Semantic Changes in Canadian Parliamentary Debates

Jungeun "June" Lim

Abstract

This project shows the possibility of using word embeddings for detecting cultural and social changes. Based on the Canadian Parliamentary debates dataset and focusing on neighborhoods of words in the word2vec models, it attempts to identify diachronic word replacement that happens due to cultural and political changes. In addition, it measures and visualizes the changing similarity among political parties to find evidence for the phenomenon of political polarization. Experiments to choose hyperparameters most suitable for these purposes and preliminary analyses with selected keywords related to social changes and political issues are discussed.

1 Introduction

When referring to the same thing or talking about the same topic, people in different time periods may use different words. Such a phenomenon, which is called diachronic word replacement, often reflect cultural and social changes. For instance, English speakers traditionally used the word men to refer to all human beings. However, such a usage is decreasing because it is now widely recognized that it is sexist language; the word "people" is used in place for "men" more often nowadays. Another example is metric units and imperial units. Canada was traditionally using imperial units, but since the 1970s, the country has converted to the metric units; "miles" has been replaced by "kilometres," for example. My first goal in this project is to see how such a replacement can be detected with word embedding in a simple way.

Distributed representation of words based on a corpus constructed by a certain group of authors or speakers may represent their mental models. If right-wing politicians and left-wing politicians exhibit different political views in their speeches, word embedding models built based on their speeches could also be different. My second goal in this project is to explore to what extent this assumption holds true and how it can be applied to quantify the increasing political polarization over time, which is a worldwide phenomenon actively studied in various fields of social science during the past couple of decades (Fiorina and Abrams, 2008; Conover et al., 2011).

2 Related Work

2.1 Using Word Embedding for Humanities and Social Science

Compared to other computational methods such as machine learning, topic modeling, and network modeling, word embedding is less used by humanists and social scientists (Schmidt, 2015). However, computer scientists have continuously suggested the possibility of applying word embedding to problems in humanities and social science. For instance, Kutuzov, Velldal, and Øvrelid (2017) showed that word embedding models trained with English news text could be used to represent and predict the temporal dynamics of armed conflicts. Garg, Schiebinger, Jurafsky, and Zou (2018) showed that the changing degree of bias against gender and ethnic minorities over time could be measured based on word embedding models.

While most of the previous work was done in the field of computer science, more and more humanists and social scientists are becoming interested in using word embeddings encouraged by the development of programming libraries readily available for implementing word embedding, in particular, word2vec. For instance, van Lange and Futselaar (2018) investigated changes in discussions about the punishment or war criminals in Dutch Parliament after World War II using word2vec. Rheault and Cochrane (2020) builds "party embeddings" with parliamentary debates from Britain, Canada, and the United States and represent politicians as vectors in a vector space with ideological dimensions by modifying word2vec. While this study does not make specific claims, it shows the potential of using word embeddings for political science research.

2.2 Quantifying Political Polarization Based on Political Texts

The increasing polarization of politics has received much academic attention in recent years. From the voting behavior of party members (Poole and Rosenthal, 1997) to the network of politicians and ideologically opinionated individuals on social media (Abul-Fottouh and Fetner, 2018), there are a variety of facts and data that have been utilized to quantify the degree of political polarization.

What people with explicit ideological positions say in political debates is one of such useful resources. Most of the studies analyzing political texts to measure the degree of partisanship assume that the difference in the languages people use when discussing political issues represent the ideological difference among them.

Gentzkow, Shapiro, and Taddy (2019), which analyzes the political polarization in the United States based on the proceedings of the United States Congressional Record from 1873 to 2009, is an example of the studies based on such an assumption. It defines the degree of political polarization as how easy it is for an observer to infer a congressperson's party solely by listening to her speech. By constructing a Bayesian structural choice model to estimate the change in the easiness over time, it shows that the language uses between Republicans and Democrats have become increasingly different from each other and suggests that it is an evidence of political polarization.

Based on the same assumption, supervised machine learning algorithms trained to classify the partisanship of the speakers or writers solely based on what they said or wrote were also used in many studies. Such studies typically define the degree of political polarization as the accuracy of their classifiers. For instance, Peterson and Spirling (2018) shows their classifier becomes more accurate when they test with more recent datasets of British parliamentary debates from 1935 to 2013, and Alexander (2019) reports a similar result with Australian parliamentary debates dataset.

However, to the best of my knowledge, there have been no studies that utilize word embeddings to specifically quantify political polarization.

3 Data

The digitized, annotated, and cleaned transcript of Canadian parliamentary debates from 1901 to 2019 can be freely downloaded (Beelen et al., 2017). The dataset consists of .csv files, each for a day when the parliamentary debate was held. Along with the transcript of speeches, extra information such as the party association of the speakers and the topic of the speech (e.g., immigration, environmental affair, income tax act) are included. Each .csv file is located in the corresponding directory of the month, and each directory of the month is located in the corresponding directory of the year.

3.1 Data Preprocessing

After lowercasing all the transcripts, tokenizing was done with regular expressions, considering only the words containing only alphabets as tokens.

For the purpose of detecting diachronic word replacement, I divided the data into six time slices, each based on 20 years' worth of scripts (1901-1920, 1921-1940, 1941-1960, 1961-1980, 1981-2000) except for the most recent time slice which contains 19 years' worth of scripts (2001-2019).

For the purpose of looking for evidence of political polarization, I focused on the past 35 years since political polarization is a recent phenomenon that began to be discussed as an academic subject a few decades ago. Three word embedding models – one for the Liberal Party, one for the Conservative Party, and one of the New Democratic Party – were constructed per each 5-years long time slice. Since the Conservative Party and the New Democratic Party have changed their names and merged with other parties throughout history, I also considered the older names of these parties. For instance, 'Conservative', 'Conservative (1867-1942)', 'Progressive Conservative', 'Conservative Party of Canada', 'Liberal-Conservative', 'Canadian Alliance', and 'Reform' were all considered as the Conservative Party.

4 Computational Methodology

The code used for this project can be found at https://github.com/JuneJLim/CSC2611_project_2020. I chose word2vec models over other approaches that I initially considered (latent semantic analysis and Glove) because of its good performance reported in previous studies that utilized word2vec models in the field of humanities and social science (van Lange and Futselaar, 2018; Rheault and Cochrane, 2020) and its reasonable amount of memory use and run time that allowed me to work on my personal computer. To construct word2vec models, the Python library `gensim` (Rehurek and Sojka, 2010) was used. The following subsections describe the consideration I took when setting hyperparameters to build the models and my approach to compare words in different models.

4.1 Word2vec Hyperparameters: Effect of Different Window Sizes

When training a word2vec model, one needs to determine two parameters: the dimensionality of the vectors and the size of the context window (i.e., the number of words that occur before and after a word).

The effect of the first parameter on the trained model is rather simple; the higher the dimensionality is, the better the quality of the embedding while the gain in the quality diminishes after some point (Mikolov et al., 2013a). Since higher dimensionality requires more memory space and longer computation time, typically the dimensionality between 50 and 1,000 is used. I set the dimensionality to 300 for all the models, considering the computational resource available to me and following the previous studies on similar topics.

On the other hand, how to decide on the window size is not very straightforward. Many of the studies in the field of computer science that implement word2vec models set the window size to 5 or 10. However, it does not mean that the window size smaller than 5 or bigger than 10 is less useful or less efficient. I find that the decision should be made based on the purpose of building the model and the characteristics of the corpus being used. For instance, Rheault and Cochrane (2020), a study in the field of political science that builds word2vec models based on parliamentary corpora, uses the window size of 20, which roughly corresponds to the average length of a sentence in British parliamentary debates (19.56 tokens per sentence on average) and in Canadian parliamentary debates (20.62 tokens per sentence on average) which is longer than, let's say, what social media users write online. The authors explain that they chose this number because the same topic tends to be discussed across several sentences in an individual speech. Because their interests lie in the political differences among different party members represented in their language use, their claim sounds persuasive.

Window Size of 5	Window Size of 20
programs	programs
goals	*expertise*
basically	overall
significantly	goals
foster	*criteria*
con	*transparency*
currently	*choices*
venture	*challenges*
overall	con
addresses	*major*

Table 1: Top 10 changed words over the past 100 years in Canadian Parliamentary debates, captured with word2vec models built with different window sizes (namely, 5 and 20). Italicized words are the words that do not overlap with the top 10 changed words captured by the other models built with the different window size.

In a broader sense, this claim implies that, while a word embedding model built with a bigger window size may be more sensitive to changes in cultural and social uses or implications of words, word embedding models built with a smaller window size may capture changes in word meanings in general and typical linguistic styles in certain time periods.

Table 1 shows the top 10 changed words over the past 100 years in Canadian Parliamentary debates captured with word2vec models built with different window sizes (5 in the left column and 20 in the right column). The result supports the hypothesis above. We can see three adverbs (basically, significantly, and currently) only in the list of most changed words suggested by the models built with the window size of 5. On the other hand, the list of most changed words suggested by the models built with the window size of 20 mostly comprises nouns. In particular, words that only appear in this list such as transparency and choices have not experienced fundamental changes of meaning but are often used in newer contexts that recently emerged from cultural and social changes.

Based on this observation, all the word2vec models built for this project used the window size of 20 because the focus of this project lies in social and cultural changes.

4.2 Measuring Semantic Change of a Word Based on Relationship with Other Words

Rather than directly using word vectors in word2vec models, I decided to focus on the relationship (as cosine similarity) between the word of interest and all the other words in the model. This is an intuitive method that eliminates the need for space alignment of two different embeddings, which does not always guarantee a reliable comparison between two models.

This approach starts with constructing a list of words shared among multiple word2vec models to be compared. Let's say one wants to compare two word2vec models. They could be one for the past and one for the present, or one for the Liberal Party and one for the Conservative Party. While all the common words between the two models could be used, one may use only *N* most frequent words to avoid a memory issue (I used 8,000). For each model, one constructs an *N* by *N* matrix, in which each row represents cosine similarity between a word and all the other words in the model. I call this matrix *neighborhood relationship matrix* and each row in this matrix *neighborhood relationship vector*.

This representation of the relationship between all the possible word pairs in a model is used for both of the goals of this project: comparing individual words in different models (to detect diachronic word replacement) and comparing the whole model to the other model (to examine the difference in the language uses of two political parties). For the first purpose, for instance, to see how the word "*miles*" in the past is being replaced, I first take the neighborhood relationship vector corresponding to "*miles*" from the neighborhood relationship matrix for the past time slice. Next, I look for the most similar vector to it in the neighborhood relationship matrix for the present time slice. If this approach can detect the actual diachronic replacement from "*miles*" to "*kilometres*," the vector I get will be the one corresponding to "*kilometres*."

The same approach can be taken in the opposite direction (present to past, rather than past to present). An illustrating example is going from "*kilometres*" to "*miles*." To do so, I take the neighborhood relationship vector corresponding to "*kilometers*" from the neighborhood relationship matrix for the present time slice and look for the most similar vector to it in the neighborhood relationship matrix for the past time slice. If I am successful, the vector I get will be the one corresponding to "*miles*".

It is important to note that going in both directions makes sense only for some cases in which the two words are almost completely interchangeable. For instance, "*men*" and "*people*" are not completely interchangeable, obviously, because the word "*men*" also means adult human males, not human beings in general.

To compute the similarity between two entire word2vec models, one may iterate over all the words to compute the similarity between each word's neighborhood relationship vector from the two models. This computation is time-consuming; if *N* most frequent words are used, one has to make *N^2* comparisons. Because of this issue, I used only 1,000 most frequent words rather than 8,000 here when comparing two political parties. Finally, with *N* cosine similarity values for each word, one may compute the mean of these values and consider it as the difference between two models.

5 Result

5.1 Diachronic Word Replacement

Table 2 shows some examples of diachronic word replacement I acquired by going from the

1901-1920: "Y"	2001-2019: "X"	Cosine Similarity
men	people	0.6188
chairman	*chair*	0.5009
miles	*kilometres*	0.5310
canada	world	0.5918
train	*plane*	0.6045
gas	electrical	0.6222
ought	*should*	0.6649
pleased	*happy*	0.6742

Table 2: Selected examples of word replacement happened across the past 100 years. X is "the new Y."

2001-2019: "X"	1901-1920: "Y"	Cosine Similarity
refugees	immigrants	0.5668
chair	*chairman*	0.5009
kilometres	*miles*	0.6537
gay	jews	0.5491
plane	*train*	0.6545
great	splendid	0.5171
should	*ought*	0.6550
happy	*pleased*	0.6650

Table 3: Selected examples of word replacement happened across the past 100 years. Y was "X in old days."

past (1901-1920) to present (2001-2019). Table 3 shows the examples I acquired by going from present to past. To extract possible cases on diachronic word replacement, I iterated over all the words looking for the cases in which (1) the word in the second model that is most similar to the word of interest in the first model is not equal to the word of interest and (2) the cosine similarity between two neighborhood relationship vectors that are most similar to each other is larger than some threshold (I used 0.5). The most illustrating examples were selected from the extracted word pairs.

Before discussing these examples further, it would be helpful to refer to the four types of diachronic word replacement suggested by Tahmasebi, Borin and Jatowt (2018): (a) lexical replacement that requires sense information to detect it (e.g., "nice" to "foolish"), (b) terms that describe the same entity/object at different times (e.g., Myanmar to Burma), (c) terms that are instances of the same type that were valid at different times (e.g., Bush to Clinton), and (d) temporal analogs which share roles, attributes, functions despite time gap (e.g., Walkman vs. iPod). Among these, (b) and (c) are mostly limited to proper nouns.

Cases of (a) and (d) are mixed in the examples I acquired. For instance, "men" to "people" and "ought" to "should" could be classified into type (a). While "ought" was used frequently in everyday English 100 years ago, it is now considered somewhat old-fashioned and "should" is more frequently used in place of "ought." (These two words have slightly different meanings in a strict sense, but they are often considered interchangeable.) Here my definition of type (a) is a little wider than that in Tahmasebi, Borin, and Jatowt (2018); I suggest that the cases in which different words are generally preferred to refer to the same thing in the different time periods could also be considered as type (a). "Train" to "plane" and "gas" to "electrical" could be considered as type (d). While we still ride trains and use gas, planes have replaced trains and electricity replaced gas in many aspects of our lives.

While my approach did capture some diachronic word replacement, it has several limitations. The limitations and the ways to address them will be discussed in the Discussion section at the end of the report.

5.2 Political Polarization Over Time

Since I take three major political parties in Canada – the Liberal Party, the Conservative Party, and the New Democratic Party – into consideration, there are three pairs of parties for me to analyze. The three charts based on those pairs are shown in Figure 1. Each chart in Figure 1 shows the changes in the similarity between two word2vec models trained based on the speeches by the members of two Canadian political parties over the past 35 years. Each bar is corresponding to one 5-years long time slice.

Increasing political polarization is only observable in the top chart that shows the similarity between the Liberal Party and the Conservative Party; the similarity between them is decreasing over time. This is not surprising because, as the names of the parties suggest, they have been the two dominant parties over the history of Canada, one traditionally represents conservatives and the other traditionally represents liberals. Because the New Democratic Party, which became the Official Opposition for the first time in 2011 and has never been the ruling party, has a socialist root and is a social-democratic party, political polarization may not be applicable to it.

I find that two other charts showing the similarity between the New Democratic Party and the other parties show interesting trends that reflect actual political facts and changes. First, the Conservative Party and the New Democratic Party are least similar to each other. This obviously makes sense and validates using word embedding models for examining the political difference between parties. Second, the similarity between the Liberal Party and the New Democratic Party seems to be corresponding to the number of seats that the New Democratic Party won. When the New Democratic Party becomes more similar to the

Liberal Party, the New Democratic Party appears to win more seats in the federal election after a few years. While a more thorough analysis is needed to

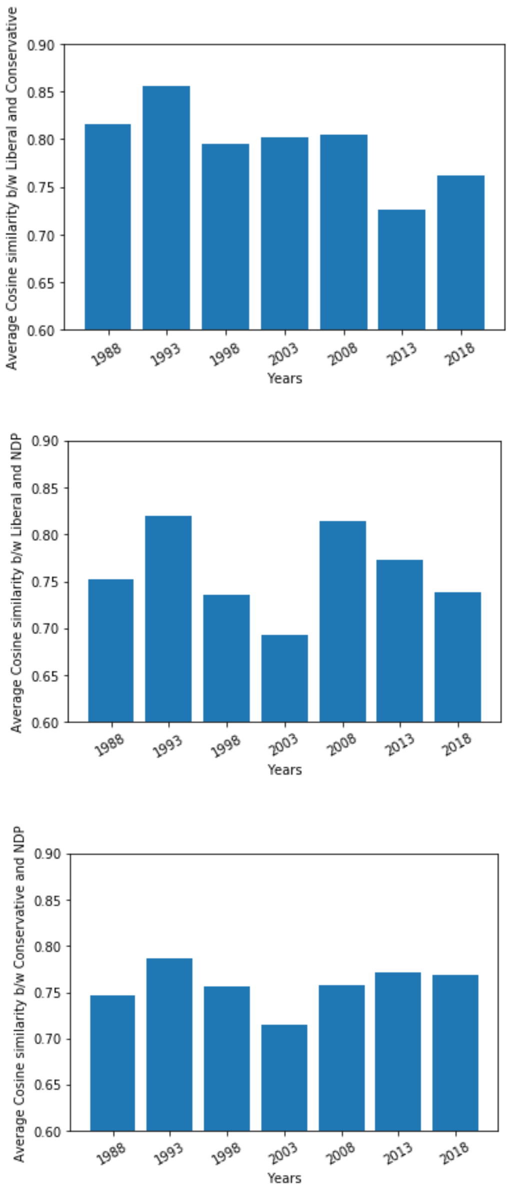

Figure 1: Changes in the similarity between two word2vec models trained based on the speeches by the members of two Canadian political parties over the past 35 years. The top chart shows the similarity between the Liberal Party and the Conservative Party. The middle chart shows the similarity between the Liberal Party and the New Democratic Party. The bottom chart shows the similarity between the Conservative Party and the New Democratic Party.

make such a claim, this may suggest that being less extreme or being closer to mainstream affects the election result positively.

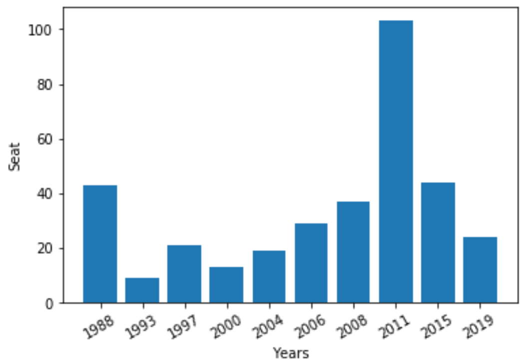

Figure 2: Changes in the number of seats that NDP has won in the Canadian parliament.

5.3 Case Studies: Changes in Politics-related Words Over Time

To see if changes in Canadian history and differences between Canadian political parties can be detected in word2vec models, I conducted a preliminary analysis with keywords related to cultural changes and political issues. To gather the keywords to use as query terms, I referred to Bennett, Grossberg, and Morris (2005) and Lakoff (1997).

Table 4 shows the difference in the nearest neighbors of the word "equality" between the time period of 1921-1940 and 2001-2019. The nearest neighbors of the word "equality" during 1921-

10 nearest neighbors of "equality"	
1921-1940	**2001-2019**
equal	inclusion
status	parity
nationhood	equal
bargaining	equity
independence	advancement
collective	acceptance
footing	diversity
harmony	pluralism
freedom	inclusiveness
enduring	tolerance

Table 4: The 10 most similar words to the term "equality" based on word2vec models trained with 20-years' worth of speeches in Canadian parliament during 1921-1940 and 2001-2019.

1940 indicate that equality was mainly discussed as something related to “status,” “freedom,” and “independence,” to pursue or fight for as a “nation” with a “collective” and “enduring” effort. They seem to be reflecting the fact that independence from Britain was still an important issue in Canada at that time. On the other hand, the nearest neighbors of the word "equality" during 2001-2019 imply that equality is something about “inclusion,” “acceptance,” “diversity,” “pluralism,” and “tolerance” that “advances” the society. They seem to be reflecting the increasing number of immigrants that Canada is receiving and how the Canadian government deals with them.

Table 5 shows the difference in the nearest neighbors of the word "abortion" among the three major political parties in Canada. In the Conservative Party members’ speeches, the word “abortion” is closer to negative words related to cruelty, death, crime, and violence. On the other hand, in the speeches of two other parties' members, “abortion” is closer to neutral words mostly related to pregnancy itself and medical procedure. This contrast reflects the difference between conservative views and libertarian perspectives on abortion.

Finally, inspired by Garg, Schiebinger, Jurafsky, and Zou (2018), I analyzed words referring to gender and races. Figure 3 shows that while “woman” and “man,” “Canadian” and “Asian,”

Conservative Party (2016-2019)	**Liberal Party (2016-2019)**	**NDP (2016-2019)**
punishment	abortions	contraception
termination	reproductive	clinic
bestiality	counselling	reproductive
inhumane	shelter	sex
animal	adolescents	medical
fraudulent	vocational	remuneration
arson	mutilation	dual
bodily	genital	insurmountable
blasphemy	psychosocial	surgery
euthanasia	contraception	specialized

Table 5: The 10 most similar words to the term “abortion” based on word2vec models trained with 5-years’ worth of speeches by members of different Canadian political parties.

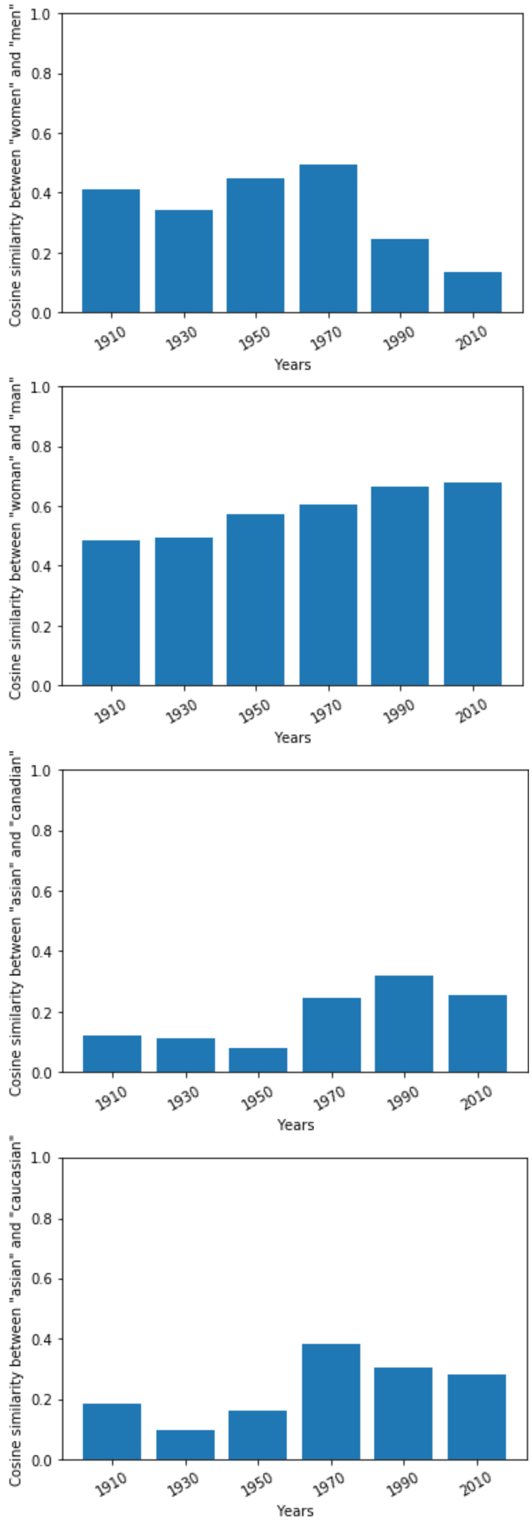

Figure 3: Changing cosine similarity between two words over time. The top chart shows the similarity between "woman" and "man." The second top chart shows the similarity between "women" and "men." The second bottom chart shows the similarity between "asian” and “canadian.” The bottom chart shows the similarity between “asian” and “caucasian.”

and "Caucasian" and "Asian" are becoming more similar to each other, "women" and "men" are becoming less similar to each other. This result well-reflects the changes in reality. While minorities are treated more equally as individuals in political debates, the issues unique to minority groups are getting more attention among politicians. For instance, while an individual woman is increasingly treated in the same way as a man, discrimination that women as a minority group experience is more actively discussed nowadays.

6 Discussion and Future Work

I have shown the possibility of using word embeddings for detecting cultural and social changes. To be more specific, using the Canadian parliamentary debates dataset and focusing on neighborhoods of words in the word2vec models, I tried to 1) detect diachronic word replacement and 2) quantify differences among political parties to confirm the phenomenon of political polarization. In addition, 3) I have conducted a preliminary analysis that shows the possibility of detecting social and cultural changes and different views of political parties utilizing word embedding models.

My approach to detect diachronic word replacement has several limitations and room for improvement. First, while the algorithm did capture some diachronic word replacement, many of the possible cases of diachronic word replacement automatically detected seem like just a pair of synonyms. Even though some of them may actually represent slight changes in trends in the language use of Canadian politicians, there is no way to validate them. One way to address this issue is to consider the frequency of words along with top *N* most similar words rather than just one most similar word. If word replacement has been actually occurring, the frequency of old words should be decreasing while the frequency of alternative words should be increasing over time. Second, while it did capture some diachronic word replacement, it cannot distinguish different types of replacement.

Even though my initial goal was to see if comparing word2vec models can be evidence for political polarization, it also yielded some interesting results that appear to reflect political situations and changes in Canada. To confirm my interpretation is reliable, more validation should be done against the historical facts and existing studies on Canadian politics. While the topic of each speech was included in the dataset I used, I did not make use of it. Word embedding models trained with the speeches about a specific controversial issue about which two political parties have different opinions could be helpful for validating my approach and may bring more interesting insights about Canadian politics and how to use word embeddings for social science studies.

The algorithm used to construct word2vec models in this project, skip-gram with negative sampling (SGNS), involves random factors such as random initialization, random sampling, and stochastic gradient descent. These random factors make it impossible to replicate exactly the same models and results from scratch. Antoniak and Mimno (2018) claims that, if such variability is not addressed, it could undermine the robustness of research utilizing word embedding models to understand the linguistic, social, or cultural characteristics of a certain group of people who created the corpus from which the models were built. For instance, there is a possibility that two models trained with the same corpus and hyperparameter could yield a great variation in top *N* most similar words to a certain word. To address this issue, I may train several models to see if the variation is in the acceptable range and take the results that only appear consistently to ensure that my analysis is reliable in the future.

References

Deena Abul-Fottouh and Tina Fetner. 2018. Solidarity or Schism: Ideological Congruence and the Twitter Networks of Egyptian Activists. *Mobilization: An International Quarterly*, 23(1):23–44.

Rohan Alexander, 2019. *A Word-Count Based Classifier of Politicians in the Australian Federal Parliament (1901–2018).* [online] Available at: < https://rohanalexander.com/academic.html> [Accessed 7 April 2020].

Maria Antoniak and David Mimno. 2018. Evaluating the Stability of Embedding-based Word Similarities. *Transactions of the Association for Computational Linguistics*, 6:107–119.

Kaspar Beelen, Timothy Alberdingk Thijm, Christopher Cochrane, Kees Halvemaan, Graeme Hirst, Michael Kimmins, Sander Lijbrink, Maarten Marx, Nona Naderi, Ludovic Rheault, Roman Polyanovsky, and Tanya Whyte. 2017. Digitization of the Canadian Parliamentary Debates. *Canadian*

Journal of Political Science/Revue canadienne de science politique, 50(3):849–864.

Tony Bennett, Lawrence Grossberg, and Meaghan Morris. 2005. *New keywords: A revised vocabulary of culture and society*. Blackwell Pub, Malden, MA.

Michael D. Conover, Jacob Ratkiewicz, Matthew Francisco, Bruno Goncalves, Filippo Menczer, and Alessandro Flammini. 2011. Political Polarization on Twitter. In *Fifth International AAAI Conference on Weblogs and Social Media.*

Morteza Dehghani, Kenji Sagae, Sonya Sachdeva, and Jonathan Gratch. 2014. Analyzing Political Rhetoric in Conservative and Liberal Weblogs Related to the Construction of the "Ground Zero Mosque." *Journal of Information Technology & Politics*, 11(1):1–14.

Dorottya Demszky, Nikhil Garg, Rob Voigt, James Zou, Jesse Shapiro, Matthew Gentzkow, Dan Jurafsky. 2019. *Analyzing Polarization in Social Media: Method and Application to Tweets on 21 Mass Shootings*. NAACL-HLT (1):2970-3005.

Nikhil Garg, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. Word embeddings quantify 100 years of gender and ethnic stereotypes. In *Proceedings of the National Academy of Sciences*, 115(16):E3635–E3644.

Matthew Gentzkow, Jesse M. Shapiro, and Matt Taddy. 2019. Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica,* 87(4):1307–1340.

Stefano Gurciullo and Slava J. Mikhaylov. 2017. Detecting policy preferences and dynamics in the UN general debate with neural word embeddings. In *2017 International Conference on the Frontiers and Advances in Data Science (FADS)*, pages 74–79.

Andrey Kutuzov, Erik Velldal, and Lilja Øvrelid. 2017. Temporal dynamics of semantic relations in word embeddings: an application to predicting armed conflict participants. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing,* pages 1824–1829, Copenhagen, Denmark.

George Lakoff. 1997. *Moral Politics: What Conservatives Know That Liberals Don't.* University of Chicago Press, Chicago, IL.

Aaron M. McCright and Riley E. Dunlap. 2011. The Politicization of Climate Change and Polarization in the American Public's Views of Global Warming, 2001–2010. *The Sociological Quarterly*, 52(2):155–194.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013a. Efficient Estimation of Word Representations in Vector Space. In *ICLR Workshop Papers.*

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems* - Volume 2, pages 3111–3119, Lake Tahoe, Nevada.

Morris P. Fiorina and Samuel J. Abrams. 2008. Political Polarization in the American Public. *Annual Review of Political Science*, 11(1):563–588.

Andrew Peterson and Arthur Spirling. 2018. Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems. *Political Analysis,* 26(1):120–128.

Keith Poole and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* Oxford University Press, New York, NY.

Gabriel Recchia, Ewan Jones, Paul Nulty, John Regan, and Peter de Bolla. 2017. Tracing Shifting Conceptual Vocabularies Through Time. In Paolo Ciancarini, Francesco Poggi, Matthew Horridge, Jun Zhao, Tudor Groza, Mari Carmen Suarez-Figueroa, Mathieu d'Aquin, and Valentina Presutti, editors, *Knowledge Engineering and Knowledge Management*, pages 19–28, Cham. Springer International Publishing.

Radim Rehurek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta.

Ludovic Rheault and Christopher Cochrane. 2020. Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora. *Political Analysis*, 28(1):112–133.

Ben Schmidt, 2015. *Vector Space Models For The Digital Humanities*. [online] Available at: <http://bookworm.benschmidt.org/posts/2015-10-25-Word-Embeddings.html> [Accessed 7 April 2020].

Miriam Smith. 2007. Framing Same-sex Marriage in Canada and the United States: Goodridge, Halpern and The National Boundaries of Political Discourse. *Social & Legal Studies*, 16(1):5–26.

Nina Tahmasebi, Lars Borin, and Adam Jatowt. 2019. *Survey of Computational Approaches to Lexical Semantic Change*. *arXiv*:1811.06278 [cs].

Milan M. van Lange and Ralf D. Futselaar. 2018. Debating Evil: Using Word Embeddings to Analyze Parliamentary Debates on War Criminals in The Netherlands. In *Proceedings of the conference on*

Language Technologies & Digital Humanities 2018, pages 147–153.

Lu Wei, Wang Jiamin, and Hu Jiming. 2020. Analyzing the topic distribution and evolution of foreign relations from parliamentary debates: A framework and case study. *Information Processing & Management*, 57(3):102191.