Std. Error (df = 168)	6.999	2.358	1.305	9.580	0.837
F Statistic (df = 4; 168)	63.461***	10.734***	2.546**	38.567***	9.057***

Note:

*p<0.1; **p<0.05; ***p<0.01