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?

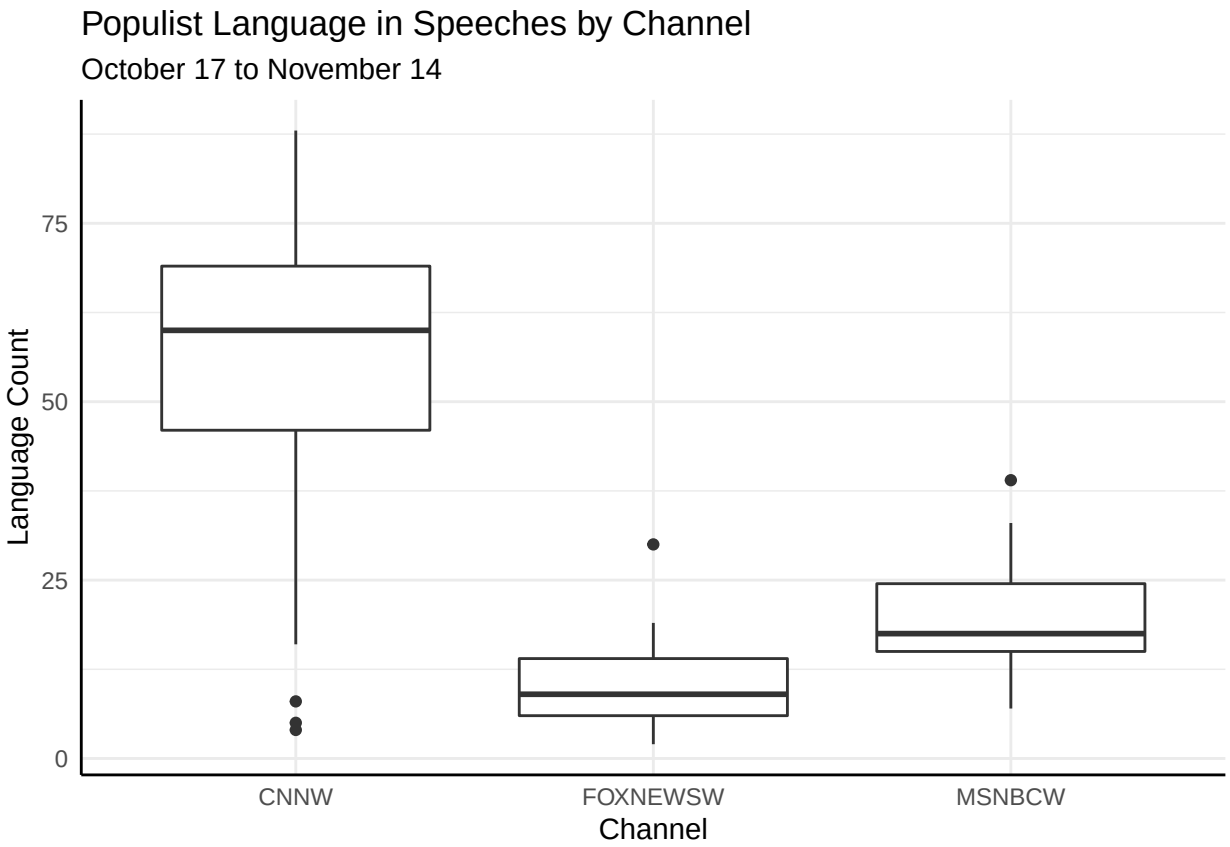
We classified the below categories using word baskets built by the study referenced in our project proposal. The table below shows the raw counts of how many times each channel used a word associated with these categories.

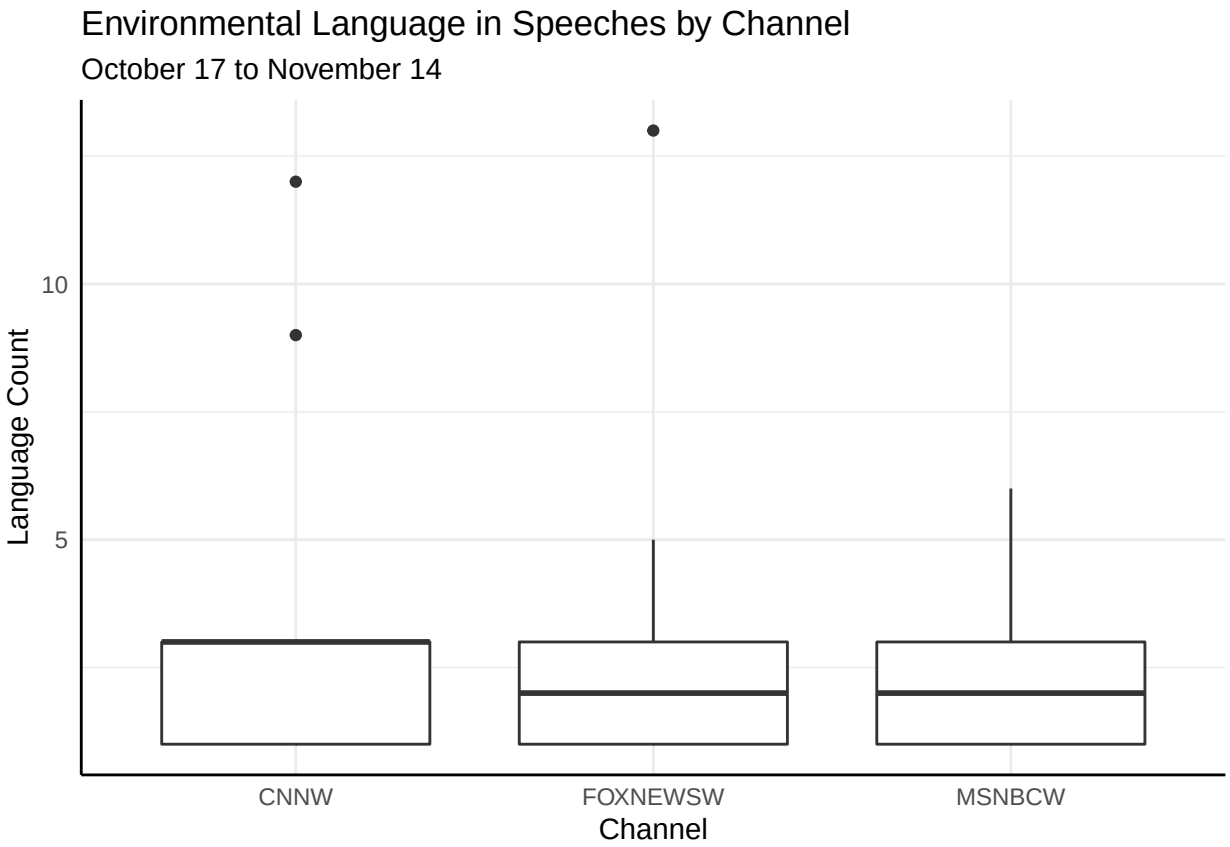
Language Sentiment Across Channels
October 17 to November 14

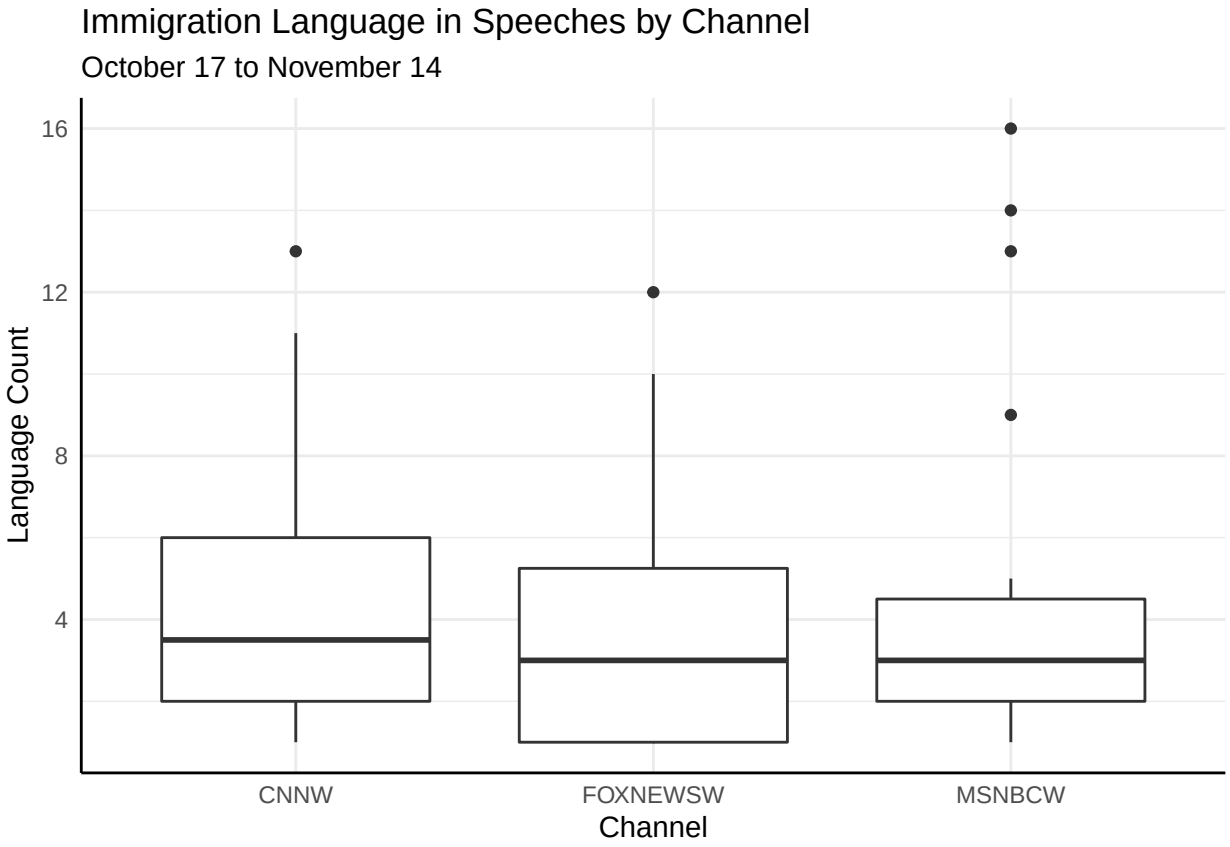
channel	total_populism	total_environment	total_immigration	total_proggresivsim	total_conservatism
CNNW	1576	34	100	209	219
FOXNEWSW	298	39	80	185	76
MSNBCW	470	38	89	117	160

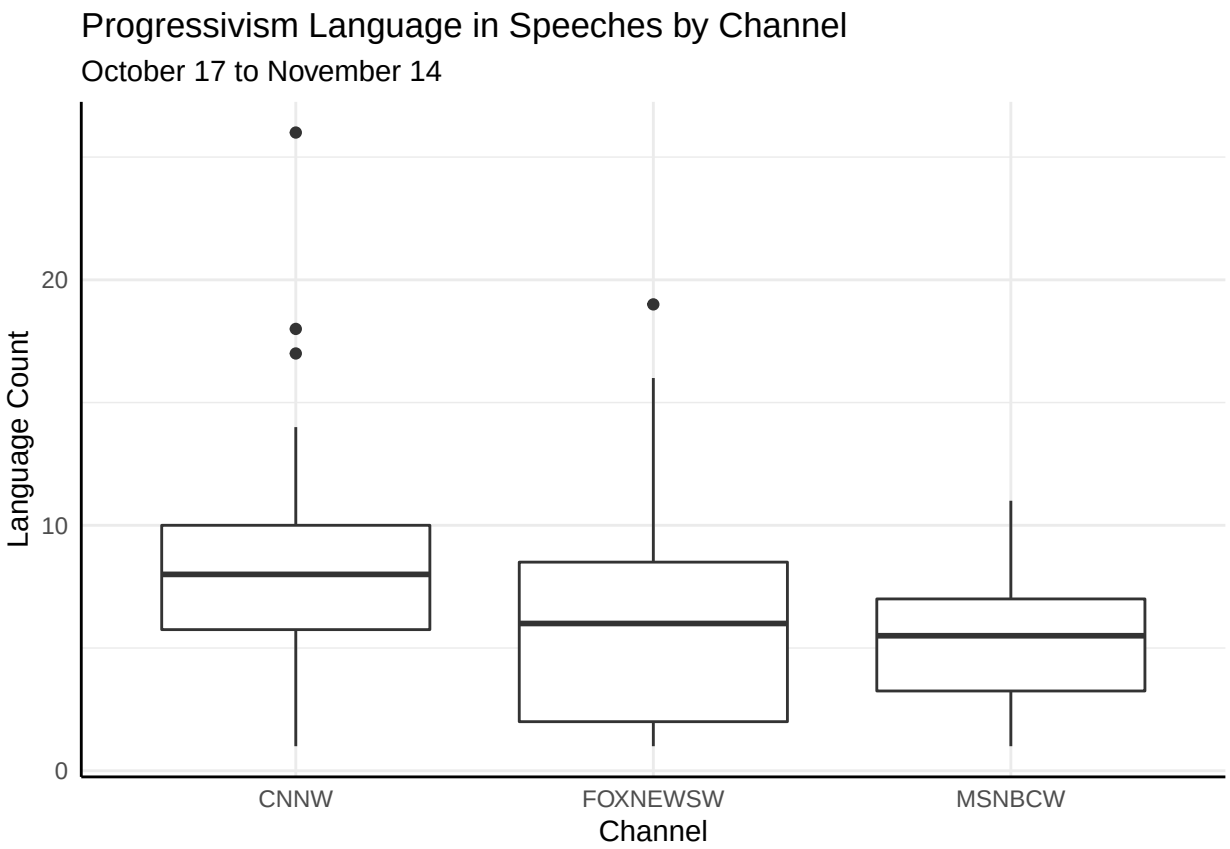
How much specific language is each channel using each day?

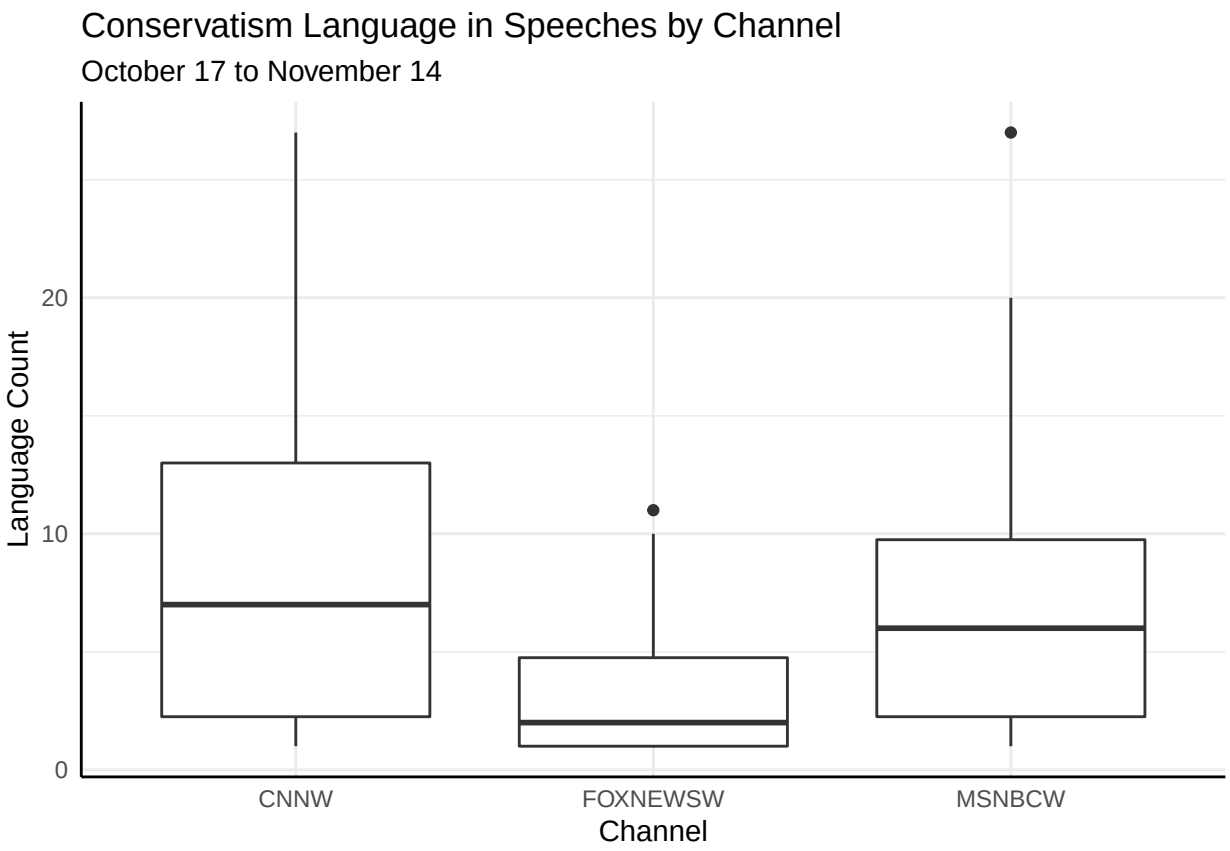
The histograms below visualize how much channels use the categorical language on a daily basis.


Populist Language in Speeches by Channel
October 17 to November 14
Language Count
75
50
25
0
CNNW
FOXNEWSW
MSNBCW
Channel


Environmental Language in Speeches by Channel
October 17 to November 14
Language Count
10
5
CNNW
FOXNEWSW
MSNBCW
Channel

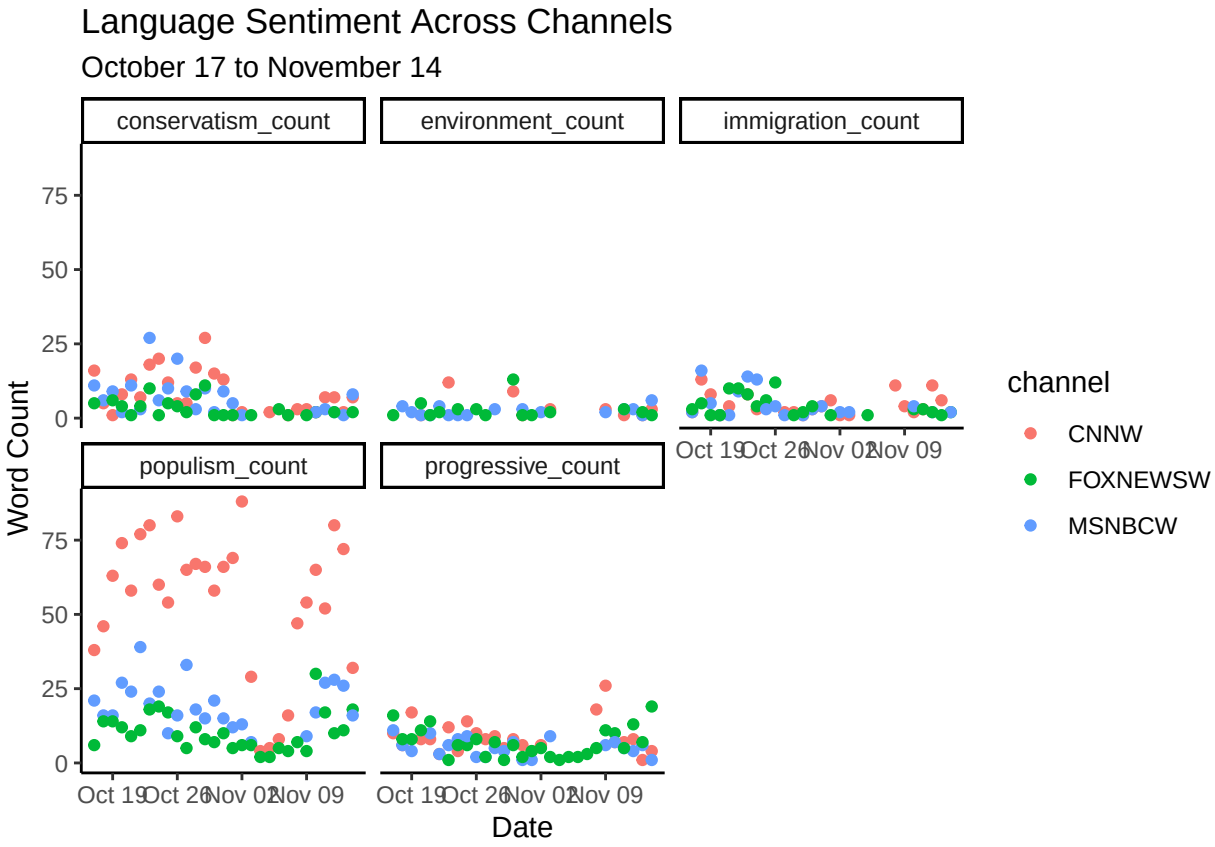

Immigration Language in Speeches by Channel
October 17 to November 14
Language Count
16
12
8
4
CNNW
FOXNEWSW
MSNBCW
Channel


Progressivism Language in Speeches by Channel
October 17 to November 14
Language Count
20
10
0
CNNW
FOXNEWSW
MSNBCW
Channel



How does usage vary across time?

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



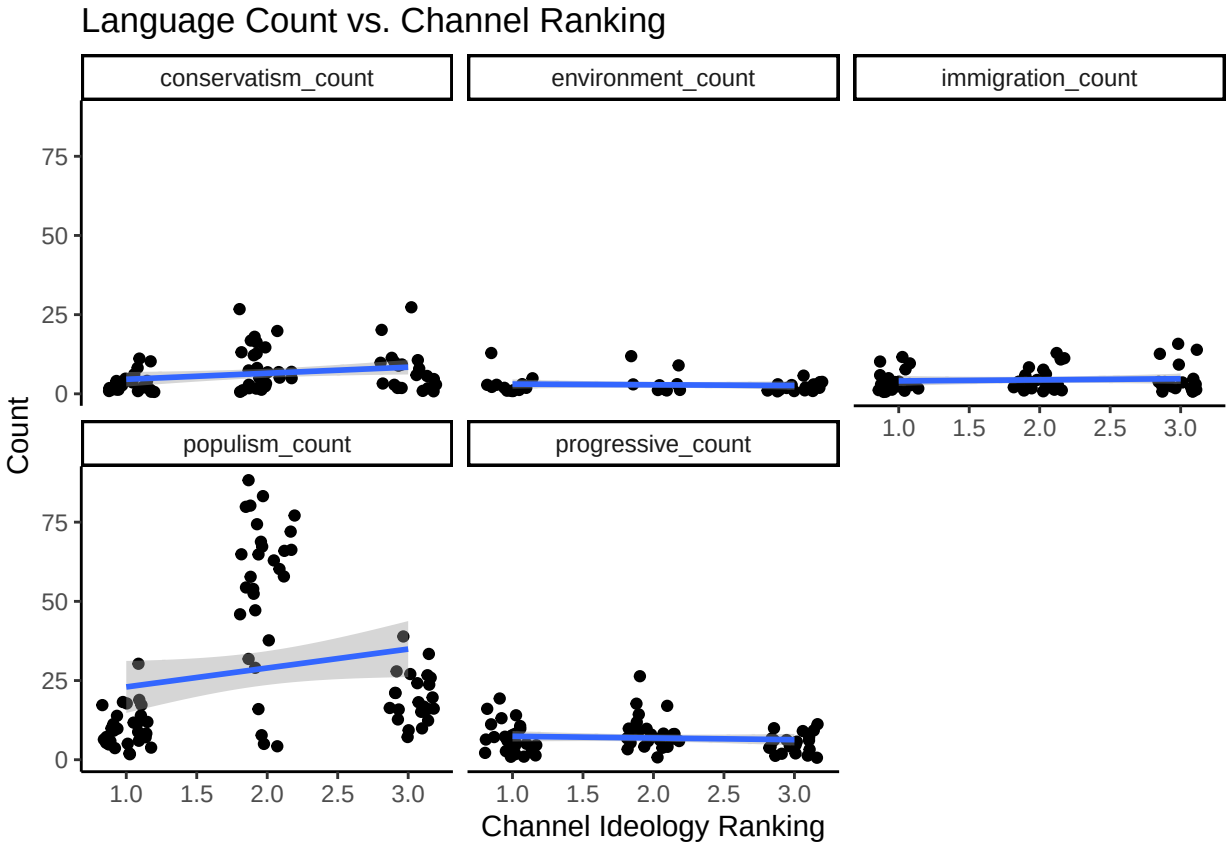
Regression: Does Ideology have an affect on language usage?

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: Tue, Nov 17, 2020 - 20:39:14

```
## `geom_smooth()` using formula 'y ~ x'
```

	Dependent variable:				
	populism_count	immigration_count	environment_count	progressive_count	conserva
	(1)	(2)	(3)	(4)	
ranking	5.976*	0.344	−0.224	−0.537	1.9
	(3.343)	(0.623)	(0.520)	(0.676)	(0
Constant	16.997**	3.728***	3.305***	7.935***	2
	(7.015)	(1.333)	(1.161)	(1.410)	(1
Observations	82	61	39	74	
R^2	0.038	0.005	0.005	0.009	0
Adjusted R^2	0.026	−0.012	−0.022	−0.005	0
Residual Std. Error	24.270 (df = 80)	3.893 (df = 59)	2.845 (df = 37)	4.755 (df = 72)	5.979 (
F Statistic	3.195* (df = 1; 80)	0.304 (df = 1; 59)	0.185 (df = 1; 37)	0.630 (df = 1; 72)	4.485** (
Note:					*p<0.1; **p<0.05



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