Word embeddings for the analysis of ideology and identity in Canadian politics

Department of Political Science, University of Toronto

Dissertation Research Proposal

Justin Savoie

2021-09-16

Contents

Chapter 1

Introduction

Patterns of political beliefs are not random. People who hold certain ideas often hold others. An individual on the "political left" is likely to hold views clustering in certain specific ways with regards to, for instance, economic redistribution, immigration policy, environment policy, or human rights. Similarly, an individual on the "political right" will hold relatively coherent beliefs, and often different, to those held by a leftist. Of course, the political space is not perfectly unidimensional or bi-dimensional, or x-dimensional. Researchers and analysts often disagree about the number of dimensions, or the best ways to label them. At minimum, they usually agree it depends on context. For instance, introductory textbooks talk of a unidimensional left-right spectrum, or a bi-dimensional (Authority/Liberty–Left/Right) spectrum (Heywood, 2017). Héroux-Legault (2016) uses statistical modeling to identify five dimensions in the Canadian context. Neither do such clusters or dimensions possess universal or time-invariant validity. In the Canadian province of Québec, it is widely recognized that there has been a realignment, from a dominant federalism/sovereignty dimension to a more traditional left-right dimension (Montigny, 2016). Nonetheless, there is a consensus that "left" and "right" are meaningful concepts when talking about politics, and that spatial representations of politics do help us make sense of the political landscape. One can make sense of ideology through some form of "empirical reduction".

Similarly, one can make sense of "identity" through similar techniques. Without a doubt, political identities are plural and wavering (Hobsbawm and Ranger, 2012, Ozkirimli (2017)). Still they can be

fruitfully studied to extract a core; to tell something about the identity's root (Schwartz, 1967). As the researcher can place the subject amid a spatial ideological representation, one can associate empirical symbols to cluster of political identities (Dufresne et al., 2019).

One important way, and in a sense the only way, to study political beliefs (about ideology or identity) is through language; through what people say orally and in writing. All human communication requires language, words, and meanings. This is apparent in all approaches to the study of politics. It is most obvious in political theory, where commentators reflect at lengths over questions of politics, justice, or institutions, and in qualitative research, where analysts examine the deep structure of some data, usually linguistic in nature. Yet, in the social sciences, language also features centrally in quantitative work; most survey statistics uses linguistic data (questions, survey choices), ideal points models estimate ideology from roll-call data, i.e., data representing whether a politician agrees with specific policies or bills.

Traditionally, scholars have studied patterns of political beliefs (political ideology or political identity) either through quantitative approaches (surveys, scaling, scores of various natures, etc.) or through theoretical approaches (political theory, interpretative approaches, critical approaches, etc.) Of course, it has been common to combine these approaches; to combine thicker and thinner descriptions, as befits the object of study.

Automatic content analysis methods have long promised to try to bridge the gap between quantitative and qualitative-theoretical perspectives. It has been difficult: extracting meaning from text is hard. It is not for lack of effort: political scientists have laboured intensively on text scaling methods. Approaches based on word frequencies in text (Slapin and Proksch, 2008, Laver et al. (2003)) allow to compare policy positions. Supervised classification of political texts into known categories help us study tone, partisan affiliation, or any pre-determined categories which an algorithm can learn and then put texts into. These methods have allowed us to make important advances. Nevertheless, they remain at a high level of abstraction. The unit of analysis is the document in the case of text scaling; classification is useful but does not really get at the sense or the meaning in the text.

Word embeddings have recently emerged as a promising way to get at meaning. Word embeddings stem from statistical semantics and can be summarize by the distributional hypothesis: words used and occurring in similar contexts tend to have similar meanings. It is then possible to model and represent

in high dimensional space each word (or each political concepts if they are represented by one or several words). Technical advances (the idea behind word embeddings has been around for decades but efficient and mainstream implementation is recent (Mikolov et al., 2013b, Mikolov et al. (2013a)) have the potential to offer new perspectives to the study of political meaning. Some researchers have started using such methods, but, in the social sciences, their use is still in its infancy. My dissertation uses statistical semantics to study ideology and political identity in Canada. Using textual data from transcripts of Parliamentary Debates (1900-2019) and Twitter data, I study patterns of word use, over time and context (e.g. elites vs non-elites, left vs right). I aim to get at word and concept meaning, using word vectors, in the context of discussions about Canadian politics.

Chapter 2

Research questions

The main research question in this study is: can we use word embeddings to document and map the changing meaning of key terms used when talking about politics in Canada? Is it possible to reproduce, even in an approximate form, a 'commonly-accepted-among-scholars' analysis of the development of the meaning of four key concepts in Canadian politics: rights, identity, citizenship, and nation. Such an approach combines the scope of conceptual approaches with at least some of the breadth of linguistic approaches, while capitalizing on quantitative method's empirical grounding. I define ideology and identity in a broad sense as the thoughts, beliefs, attitudes, behaviours, and values that are related to politics.

Chapter 3

Literature review

3.1 From quantitative large-N surveys to word embeddings

People holding certain political beliefs often hold others. In this dissertation, I call ideological analysis the general study of those political beliefs. Ideological analysis, in this context, includes the study of what empirical political science usually calls ideology (latent policy dimension), but also symbols and markers of political identity. Ideology, thus understood, is a loaded term; it can mean very different things. These meanings range from the French ideologues's "objective" study of ideas just before and during the French revolution (Kennedy, 1979; Head, 1985) to illusion, mystification or false consciousness in Marx and Engels' German Ideology (1970), or a morphological, conceptual, non-pejorative map of the sociopolitical world (Freeden et al., 2003), and include various spatial scaling methods; measuring what is called the left and right, conservative and liberal positions; in one, or multiple dimensions; measurable in real life (Downs, 1957; Aldrich and McKelvey, 1977; Hare et al., 2015; Cochrane, 2015) or online (Barberá, 2015; Temporão et al., 2018). Understanding precisely the patterns that hold these ideas together is a matter of considerable debate. Maynard (2013) divides in three broad groups these motley vantage points: conceptual approaches, discursive approaches and quantitative approaches.

Conceptual approaches focus on systems of ideas: concepts and their complex interrelations function as the unit of analysis. Ideologies provide frameworks from which individuals understand the world.

Freeden's (2006) morphological approach is an influential example. Foundational principles such as "liberty", "security" or "equality" are shared by the great vast majority, yet adherents of different ideologies understand those principles in a different and sometimes conflicting way. Political concepts acquire meaning through discourse, culture and positionality. Ideologies evolve through conceptual competition; through "decontestation". Other scholars working in this tradition include Reinhart Koselleck (conceptual history), and Quentin Skinner (Cambridge School of historicist and contextualist interpretation). Some have formalized this analysis of the structures that underlie social interactions, via, for instance, complex systems (Thagard, 2012; Homer-Dixon et al., 2014).

Discursive approaches focus on communicative practices. They are less concerned with content, and more concerned with visibility, power, and influence. Crucially, they usually involve a critique of ideology from the left. While epistemological disagreement exists on matters of relations to truth, with post-structural discursive approaches (Mouffe, 1985; Žižek, 1989) more sceptical than critical discourse approaches (Van Dijk, 2006; Fairclough, 1992), both strands share a commitment to the critique of naïve conceptions of truth (Kellner, 1989). Vast methodological differences also cleave the field, from the study of semantics (Pêcheux, 1975), to the post-Marxist attention paid allegory (Jameson, 2020), speed (Baudrillard, 1994) , spectacle (Debord, 2012), or empty signifiers (Laclau, 2007). Through the study of linguistics or discursive constructions, critical approaches hence shed light on the origins of, and the process by which, ideas, which may or may not be, but usually are, problematic and alienating, come to reproduce themselves to the point of insidious symbolic hegemony. Recently, discursive analysis has been extended to include statistical based approaches (Nafstad and Blakar, 2012) relying largely on word count. In the context of ideological analysis, frequency usage of communality words (solidarity, welfare) or words linked to consumerism can provide insight in possible changes in the way individuals construe and reproduce political behaviours (Nafstad and Blakar, 2012).

Quantitative approaches use large-N analyses to study patterns of beliefs. A divide exists between observational studies of ideology studying citizens' stated preferences (Bonica, 2013) or legislators' voting behaviour (Martin and Quinn, 2002), and approaches in social psychology, rather focusing on cognitive biases, links between belief systems and personality; ideology as social cognition, for instance by linking political preferences to psychological inclinations toward uncertainty or threat (Jost et al., 2013). Quantitative approaches also include traditional dimensionality reduction, where values are

reduced to dimensions: see Héroux-Legault (2016) for an investigation of the Canadian case; Cochrane (2010) for a study of ideological asymmetries.

As such, ideological analysis is fragmented. For instance, experimental social psychology can look down on non-experimental approaches, where falsification is less straightforward. Conversely, non-quantitative scholars can look down upon the poor external validity of experiments, or their theoretical levity. Calls have been made to see, use and combine the full variety of available approaches. Especially, to see ideological analysis as "a field in its own right": researchers from various strands have common core objectives, how people talk about politics, and how it is significant (Maynard, 2013).

My dissertation proposes such a combination of frameworks to study ideology. Substantively, I aim to map the changing meaning of political words over ideology, time and ideology/time interactively. Methodologically, I will call upon recent advances in the computer science field of natural language processing. Namely, the dissertation's methodological workhorse will be word embeddings: a set of techniques where words, phrases or documents are mapped to vectors of numbers. Conceptually, word embeddings organize words, phrases or documents in high dimensional space, where similar items are close to each other and dissimilar items are not. Perhaps, 'immigrant' will be close to 'refugee' but not to 'parliament'; 'parliament' will be close to 'institution', but not to 'climate change'. Embeddings are calculated based on distributions of words in large corpora, and their proximity to other words. In other words, in linguist J.R. Firth's (1957) now celebrated terms: "You shall know a word by the company it keeps".

Most work in the quantitative social sciences has studied ideology from the point of view of dimensionality reduction and regression. Dimensionality reduction is ambiguous in what it reveals. Dimensionality reduction tells us if there is a fit between political beliefs (Gregg and Banks, 1965), but not how it forms or where it's from. Regression analysis finds bivariate or multivariate links. For instance, in average, education is related to views on abortion. Neither approaches are perfect as they often fail to explain family resemblance-like phenomena. Left-right positioning is one such phenomena (Cochrane, 2015). It arises when it's not that members of a family have significantly bigger noses, or lighter-eyes, or darker hair than the rest of the population. Rather, it's when some combination of those characteristics makes it likely to hold others. In politics, such combinations are frequent. Studying the emergence of fringe communities and political echo chambers using closeness centrality in a network can reveal multiple

communities and their interactions in a way impossible to traditional statistical methods (Boutyline and Willer, 2017). Ideological scaling and spatial models of politics using traditional statistical methods or network approaches then allow to address one of the most salient questions for academics and citizens alike: to what extent is society ideologically polarized, and is there increasing polarization?

Work on ideology and spatial representation has mostly used likert-type items. Dimensionality reduction is often used on categorical, ordinal or numerical variables. This way of working limits possible nuance for the respondent but simplifies the researcher's task. Working on a set of questions with answers varying only over a certain range or a certain support entails postulating that the ideological world can be conceptualized as such. Yet, making this chief assumption allows for useful and enlightening insight on the structure of politics (for a classic example see Aldrich and McKelvey (1977)). It certainly would involve some work, but it appears reasonable to believe that replacing close-ended question by longer open-text when studying political ideology, would allow to extract more substance, and increase the precision of our ideological model (Bauer et al., 2017).

Recent work has focused on exactly this: fitting models predicting the use of words in context to generate estimates of ideological placement (Rheault and Cochrane, 2020; Rodman, 2020). Thus, I propose to extend such work. These models often use word embeddings (Rumelhart et al., 1986; Mikolov et al., 2013a,b). Each word is mapped to a large numerical vector. The given vector represents the context in which the word is generally used. The method depart significantly from bag-of-words methods traditionally used in political science where word count is the main predictor of various outcome measures. Rheault and Cochrane (2020) have argued that taking into account the role of words in context is especially promising for scholars attempting to measure political ideology. Patterns of ideas often are tangible, but irreducible as ideology functions in a network-like way: ideas are loosely associated, clusters or patterns emerge, people know ideology when they see it, yet struggle to name things as they are. Ideology functions in a plastic way. (Cochrane, 2015) Word embeddings, since words are at the center of the network of words with which they are used, are designed to capture such interplay.

Usually, researcher study only one single type of texts. Rheault and Cochrane (2020) study parliamentary texts. Rodman (2020) analyses newspaper articles from the New York Times, Reuters and Associated Press. Rodriguez and Spirling (2021) model words in congressional speeches. This disserta-

tion aims to model word use and ideology from various sources: citizens, journalists, and parliamentary text. Political ideology does not arise from debates in parliament only. Similarly, pundits or political theorists do not have the monopoly over the way language is used to express political belief. Ideology emerges not within one social group or institution, but through social practices shared by citizens, politicians, pundits, theorists alike. There will be some differences, but a language-based theory of ideology must be applicable to more than one case. The study of ideology should be pluralistic in its objects: various actors, from all over society, produce ideological speech, in turn influencing the speech of various actors.

One common difficulty arising from the modelling of ideology through text motivates this research. Often, such modelling identifies other dimensions, clearly non-ideological (Hirst et al., 2010; Lauderdale and Herzog, 2016). Instead of capturing ideology, statistical models use words to learn "something else", like expressions of attack and defence, opposition and government, geographical differences, linguistic complexity, and so forth. Such findings emerge mostly from studies using word-count-like features, not word embeddings. Rheault and Cochrane's findings suggest embeddings might be a solution to this problem of ideological identifiability. The dissertation aims to test this solution on text non-parliamentary texts, and in more than one language.

One further difficulty concerns the measure of uncertainty. Producing estimates without measuring uncertainty can be misleading. Estimates without a measure of uncertainty can suggest a precise result that should actually be interpreted quite differently if a large uncertainty measure was reported. Han et al. (2018) propose a variational Bayesian approach for estimating parameter uncertainty. Several approaches to get at uncertainty can be tested. A first is to leverage the fact that embeddings are usually underdetermined. This means that running the same model twice gives different results. The numerical estimates will not be the same. However, if the model is successful, the relative results will be the same, i.e. word will have the same "meaning" as measure through distance. Hence, one might be able to measure uncertainty through some form of bootstrapping The second method would be similar, also using bootstrapping but with replacement, changing the text slightly each time; but replacing words with similar words. Estimates with uncertainty will be measured against "objective measure of ideology". In the case of citizens, this can take the form of vote choice, or self-placement on a numerical axis. For pundits or politicians, it can be more complicated to find an objective measure.

Comparing estimates and their uncertainty to expert judgments, the idea is to produce estimates that generally make sense. For example, measuring the ideology of politicians in the Canadian parliament, members of the New Democrat Party should be to the left, members of the Liberal Party of Canada in the center, and members of the conservative party on the right, uncertainty should overlap between and within parties.

3.2   Distributional semantics

The previous subsection has introduced the idea of word embeddings: high dimensional vectors encoding the meaning of a word such that similar words have similar positions; are close to one another. In that section, word embeddings have been presented as one interesting way to produce some empirical ideological analysis; as the latest in a series of quantitative methods to study political behaviours and ideas. The present section shifts the focus on the linguistic properties and foundations of word embeddings: word embeddings as part of the research area called distributional semantics, or distributional approaches generally. Many of the ideas explored are largely inspired by Sahlgren (2008).

Sahlgren suggests distributional approaches rely on a set of assumption referred as the "distributional hypothesis". This hypothesis can be stated in various forms:

"words which are similar in meaning occur in similar contexts" (Rubenstein and Goodenough, 1965); "words with similar meaning will occur with similar neighbors if enough text material is available" (Schütze and Pedersen, 1995); "a representation that captures much of how words are used in natural context will capture much of what we mean by meaning" (Landauer and Dumais, 1997); and "words that occur in the same contexts tend to have similar meanings" (Pantel, 2005).

Sahlgren summarises: the distributional hypothesis is about "a correlation between distributional similarity and meaning similarity which allows us to utilize the former in order to estimate the latter".

To examine this "distributional hypothesis assumption", two questions can be asked. The first question is about the kind of distributional properties we are looking for (and in practice a question about algorithms). The second question asks "in what sense it is meaning that is conveyed by distributional patterns" (Sahlgren, 2008). The first question, about the choice of algorithm, will be covered in the

next section on methodology. What it means to get at meaning through word embeddings is explored in this section. This will be expanded in the dissertation.

The key idea discussed will be the common critique raised against distributional semantics and word embeddings: that they do not get at meaning because they only study the text, and not the world. Sahlgren's narrative is about the structuralist foundations of distributional approaches. Theories of meaning often ask if meaning is to be found in the mind (e.g. idealism), in the world (e.g. materialism, physicalism) or in the text (e.g. structuralism word embeddings). A model of meaning can situate it in anyone of mind, world or text, but situating it in the text has the pragmatic advantage of predictive power and of current large scale use. Work in political science usually mentions Firth and the distributional hypothesis (Rheault and Cochrane, 2020; Rodman, 2020; Rodriguez and Spirling, 2021; Rodriguez et al., 2021) but, usually stops at Firth. The point of this grounding in structuralism is that it makes explicit what can be shown with word embeddings and what cannot (i.e. word embeddings are not very useful from the perspective of a Correspondence Theory of Truth or a Verificationist Theory of Truth, but useful from the perspective of a Pragmatist or Coherentist theory of Truth; they make sense if you accept meaning holism and they don't if you swear by atomism about meaning). It is not necessarily evident why such a discussion belongs in political science, rather than in philosophy, but I think it is useful to reply to one frequent objection to text-as-data methods: that they don't work in practice. I think that sometimes they work well and sometimes they don't, but that mostly they are useful. In particular, I think that discussing some arguments in favor and against meaning holism can help alleviate problems associated with the intuition that words should have one-to-one relations with objects outside of the text.

3.3 Ideology and identity in Canada: changing meanings

This theoretical and methodological setup is necessary although secondary to the main objective: describing and measuring the changing meanings of key concepts used to talk about Canadian politics. Canada has a different history than those countries we usually think of as most similar, namely the United States and countries in Western Europe. Living next to the U.S. is like "sleeping with an elephant" said Pierre Trudeau, at the same time its European origins remain evident (Resnick, 2020).

Canada's identity has always been ambivalent, is it mono-national or pluri-national, does it have a culture, a "biculture", or "multicultures"; what about aboriginal cultures? (Wiseman, 2011) And, of course, is this changing, and how, over time? To cite a recent paper, describing this ambivalence (Dufresne et al., 2019):

> "In a nutshell, Canada is an immigrant nation founded, quite recently, on native land over an enormous territory by two linguistically distinct colonial groups – the English and French. If we add to this a hegemonic threat from a powerful American neighbour, a Quebec separatist movement, strong regional resentment in the West, and Aboriginal claims, Canada's difficulty in finding shared symbols to define the nation seems unsurprising."

Studying Canadian ideology and identity, we will focus on what it means to be Canadian. The question will be addressed both from the angle of identity (what it means culturally, origins, symbols) and ideology (does citizenship "involves" redistribution, openness to immigrants, justice, etc.). We identify four key concepts: citizenship, identity, nation and rights. A more extensive literature review will be included in the dissertation. For now, we can say that Canadian identity and nationhood was intrinsically linked to a "British connection" for close to 100 years following Canada's founding in 1867. This "British connection" of course was not part of the identity of those identifying as French Canadians. Starting in the 1960s, following similar developments in other Western countries, a diversification of identity claims began. In many ways, being Canadian was not necessarily about being French, English/British or Indigenous, but from a more plural and heterogeneous experience, some would say a post-national experience. Around the same time, Canada experienced a constitutional maturation with a Rights Revolution. Between 1931's Statute of Westminster, when Canada achieved independence from Britain and 1982's patriation when Canada obtained full sovereignty over it's constitution, there was the emergence of a "rights mentality": individual rights were becoming less abstract, and more something that Canadians could rely on in their day to day lives.

Where were politicians in the House of Commons standing on these developments? Empirically, around what period can we start to see a decline of "Britishness talks"? Similarly, can we observe the rise of "rights talk"? These questions are explored in the chapter on semantic change in the House of Commons.

Chapter 4

Methodology

The use of word embeddings in social scientific research is relatively recent. Kozlowski et al. (2019) are the first to use them to analyze the meaning of words in social discourse, producing a cultural analysis of discourse in a major social science journal. In political science, Rodman (2020) is the first to produce a thorough analysis of the use of diachronic embeddings to study semantic change. Rheault and Cochrane (2020) use word embeddings to study political discourse, but their focus in on ideological placement rather than on the use of embeddings for semantics or sematic change. Recent work also provides guidelines to the applied researcher. Rodriguez and Spirling (2021) focus on general guidelines to train embeddings efficiently: algorithms, parameters, pre-trained vs newly trained embeddings, etc. Their focus is on the general machinery of word embeddings rather than on semantic measurement per se. In another methodological paper, Rodriguez et al. (2021) focus on semantic change across groups in time. Their method, called à la Carte on Text (conText), extends a linear model, through multivariate regression, refitting pre-trained embeddings to local contexts. One objective of this methodological chapter is to present possible approaches to measuring how a word's meaning varies over circumstances; over time or other covariates. Five models in particular will be tested, the first four first presented by Rodman (2020), the fifth presented in Rodriguez et al. (2021): (1) naïve time series, (2) overlapping model (3) aligned model (4) chronologically trained model (5) à la Carte on Text. This chapter does not present any new results, it rather presents in detail methods used to produce embeddings in social science. An overview is presented in this proposal; more technical details will be given in the

dissertation.

4.1 Word embeddings

Using word embeddings, each word from a given text is mapped to a large numerical vector. The vector's dimension is determined by the researcher; usually between 100 and 500 Rodman (2020). The given vector represents the context in which the word is generally used. The model can be trained by predicting the word with its context, i.e. the other words around it, on a given window; this is the CBOW model. Alternatively, one can predict the context with the word using the skip-gram model. The GloVe method similarly models word and their context. The main difference is that it takes into account explicitly a matrix of co-occurrences. That is, word frequency enters in the modelling. In practice, these methods are similar, it is not entirely clear when and why they perform better, and GloVe is generally the preferred option in social science Rodriguez and Spirling (2021). Yet, in general, one of the most popular implementations is through the well-known gensim Python library. As a default, gensim uses CBOW. For instance, the two most recent use of word embeddings in political science use gensim (Rodman, 2020, Rheault and Cochrane (2020)).

4.2 doc2vec

There exists an extension to word2vec called doc2vec which embeds paragraphs or documents in the same vector space as words. The paragraph or the document can be thought of as an extra word; and items can be compared. Doc2vec is particularly useful to map the ideological position of a politician or a citizen in this vector space. An example of logical result would be that conservative politician's embedding will represent a point in space closer to words about the economy or a traditional lifestyle, while a liberal or progressive politician will be mapped to a point closer to words about social justice or redistribution.
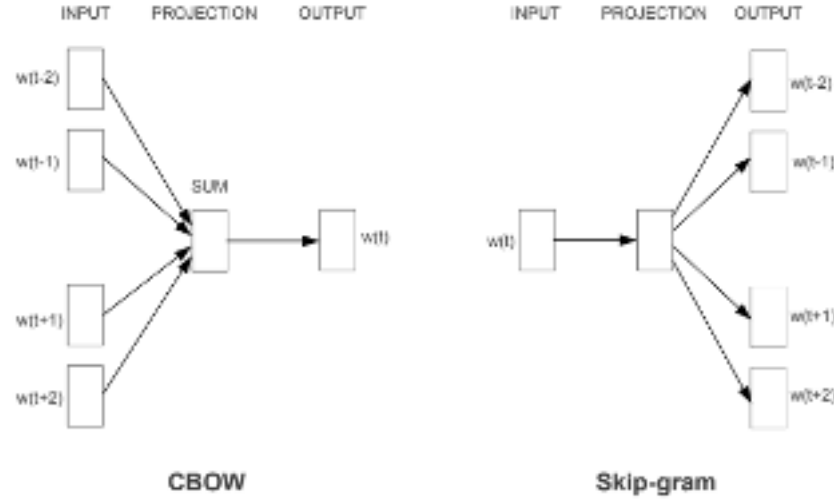
Figure 4.1: Continuous Bag of Words and Skip-gram architectures

4.3 Diachronic embeddings and semantic change

Measuring changing meaning implies that different words embeddings are calculated at different moments in time. In essence, this is what Rodman (2020) does. In her paper, she raises several issues. First, naturally, cut points will be arbitrary. This might or might not make sense. For instance, changing meaning on equality during the twentieth century might be a gradual process, but changing meaning on terror politics might be dramatic around events like September 11, 2011 or October 29, 1929. Arbitrary cut points therefore require careful thinking. Second, language often is polysemic or includes homonymies. For instance, "black" might be political in some contexts, and only about a color in others. Producing valid embeddings must take this into account. Third, and most importantly, there is the issue of non-comparability between models. Models position words in vector space. Distance can be calculated between words or documents. However, when looking at change in meaning over time, distances from words in many models are compared. This can be a problem as words across models are not aligned (Hamilton et al., 2016). Rodman provides the analogy to factor analysis. When ran on multiple datasets, it is likely that dimensions produced by factor analysis are similar, yet, factor order might change, or loadings might be on scales which are, in the absolute, non-comparable. For these reasons, while measuring changing word meaning over time, the analyst must proceed with caution.

Several methods exist to deal with non-comparability. Rodman identifies four. A naïve model, where the non-comparability problem is ignored, an overlapping time series models where differences are smoothed across time by including some text from the previous period, a chronologically trained where the model is initialized using results from the previous period and aligned time series where naïve models are ran, then aligned using a matrix alignment. Rodman's work suggests that a chronologically trained model is the best. The dissertation will validate these assumptions on new datasets.

Other methods exist to deal with semantic change. Long Short Term Memory units (LSTMs) on word embeddings corresponding to time-periods (Boukhaled et al., 2019). Contextualized word embeddings like BERT and ELMo can also model diachronic semantic change (Kutuzov, 2020). These methods are more complex to implement and yet to be really used in the social sciences. I will say a few words about them but they will not focus centrally in this dissertation.

4.4   Uncertainty

Measuring uncertainty when using word embeddings is non-trivial. Yet, when social science researcher make inference, they must accompany with uncertainty statements. The methodology section will discussion the question of uncertainty. The focus will be on bootstrapping methods (Rodman, 2020) and bayesian methods (Han et al., 2018; Lauretig, 2019)

Chapter 5

Capturing the dynamic meanings of political keywords in the Canadian House of Commons

Can we trace the changing meaning of keywords used to talk about politics? Using data from the Canadian Parliamentary debates dataset (Hansard) we model the changing meaning of the words 'nation', 'rights', 'citizenship' and 'identity'. This choice of words is only preliminary, the list will likely be expanded.

Rodman (2020) traces the meaning of the word equality in America between 1855 and 2016. In a first step, she uses embeddings to document the underlying structure of equality by asking: equality of what? With word embeddings, she shows how equality is discussed alongside race, gender, international relations and German affairs (for news related to World War II). We will provide a similar analysis for 'nation', 'rights', 'citizenship' and 'identity', in the Canadian context.

A working hypothesis is that prior to 1982 discussions of rights were less formal and more abstract, while after 1982, they were grounded in concrete examples. Similarly, debates on identity at the beginning of the twentieth century would have been centered on language, bilingualism or biculturalism, and only in recent years have moded to issues, of sex and gender, sexual minorities, race and multiculturalism.

Recent work (Rheault and Cochrane, 2020) has shown that word embeddings and party embeddings can capture a latent concept like ideology. The chapter tests whether it is possible to measure the evolving meaning of concepts central in the history of Canadian politics. Two examples can help illustrate.
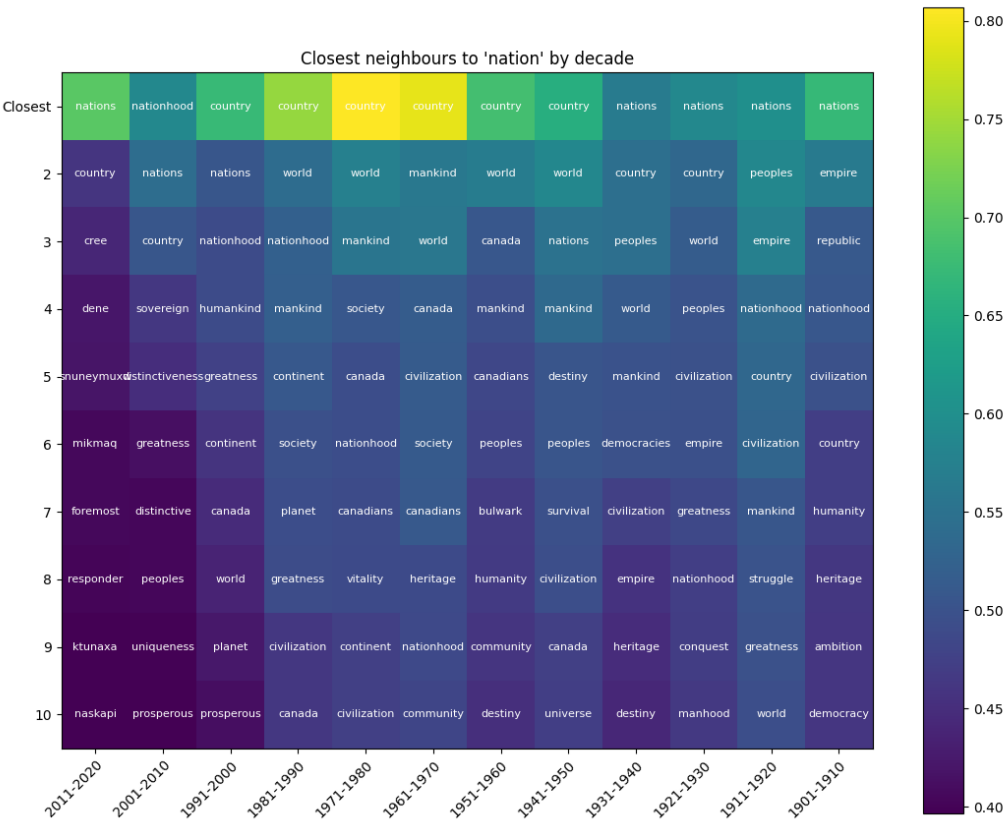

Figure 5.1: Co-sine similarity to 'nation' by decade

We use model data from the Hansard dataset to calculate a set of embeddings for each decade during the period 1901-2020. The model use is the default gensim implementation of CBOW with dimensions 300, window 6, using 15 iterations. This is equivalent to Rodman's "naïve model". Even with this simple specification, interesting results stand out. Figure 2 shows that during the first decade of the twentieth century, the closest words to 'nation' using co-sine similarity were 'nations', 'empire',

'republic', 'nationhood' and 'civilization'. By contrast during the most recent decade, the closest words to 'nation' were 'nations', 'country', 'cree', 'dene' and 'snuneymuxw'. During this most recent decade, six out of the 10 closest words to 'nation' were Indigenous nations. It is the first decade that at least one Indigenous related word appears. As we can see, words such as 'empire', 'struggle', 'conquest' or 'survival' are semantically close to 'nation' in early decades, but not in recent decades.
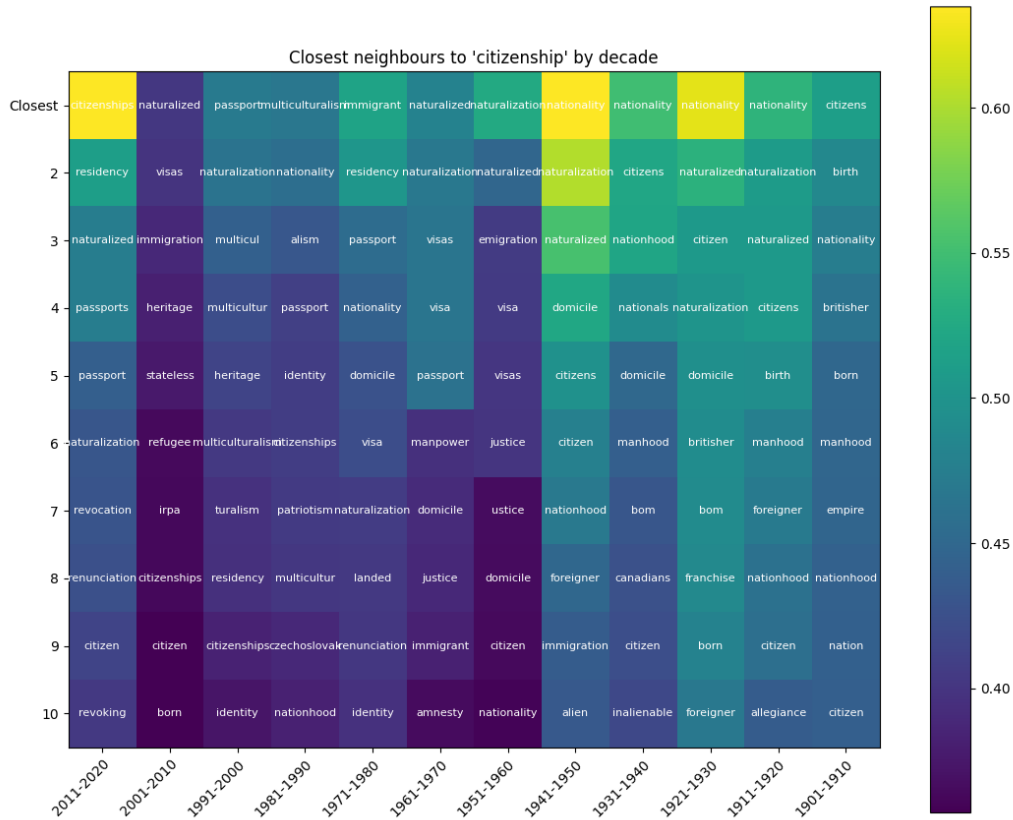

Figure 5.2: Co-sine similarity to 'citizenship' by decade

Figure 3 shows results that parallel those just discussed. Using the same word embeddings model, we visualize the closest neighbours to 'citizenship' by decade. During the first decade of the twentieth century, omitting the obvious 'citizens' the closest words by semantic proximity are 'birth', 'nationality', 'britisher', 'born', 'manhood', and 'empire'. In the most recent decade, the closest words are 'resi-

dency', 'naturalized', 'passports', 'passport', and 'naturalization'. It seems obvious that a slight shift of meaning for the term 'citizenship' has occurred, from an emphasis on 'birth' and 'britishness', and even 'maleness'. Mapping these changes, and when exactly they occurred historically, is one objective of this chapter.

Chapter 6

A model of author style with word embeddings

The previous chapter asked if the way political keywords were used over time changed. This chapter asks if backbenchers members of parliament use political keywords differently. This chapter is a replication and an extension of Huang et al. (2020). It is first a replication: the authors establish a model of Author "Style" in the UK House of Commons (1935-2018). They ask if backbenchers have "become less distinctive from one another in terms of their speech". Providing a model of "distinctiveness", they find no evidence of increased homogeneity, but they do find that senior backbenchers used to be less different, but that this is changing in recent years.

This chapter is, hence, first a replication, but it is also an extension. As the authors mention, their model is an implementation of a traditional Poisson distribution model (Mosteller and Wallace, 1963). The authors go on to say that the use of a word-embeddings approach could be used is the intention was to focus on feature representation (p.9). The objective is, hence, to extend this model of author style with the use of word embeddings to focus on the representation of meaning inside the model. Whereas they model author style using word frequencies, I will try to model it using a model averaging word vectors for users. I will further compare results from the Poisson model to results from the word embeddings model.

The data used will be the same as for the previous chapter: the Canadian Parliamentary dataset (Hansard). Strategic coding choices will closely follow the one adopted by Huang et al. (2020). For instance, "experience" can be calculated as the number of sessions served in parliament since first speech.

Chapter 7

Politicians and Twitter users talking about Canadian politics: how are they similar; how are they different?

This chapters focuses on semantic differences in the way politicians and users talk on twitter. The previous chapter using texts from the Hansard dataset used diachronic embeddings to study semantic change. The current chapter focuses on synchronic measures to capture differences in the way elites (politicians) and non-elites (users) express themselves.

The chapter asks two research questions. First, when looking at word-use, are there significant differences in the ways political keywords such as "nation", "citizenship", "rights" and "identity" are used by elites and non-elites. Using methods such as à la Carte on Text, can we make statistical statements about how these word's use varies. Second, the chapter will try to use this preliminary work to build a model of partisan associations. After having looked at the different meanings elites and non-elites assign to political keywords, can we use those to get at partisanship or ideology. More concretely, does the use of word vectors, representing meaning, increase the accuracy of a model of partisan affiliation or ideology, when compared to alternatives using bag of words approaches or approaches based on followers.

Following Wu et al. (2019), the idea is to learn associations between known partisan words and non-partisan words. These authors obtain partisan subspaces using analogy property of embeddings: a gender dimension, for instance can be calculated by substracting the embedding for "she" and "he", or "woman" and "man", or as showed in the section on chapter 4, by adding those differences. In the Canadian case, then how and where do words studied in chapters 4 and 5 appear within the conservative, liberal or progressive subspace. For instance, where is "bourgeois" in conservative subspace; in the liberal subspace. Where is "bureaucracy"?

Chapter 8

Case justification and contribution

Why Canada? The use of word embeddings to study meaning is recent (the idea is old, but efficient implementation is recent (2013)). Work in applied political science is very recent (less than five years old). As it is often the case in such a situation, almost all applied work focuses on the American case (Rodman, 2020, Kozlowski et al. (2019)). While these studies find interesting results for the study of culture and meaning, these results need to be tested outside the American case. Rodman, for instance, finds semantic change around the term "equality" that is specific to the United States. Other countries certainly saw the term evolve in a slightly different way. Furthermore, most of the work studying meaning with embeddings is in the English language. Canada is an interesting case as conversations about politics usually take place in English and French. While this dissertation will not explicitly rely on the analysis of French text, this certainly is the next step in the Canadian context. Hence, studying the Canadian case opens the door to a bilingual analysis, something more difficult in the United States or in Europe (with evident caveats in Europe like Belgium).

Next, as discussed in the third section of the literature review, the Canadian case is interesting due to the nature of it's shifting, ambivalent, political identity. Major cases studied (e.g. Britain, United States) are countries with a history of (more) explicit self-understanding. Thinking of the United States, people usually think of a mature world leader, of "Life, Liberty and the pursuit of Happiness", of the American nation, of the "extremes". The political science literature is more ambiguous as regards to what constitutes Canadian identity. Does this ambivalence still allow to capture meaningful

associations using word embeddings?

Finally, the main contribution of this dissertation is to apply word embeddings models to the study of Canadian politics. Word embeddings models offer a new and highly attractive way to study meaning. While it has been used by some social researchers in the last five years, my belief is that the full power of the method has not been utilized yet, and it's thorough application, especially focusing on diachronic embeddings, has the potential to shed new light on our understanding of how people talk about politics. Natural Language Processing (NLP) methods are perhaps the next (already started) technological revolution (Floridi and Chiriatti, 2020; Lakhotia et al., 2021). This dissertation is a humble contribution to the adavancement of their understanding in applied social science.

Chapter 9

Timeline

Item	Dates
Methodology chapter with examples using (1) naïve time series, (2) overlapping model (3) aligned model (4) chronologically trained model (5) à la carte on text	September – October 2021
Hansard chapter (analysis and write up)	October – December 2021
Backbencher chapter (analysis and write up)	December – January 2021-2022
Twitter chapter (analysis and write up)	February – April 2022
Conclusion/editing	May 2022
Peer-review process, edits, back and forth for feedback and criticism supervisor, committee, etc.	June 2022 -

Bibliography

Aldrich, J. H. and McKelvey, R. D. (1977). A method of scaling with applications to the 1968 and 1972 presidential elections. *American Political Science Review*, 71(1):111–130.

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political analysis*, 23(1):76–91.

Baudrillard, J. (1994). *Simulacra and Simulation*. University of Michigan press.

Bauer, P. C., Barberá, P., Ackermann, K., and Venetz, A. (2017). Is the left-right scale a valid measure of ideology? *Political Behavior*, 39(3):553–583.

Bonica, A. (2013). Ideology and interests in the political marketplace. *American Journal of Political Science*, 57(2):294–311.

Boukhaled, M., Fagard, B., and Poibeau, T. (2019). Modelling the semantic change dynamics using diachronic word embedding. In *11th International Conference on Agents and Artificial Intelligence (NLPinAI Special Session)*.

Boutyline, A. and Willer, R. (2017). The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political psychology*, 38(3):551–569.

Cochrane, C. (2010). Left/right ideology and canadian politics. *Canadian Journal of Political Science/Revue canadienne de science politique*, 43(4):1007–1008.

Cochrane, C. (2015). *Left and right: The small world of political ideas*, volume 3. McGill-Queen's Press-MQUP.

Debord, G. (2012). *Society of the Spectacle*. Bread and Circuses Publishing.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy*, 65(2):135–150.

Dufresne, Y., Tessier, C., Blanchet, A., and Temporão, M. (2019). The symbolic mosaic: an empirical typology of national symbolic webs and its effect on vote choice in canada. *National Identities*, 21(4):329–345.

Fairclough, N. (1992). Discourse and text: Linguistic and intertextual analysis within discourse analysis. *Discourse & society*, 3(2):193–217.

Firth, J. R. (1957). *Papers in linguistics, 1934-1951*. Oxford University Press.

Floridi, L. and Chiriatti, M. (2020). Gpt-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4):681–694.

Freeden, M. (2006). Ideology and political theory. *Journal of Political Ideologies*, 11(1):3–22.

Freeden, M. et al. (2003). *Ideology: A Very Short Introduction*, volume 95. Oxford University Press.

Gregg, P. M. and Banks, A. S. (1965). Dimensions of political systems: Factor analysis of a cross-polity survey. *American Political Science Review*, 59(3):602–614.

Hamilton, W. L., Leskovec, J., and Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. *arXiv preprint arXiv:1605.09096*.

Han, R., Gill, M., Spirling, A., and Cho, K. (2018). Conditional word embedding and hypothesis testing via bayes-by-backprop. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4890–4895.

Hare, C., Armstrong, D. A., Bakker, R., Carroll, R., and Poole, K. T. (2015). Using bayesian aldrich-mckelvey scaling to study citizens' ideological preferences and perceptions. *American Journal of Political Science*, 59(3):759–774.

Head, B. (1985). *Ideology and Social Science*. Springer.

Héroux-Legault, M. (2016). Substate variations in political values in canada. *Regional & Federal Studies*, 26(2):171–197.

Heywood, A. (2017). *Political Ideologies: An Introduction*. Macmillan International Higher Education.

Hirst, G., Riabinin, Y., and Graham, J. (2010). Party status as a confound in the automatic classification of political speech by ideology. In *Proceedings of the 10th International Conference on Statistical Analysis of Textual Data (JADT 2010)*, pages 731–742.

Hobsbawm, E. and Ranger, T. (2012). *The Invention of Tradition*. Cambridge University Press.

Homer-Dixon, T., Milkoreit, M., Mock, S. J., Schröder, T., and Thagard, P. (2014). The conceptual structure of social disputes: Cognitive-affective maps as a tool for conflict analysis and resolution. *Sage Open*, 4(1):2158244014526210.

Huang, L., Perry, P. O., and Spirling, A. (2020). A general model of author "style" with application to the uk house of commons, 1935–2018. *Political Analysis*, 28(3):412–434.

Jameson, F. (2020). *Allegory and Ideology*. Verso.

Jost, J. T., Federico, C. M., and Napier, J. L. (2013). Political ideologies and their social psychological functions. *The Oxford handbook of political ideologies*, pages 232–250.

Kellner, D. (1989). *Critical Theory, Marxism, and Modernity*. Johns Hopkins University Press.

Kennedy, E. (1979). " ideology" from destutt de tracy to marx. *Journal of the History of Ideas*, 40(3):353–368.

Kozlowski, A. C., Taddy, M., and Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, 84(5):905–949.

Kutuzov, A. (2020). Distributional word embeddings in modeling diachronic semantic change. *Ph.D. Dissertation: University of Oslo*.

Laclau, E. (2007). Emancipation (s).

Lakhotia, K., Kharitonov, E., Hsu, W.-N., Adi, Y., Polyak, A., Bolte, B., Nguyen, T.-A., Copet, J., Baevski, A., Mohamed, A., et al. (2021). Generative spoken language modeling from raw audio. *arXiv preprint arXiv:2102.01192*.

Landauer, T. K. and Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, 104(2):211.

Lauderdale, B. E. and Herzog, A. (2016). Measuring political positions from legislative speech. *Political Analysis*, 24(3):374–394.

Lauretig, A. M. (2019). Identification, interpretability, and bayesian word embeddings. *arXiv preprint arXiv:1904.01628*.

Laver, M., Benoit, K., and Garry, J. (2003). Extracting policy positions from political texts using words as data. *American political science review*, 97(2):311–331.

Martin, A. D. and Quinn, K. M. (2002). Dynamic ideal point estimation via markov chain monte carlo for the us supreme court, 1953–1999. *Political analysis*, 10(2):134–153.

Marx, K. and Engels, F. (1970). *The German Ideology*, volume 1. International Publishers Co.

Maynard, J. L. (2013). A map of the field of ideological analysis. *Journal of Political Ideologies*, 18(3):299–327.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.

Montigny, É. (2016). La fin des oui et des non au québec? un clivage en déclin. *L'Idée fédérale*, 7(1):1–10.

Mosteller, F. and Wallace, D. L. (1963). Inference in an authorship problem: A comparative study of discrimination methods applied to the authorship of the disputed federalist papers. *Journal of the American Statistical Association*, 58(302):275–309.

Mouffe, C. (1985). *Hegemony and socialist strategy: towards a radical democratic politics*. Verso.

Nafstad, H. E. and Blakar, R. M. (2012). Ideology and social psychology. *Social and Personality Psychology Compass*, 6(4):282–294.

Ozkirimli, U. (2017). *Theories of Nationalism: A Critical Introduction*. Palgrave.

Pantel, P. (2005). Inducing ontological co-occurrence vectors. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 125–132.

Pêcheux, M. (1975). *Language, Semantics and Ideology*. Springer.

Resnick, P. (2020). *The European roots of Canadian identity*. University of Toronto Press.

Rheault, L. and Cochrane, C. (2020). Word embeddings for the analysis of ideological placement in parliamentary corpora. *Political Analysis*, 28(1):112–133.

Rodman, E. (2020). A timely intervention: Tracking the changing meanings of political concepts with word vectors. *Political Analysis*, 28(1):87–111.

Rodriguez, P. and Spirling, A. (2021). Word embeddings: What works, what doesn't, and how to tell the difference for applied research.

Rodriguez, P. L., Spirling, A., and Stewart, B. M. (2021). Embedding regression: Models for context-specific description and inference. *Working paper*.

Rubenstein, H. and Goodenough, J. B. (1965). Contextual correlates of synonymy. *Communications of the ACM*, 8(10):627–633.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088):533–536.

Sahlgren, M. (2008). The distributional hypothesis. *Italian Journal of Disability Studies*, 20:33–53.

Schütze, H. and Pedersen, J. O. (1995). Information retrieval based on word senses.

Schwartz, M. A. (1967). *Public opinion and Canadian identity*. University of California Press.

Slapin, J. B. and Proksch, S.-O. (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science*, 52(3):705–722.

Temporão, M., Kerckhove, C. V., van der Linden, C., Dufresne, Y., and Hendrickx, J. M. (2018). Ideological scaling of social media users: a dynamic lexicon approach. *Political Analysis*, 26(4):457–473.

Thagard, P. (2012). Mapping minds across cultures. *Grounding social sciences in cognitive sciences*, pages 35–62.

Van Dijk, T. A. (2006). Ideology and discourse analysis. *Journal of political ideologies*, 11(2):115–140.

Wiseman, N. (2011). *In search of Canadian political culture.* UBC Press.

Wu, P. Y., Mebane Jr, W. R., Woods, L., Klaver, J., and Due, P. (2019). Partisan associations of twitter users based on their self-descriptions and word embeddings. In *2019 Annual Meeting of the American Political Science Association (APSA 2019).*

Žižek, S. (1989). *The Sublime Object of Ideology.* Verso.