Canadian politicians in their own words: A corpus analysis of affective language use in the Canadian House of Commons

Jordan Gallant – Brock University

Sarah McDonald – Brock University

Gary Libben – Brock University

Politicians must choose their words. This is particularly so in settings such as the House of Commons. When Canadian Members of Parliament speak in the House of Commons, they must use words to perform political acts. We report on a study of how some simple tools of psycholinguistics and corpus linguistics can be used to characterize those words and to link them to the 'political culture' within which they are produced and the political dynamics of the day.

To investigate these questions, we constructed a corpus of all words used by all represented parties (BQ, CPC, GP, Lib, and NDP) in the Canadian House of Commons from 2006 to the present. This was done using the official records of the House debates, known as the Hansard. The Hansard are lightly edited, "substantially verbatim" (Caruso, Frohman, Kinsman & Sutherland, 2015) transcriptions made freely available online in pdf, txt, and xml (Extensible markup language) formats. Using RStudio (RStudio Team, 2020) and the XML package (v 3.99-0.5, Temple Lang, 2020), we parsed 1,640 Hansard documents into a corpus of individual tokens (n=50,388,730) tagged by date, language, speaking turn, member, and political party. Only tokens produced by represented parties were included (n=48,234,204). We then linked those words to lexical ratings provided by the Warriner et al., (2013) database. This key step provided us with ratings for each word regarding whether it is positive or negative (valence), emotionally arousing or calming (arousal), and whether it makes you feel helplessness or in-control (dominance). In total, we obtained ratings for 72% of our tokens (n=34,743,463).

Rather than analyze individual tokens, we chose an individual member's speaking turn as our fundamental level of analysis. To this end, we calculated the mean valence, arousal, and dominance ratings for each speaker turn (n=369,556). Turns were tagged with information regarding the political dynamics at the time of the speech. We included as variables: 'party in power', 'minority or majority government', and 'the opposition party'. We also analyzed whether the speech occurred during or prior to the COVID-19 pandemic. Using these tags, we were able to quantitatively analyze how shifting political dynamics impacted the affective language choices made by Members of Parliament.

Our results indicate several differences in the political culture among parties as well as consistent effects of power dynamics. For instance, members of the party in power tend to use language that is more positive, and which instills a greater sense of control, whereas members of the opposition tend to do the opposite. In our analysis of language produced before and after the onset of the COVID-19 pandemic, we found that MPs from all represented parties, excluding the opposition, tended to use language that is more positive, less emotionally arousing, and which instills a greater sense of control. We interpret this as a sign that Canadian politicians come together in times of crisis and that this is reflected in their language use.

In conclusion, our findings show how complex cultural and political phenomena may be reflected in the lexical choices made by politicians. Recent psycholinguistic databases and corpus analysis tools have made it possible to probe those lexical choices and thus offer a way to capture the correlates of political culture.

References

Caruso, D., Frohman, L., Kinsman, R., & Sutherland. R. (2015). Some editing required: Producing Canada's Hansards. *Canadian Parliamentary Review, 35(2)*, 7-14.

Duncan Temple Lang (2020). *XML: Tools for Parsing and Generating XML Within R and S-Plus*. R package version 3.99-0.5. https://CRAN.R-project.org/package=XML

RStudio Team (2020). *RStudio: Integrated Development for R. RStudio*. PBC, Boston, MA URL http://www.rstudio.com/.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45(4), 1191- 1207.