

Title: Deconstructing Populist Evolution, A Temporal Word Embedding Approach

Author: Evelyn Simoni

Abstract: *This study introduces a novel approach to analyzing the discursive evolution in populist regimes, focusing on Hugo Chavez's anti-elitism. Employing word embeddings to examine 384 transcripts of Chavez's broadcasts, the research uncovers specific patterns across a 13-year period. The proposed technique addresses a gap in current populism studies by adopting a computational method to map discursive changes in a continuous framework. The findings reveal a transition in Chavez's anti-elitism, initially directed at the political class, developing into an anti-imperialist stance and ultimately targeting the economic elites. It also sheds light on the timing of these rhetorical shifts, providing evidence in favor of strategic over programmatic motives and illustrating how populists adjust their discursive frames to serve their shifting objectives. Beyond its contribution to the literature of populism in general, and Hugo Chavez in particular, the paper aims to open avenues for similar methodological analyses across other political texts.*

Introduction

The political science literature has made significant progress in classifying and quantifying populism through computational text analysis. Classical content analysis, which involves manually analyzing the content of a text, can be time-consuming, costly, and vulnerable to adverse selection. Early automated methods, such as those by Rooduijn and Pauwels (2011) and Bonikowski and Gidron (2016), relied on dictionary-based approaches with predetermined lists of words to capture populism intensity. However, these methods have their limitations, including language, regional, and time specificities. Furthermore, such approaches may not account for the variability in word meanings across different subjects and contexts, which is critical in textual analysis (Quinn et al., 2010).

Recent advancements in machine learning have accelerated quantitative research of political texts by offering tools that automatically scale and classify texts with minimal pre-set parameters (Rodman, 2020). But, the bulk of this research in populism studies has concentrated on cross-comparative analyses for which the outcome variable is a result of the subjects' (political actors) characteristics, such as the host's right-left ideology, political system, region, and institutional development (Hawkins et al., 2018; Di Cocco & Monechi, 2021). A gap still exists in adopting an approach that tracks the evolution of populist discourses in which the subject remains constant. Moreover, previous studies often adopt a unidimensional view, measuring degrees of populism, and thus overlook the complex, multifaceted nature of populist discourse.

The deconstruction of discourse in populist regimes is crucial to identifying its role not only when seizing power but also in exercising such power. Despite much discussion on the harmful link between populism and liberal democracy (Ruth-Lovell et al., 2019), the precise mechanisms by which populism may affect democratic outcomes remain to be tested. If we want populism to be able to talk to theories of democratic backsliding (Haggard & Kaufman, 2016; Galston, 2019), accountability (Ruth, 2018), and pluralism (Abts & Rummens, 2007), we need to be able to come up with a multidimensional measurement that tracks shifts of populists' discursive frames in a way that allows us to jointly test alternative explanations.

Furthermore, it is essential to develop an approach applicable across different contexts, including types of regimes, languages, regions, and historical periods.

In this paper, I build on recent advancements in natural language processing (NLP) and word embeddings. Drawing on the innovative approach of Rodriguez et al. (2023), I employ ALC (à la carte) word embeddings to detect semantic changes in populist discourse. This method allows us to capture patterns in anti-elitism rhetoric by analyzing changes in the frequency of words and their evolving meaning over time.

To create the list of words to analyze, I rely on a semi-supervised method by Chester (2023) that creates dictionaries based on a word embeddings matrix and a short list of seed words. One of the main advantages of this method is that it eliminates the need to use unreliable pre-constructed dictionaries, especially in non-English languages, and/or reduces the work and bias of constructing them manually.

I present an application of this method by examining Hugo Chavez's discourse evolution over his 14-year presidency. By analyzing 384 transcripts of the radio show *Aló Presidente*, I tap into Chavez's weekly direct communication with the Venezuelan public. This method is able to discern changes in the conceptualization of who 'the people' and 'the elites' represent. To do so, I categorize anti-elitism into three types. The first category is the political elite, representing the political class that governed Venezuela since the reestablishment of democracy in 1958. The second is the economic elite, comprising wealthy capitalists such as bankers, businessmen, owners of factories, and large landowners. The third category encompasses international elites, spearheaded by the United States and influenced by liberal multilateral institutions like the International Monetary Fund. This tripartite categorization facilitates a nuanced analysis of how Chavez's rhetoric evolved in targeting different groups as his presidency progressed and his political strategy adapted to domestic and international changes.

The findings discount populist discourse as a static expression rooted in predefined social categories (Roberts, 2022). It emphasizes the need to deconstruct populist discourse rather than merely focusing on the degree to which political actors are populists or not. Moreover, it shows empirically how populists in power construct and reconstruct who 'the people' and the 'evil elites' are, emphasizing the potential lasting role of populists as entrepreneurs of social identity (Reicher & Haslam, 2017).

This study opens new paths to studying the causes and effects of populist evolving discourse. I also discuss some limitations of the proposed method, its application to other types of research on political texts, and outline directions for future research.

Conceptualizing Populism

The term populism is highly contested and has been defined through various theoretical perspectives and disciplines, leading to broad generalizations about populism's universal features (Gidron & Bonikowski, 2013). One of the most common definitions used by political scientists describes populism as an ideology. Cas Mudde understands populism as a "thin-centered ideology that considers society to be ultimately separated into two homogenous and antagonistic groups, 'the pure people' versus 'the corrupt elite,' and

which argues that politics should be an expression of the *volonté générale* (general will) of the people." Rather than a complete ideology, populism is a thin ideology (Freeden, 1998): "employed as a framing device for other, more comprehensive ideologies" (Bonikowski et al., 2018). Thus, populism can take both left- and right-wing forms and can be found in established democracies and authoritarian regimes (Mudde, 2004).

A different perspective defines populism as a discursive style. Kazin (1995) argues that populism is not an ideology that attempts to capture core beliefs about society but rather a strategic way of making claims about politics. Instead of being an identity of political actors or a type of movement, populism is defined as a way of constructing the political (Laclau, 2005) and a type of political discourse (Deegan-Krause & Haughton, 2009). In other words, populism is not necessarily tied to any particular political actor but rather is a way of framing political issues. According to this discursive perspective, populism is a way of mobilizing support and shaping public opinion by presenting oneself as a champion of the "ordinary" people against the elites (de la Torre, 2017).

Finally, the social psychology literature has shed light on populist enduring appeal among voters by tapping into questions of social identity and intergroup behavior. The "us" versus "them" rhetoric of populist discourse resonates with humans' universal concern for status (Anderson et al., 2015). Within this perspective, populists can be understood as 'entrepreneurs' of social identity (Reicher & Haslam, 2017). By aligning themselves with ordinary voters and embodying the prototypical citizen (Mols & Jetten, 2016), they use existing or artificial threats to citizens' status to cut across social cleavages and recreate social categorizations in the form of in-group, namely the people, and out-group, the elites.

The Need to Deconstruct Populism

Traditionally, political science literature has been busier analyzing populists as outsiders, opposition parties, and protest movements than analyzing their activities in office (Barr, 2009). For instance, numerous studies examined the factors that contribute to the emergence of populism. Some focused on the role of economic factors, such as income inequality, economic insecurity, and changes in the labor market (Bornschieer, 2010; Rydgren, 2007). Others have examined the role of cultural and ideological factors, such as the appeal of nationalist or anti-establishment messages (Pop-Eleches, 2010). Still, other studies have examined the role of political factors, such as the performance and legitimacy of democratic institutions and the behavior of established political parties (Canovan, 1999). All of these findings, however, can be understood through Laclau's theory. Laclau (2005) poses that populists construct a superordinate category ("the people") from a heterogeneous group of people by combining various unabsorbed social demands into a unified chain of equivalence, creating two opposing camps within society. This new identity does not erase the distinctiveness of individual demands but establishes a common ground among them. The rise of populism is *necessarily* a product of the relationship between the people and the elites. Accordingly, using unidimensional scales for measuring the degree of populism of political actors is effective for empirically testing these theories concerning its emergence.

In this paper, I claim that it is crucial to also examine the role of populist discourse in wielding power. While studies linking populism and democratic backsliding abound, it is still contested how the rise of

populist actors to public office affects the quality of democratic regimes. On the one hand, scholars identified the negative consequences of populism on democracy's procedural and liberal elements (Chesterley & Roberti, 2018; Houle & Kenny, 2018; Ruth, 2018), while others show a positive impact on representation and participation (Mudde & Rovira Kaltwasser, 2017). These seemingly contradictory findings suggest that populists in power often try to dismantle horizontal accountability mechanisms while improving the vertical accountability component that makes democracy more participatory (Ruth-Lovell et al., 2019). If that is so, what role does populism discourse play in this phenomenon, if any?

To answer this question, this research aims to dissect populist rhetoric to determine how references to 'the elite' evolve and the implications of these shifts. This approach, which I term the deconstruction of populist discourse, seeks to uncover the strategies populist leaders employ to maintain relevance and appeal, thereby shedding light on the dynamics of populist regimes and the factors behind their success or demise.

Populists as Entrepreneurs of Social Identities

According to Roberts (2022), populism should not be seen as an expression arising from fixed or pre-existing social categories. Instead, the intended division of society as an in-group/out-group perspective is a dynamic process, a process in which populists actively construct a new, unified entity, a 'popular subject,' from a complex social fabric characterized by a multitude of subjects, fluid identities, and intersecting interests. In this sense, "the construction of a populist subject is a context-specific product of political conflict, mobilization, and discourse" (Roberts, 2022, p. 683).

The interplay between identity formation and dynamic populist discourse, as theorized by Roberts, forms the basis of my empirical research. By analyzing evolving patterns of anti-elitism, this study aims to test two hypotheses. The first hypothesis is that populists in power alter their depiction of 'the elite' over time. In other words, I do not test whether "the composition of 'the people' and 'the elite' *can* take a variety of different forms" (Roberts, 2022), but if such variation indeed occurs. Drawing from Hawkins' (2010) assertion that populism is characterized by a negative construction of 'the people', defined in opposition to an enemy, my analysis primarily focuses on anti-elitist rhetoric.

The second hypothesis examines whether the observed shifts in anti-elitist rhetoric are driven by ideological constraints or strategic motivations. According to Pappas (2019), populist regimes escalate polarization. But unlike ideological polarization, which is rooted in concrete policy issues, populism lacks a comprehensive ideology. As a result, it uses polarization deliberately as a strategy to amplify conflict over consensus. This implies that changes in rhetoric are not merely reflections of social dynamics and existing cleavages but are tactics to manipulate anti-elitist sentiments to reshape social identities to populists' advantage (Reicher & Haslam, 2017). The central premise of this hypothesis is that changes in anti-elitism are often less about emerging threats to populists' electoral base and more about addressing challenges to their own authority and power.

The questions posed in this paper are mostly exploratory. Despite being primarily descriptive, the literature still lacks an efficient method that is able to answer these questions with minimum bias and that

is applicable to populist regimes across different regions, languages, and periods. Using the presidency of Hugo Chavez in Venezuela, I build on recent advancements in computational processing of political texts to build a strategy that surpasses this identification problem. To my knowledge, this is the first study using computational methods over purely qualitative analysis to investigate such issues.

Temporal Word Embeddings

Recent advancements in machine learning and natural language processing allow researchers to devise new measurement tools to study populism (Hawkins et al., 2018; Di Cocco and Monechi, 2021). In the realm of machine learning, word embeddings is a new set of unsupervised techniques that represent each word in a given vocabulary as a vector of real numbers in a multidimensional space. Word embeddings predict a target word as a function of the other words that appear within a small window of that target word in the corpus. The distributional hypothesis is the backbone assumption that drives word vectorization models (Turney and Pantel, 2010). This hypothesis posits that we can discern semantic meaning by systematically collecting word collocations, regardless of other contexts or grammatical structures (Rodman, 2020). The power of the word embeddings approach lies in its generalizability. It can be applied to any language or individual corpora within a given language; it requires no prior input of dictionaries or grammatical structures. It deduces the meanings of words based on considering the collocations of all words that occur in a given corpus of texts (Rodman, 2020).

The applicability of word embeddings in social science relies on the concept that words with similar semantics are geometrically closer to each other in the multidimensional space. For example, in the Chavez corpora, we would expect the word ‘fascists’ to be closer to the word ‘traitors’ than to the word ‘allies’. This characteristic of word embeddings allows the researcher to measure changes in meanings by analyzing “nearest neighbors” (words closer together in space) of specific words. Compared to other popular methods in text-as-data analysis, such as topic modeling, word embeddings allow researchers to extract meaning from the window where the target word occurs rather than from the entire document.

Previous work that uses word embeddings in social science includes Kozlowski et al. (2018) exploring cultural dimensions through word embedding analysis, Rheault and Cochrane (2019) examining ideological positions in parliamentary texts, and, more recently, Bellodi (2023) who presents a dynamic approach to measuring bureaucratic reputation.

Word embeddings will be instrumental in deconstructing populist discourse in two ways. Firstly, they enable us to discern the nuanced shifts in the meaning of populist terms that, while individually ambiguous, gain specificity within broader textual frameworks. For example, 'elites' may signify political figures, affluent citizens, or media entities, depending on the speech context. By analyzing how these terms are embedded in various contexts, we can track the evolution of populist rhetoric and its targets over time.

However, using word embeddings to understand the evolution of meaning in a continuous time frame creates several challenges. For instance, if we train a set of word embeddings on the entire of Chavez’s available transcripts from the year 1999 until 2012, we will end up having only one meaning for each target word, which fails to achieve our objective of capturing the different meaning of such word over

time. One could simply divide a single corpus into time intervals to model distinct semantic universes and track changes in word vector relationships over time. Yet, this method introduces validity problems, most notably 'spatial noncomparability'. Temporal analysis compares cosine similarities (vector closeness) across different models. However, word embeddings trained from stochastic processes in different periods result in words being placed in non-aligned spaces, making direct comparison difficult (Hamilton et al. 2016). This issue also complicates comparing different runs on the same corpus, as nearest neighbors might stay the same, but their coordinates shift (Kulkarni et al. 2015). This aligns with the core concept of temporal semantic analysis: word meanings change over time. Therefore, it is expected that both the meanings of words and the structure of the space they occupy will shift over time (Rodman, 2020).

In this paper, I employ Rodriguez et al. (2023) 'a La Carte' (ALC) embedding regression model to identify semantic changes in populist discourse over time. Their approach is grounded in Arora et al. (2015) concept of documents being a 'random walk' in a discourse space, where a sequence of words is more likely to form if each word is closer to the next one in the embedding space. Given the random walk theory and a set of word embeddings trained on whole corpora (pre-trained word embeddings), one could average the embeddings (vectors) of the words surrounding the target word in document A to obtain a context-specific word embedding of the target word for document A. You could then repeat such a process for each document in your corpora.

Unfortunately, as Khodak et al. (2018) demonstrate, averaging embeddings tends to excessively emphasize common elements linked to frequent words, such as stop words. The solution is multiplying the averages obtained by a square matrix to reduce the influence of prevalent yet uninformative words. Rodriguez et al. (2023) estimate this 'transformation matrix' from the pre-trained word embeddings by using a modified linear regression model, which changes the weight (importance) of the words according to how often each word appears in the corpus.

$$v_w = Au_w$$

To obtain the context-specific word embeddings we estimate equation 1 (Rodriguez et al., 2023), where v_w is a single instance of the target word, A is the square transformation matrix, and u_w is the average of the pre-trained embeddings of the words surrounding the target word in that specific instance. The main contribution of this approach is that it "allows us to infer the meaning of words that rarely occurred in a small corpus without having to build embeddings for those rare words in that small corpus: we can 'borrow' and transform the embeddings from another source." (Rodriguez et al., 2023).

A Semi-supervised Method to Creating an Anti-Elitism Dictionary

With the ALC year-embeddings we are still left with the non-trivial task of choosing the keywords for each category of anti-elitism, that is, of creating the dictionary. Pre-constructed dictionaries can quickly become outdated, failing to incorporate new words, phrases, or shifts in usage. Additionally, these dictionaries are often built for general use. They may not suit specific research or context typical of smaller corpus and might not account for regional variations or cultural specifics in language use (Chester, 2023). Finally, these dictionaries are only available for a small subset of languages. Conversely, using available methods to construct your dictionary has its limitations. Manually creating a dictionary or

training a model to create a dictionary through a machine learning algorithm requires substantial human effort, either to identify conceptually related terms or manually code documents (Chester, 2023). Moreover, adverse selection might compromise the validity of the dictionary if crucial keywords are omitted. The absence of commonly used words when measuring anti-elitism intensity can result in a skewed understanding of the text.

In this scenario, word embeddings allow us to construct a dictionary that minimizes human error and significantly reduces the cost. I employ a recently developed semi-supervised method that creates dictionaries for specific tasks and specific vocabularies with minimum input from the researcher. The algorithm by Chester (2023), called `conclust`, uses embedding methods to generate a list of keywords based on a small number of seed words. Using word embeddings to create dictionaries has been prevalent in the realm of sentiment analysis, where the aim is to categorize words based on their positive or negative connotations, in a given polarized universe (Rice & Zorn, 2021). Yet, the use of word embeddings for broader conceptual tasks in text analysis remains limited (Chester, 2023).

Conclust requires a cosine similarity matrix and a set of seed words. The similarity matrix measures the geometric distance between pairs of word vectors in an embedding matrix. By calculating the cosine similarity, which evaluates the cosine of the angle between two vectors, the matrix determines how semantically similar each word is to another for all possible pairs of words in a given vocabulary. Seed words refer to a collection of words supplied by the user that reflect the targeted concept.

Algorithm 1: `conclust`

```

Input: Seed words:  $S$ ; Embedding model  $M$ ; Size threshold:  $n$ ;
Similarity threshold:  $t$ 
Result: Keyword set:  $K$ 
 $K = S$ ;
while  $|K| \geq n$  do
     $\bar{m} = \forall m \in M \max(\text{sim}(K, m))$ ;
    if  $\text{mean}(\text{sim}(K, \bar{m})) \geq t$  then
         $K = K \cup \bar{m}$ ;
    else
        break;
    end
end

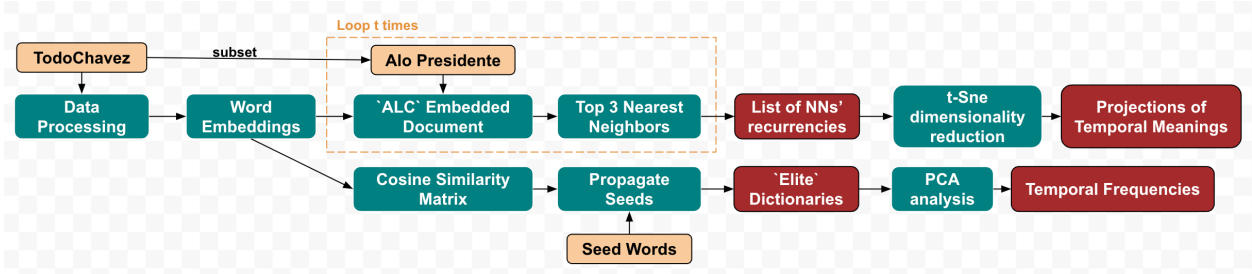
```

Starting with the user-provided seed words, the algorithm iteratively searches for a word with the highest average similarity to the seed words within the similarity matrix as long as this similarity meets a predefined threshold. If the word's average similarity is above this threshold, it is added to the keyword set. This process repeats itself until the keyword set reaches a specific size or no more words meet the similarity requirement. The outcome is a refined set of keywords closely related to the original seeds (Chester, 2023).

Application: Embedding `TodoChavez` and Creating an Anti-elitism Dictionary

The `TodoChavez` corpora is a collection of 5613 transcripts of every recorded speech, interview, and broadcast of Hugo Chavez as President of Venezuela. The 'Alo Presidente' is a subset of such a collection. It was a weekly television program hosted by Hugo Chávez during his tenure in office. The program aired

on Sundays from 1999 until 2012 and was known for its lengthy, unscripted format, allowing Chávez to address various topics and directly interact with the Venezuelan public. From the first episode in May 1999 until the last episode in January 2012, there are 384 available transcripts.



In this case, and given the size of the dataset, I created the word embeddings from the TodoChavez corpora rather than using pre-trained word embeddings. The advantage of training the embeddings in these transcripts is that the algorithm is trained in the particular idiosyncrasy, slang, and vocabulary of Hugo Chavez’s discourse. I train the word embeddings using the word2vec algorithm, a commonly used model that learns distributed representations of words by training a neural network. Its objective is to predict the likelihood of a word occurring in a specific context given its neighboring words (Rodman, 2020).

Without the availability of such large corpora, researchers can utilize pre-trained embeddings developed from extensive corpora containing billions of tokens. These pre-trained models are readily available in a multitude of languages. Researchers can then refine and fine-tune these pre-trained vectors using a smaller, more specific corpus that aligns with their research focus. This fine-tuning process allows the embeddings to adapt to the nuances and specificities of the smaller dataset, thus becoming more relevant and accurate for the researcher's particular area of interest (Howard and Ruder 2018; Rodman 2020).

The second step is to create the dictionaries for the three previously defined categories: political, economic, and international anti-elitism. Using between three and five seed words for each category, I propagate the seeds using the cosine similarity matrix of the pre-trained word embeddings. I set a maximum of 25 words for each list and a .4 similarity threshold. This produces a dictionary of words that are highly semantically similar to each other. I only include words mutually exclusive to each category and avoid vague expressions that might represent multiple types of anti-elitism. I also drop words that appear less than 100 times within the Alo Presidente corpus.

The next step is to subset the corpora and tokenize only the ‘Alo Presidente’ corpus. I conduct the typical process used in social sciences to process the raw text, including converting the text to lowercase and removing punctuation, special characters, and numbers. The next step is lemmatization, reducing words to their base or dictionary form, and grouping different words into one term. For instance, ‘united’ and ‘states’ were lemmatized to ‘united_states.’ Finally, I removed common words (stop words) that do not add meaning to the text.

Using the constructed anti-elitism dictionary and the processed 'Aló Presidente' corpus, I examine two distinct dimensions: intensity and meaning. The intensity is gauged by the frequency of anti-elitist terms

within the corpus, providing quantitative data on the extent of anti-elitist rhetoric employed. The meaning, on the other hand, is assessed by analyzing the context in which these terms are used, offering qualitative insights into the nuances of the discourse. This dual analysis enables a deeper understanding of both the overt and the subtle ways in which anti-elitist sentiment is communicated and evolved throughout the corpus.

Application: Measuring Intensity

To measure the intensity of each anti-elitism category, I calculate each keyword's quarterly frequencies using the Term Frequency (TF) metric. The TF is the number of times the word appears in a document divided by the total word count in that document, which is the proportion of usage of a term in a given quarter. I create the composite index by adding up the TF of each keyword.

I plot the normalized TF of all words and the composite index. I also plot the keywords in the quarter in which their normalized TF is equal to 1.

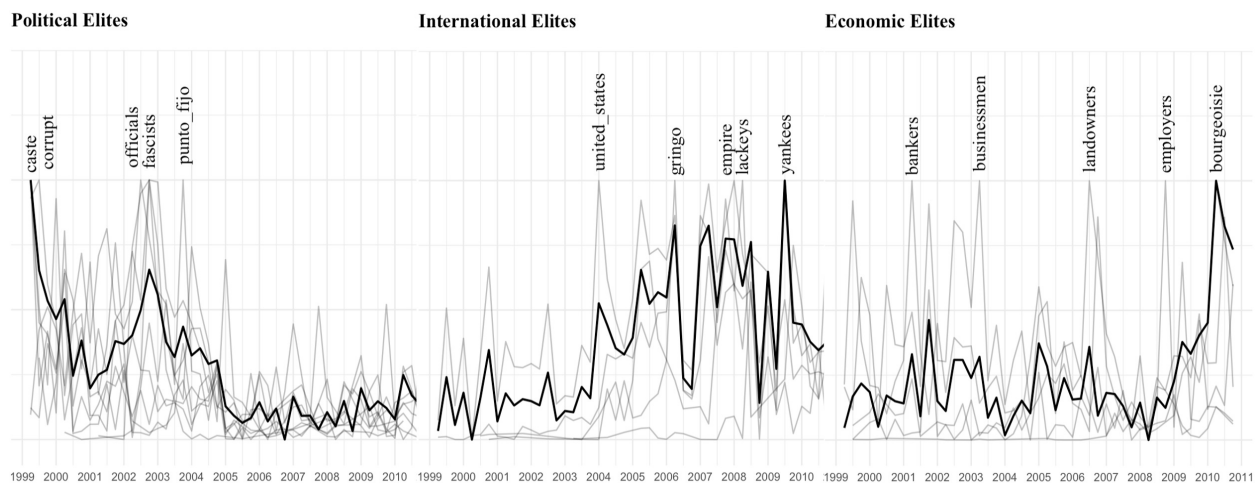


Figure 1. Normalized Index and Individual Words

Previous research on the role of populist discourse in Chavez's rise in the 1990s has been linked to the widespread dissatisfaction among voters with the political system during the Punto-Fijo Pact (Stavrakakis et al., 2016). This period, which began in 1958, denotes the ruling of Venezuela by the two main parties, Acción Democrática and Partido Socialcristiano. Therefore, it is not surprising that the 'political elite' index is more pronounced in the early years of his tenure. Despite Chavez's ascension to office, the old political class still represented a threat to Chavismo's bases and symbolized the out-group that Chavez's ideals were actively combating. This threat was not only present when the traditional political parties still held the majority of seats in both chambers of Congress but became evident after the failed coup against Chavez and the national strike in 2002.

Figure 1 also shows that Hugo Chavez did not escalate his international anti-elitism stance following the April 2002 failed coup attempt, as previously suggested by qualitative studies on Chavismo (Hawkins,

2010). Instead, he used words such as `golpistas` and `fascists` to represent the enemies of the people during this time. The rise in `United States` usage began two years later. Furthermore, he did not widely adopt the term `empire,` which carries a more negative connotation and ideological weight when referring to international elites, until three years later.

Finally, the findings suggest that it was not until the later period of his presidency that the economic anti-elitism rhetoric permeated into Chavez's discourse, mostly pushed forward by the heavy usage of words such as `patrons,` `capitalists,` and, most notably, `bourgeoise.`

Application: Tracking Meaning

To measure changes in meaning over time, I first construct the ALC year-embeddings. For 1999-2012, I searched for the top three nearest neighbors based on their closeness to the keyword and constructed a measurement of self-cosine similarity. Self-cosine similarity refers to the degree of similarity of a word to itself across different periods. Essentially, it compares the vector representation of a word in one slice of time with its vector in the following slice of time. By assessing self-cosine similarity, I can measure how the semantic representation of a word has shifted or remained stable over time.

I plot the trajectory of those words whose meaning changed significantly, together with the top three nearest neighbors for each year. To do so, I employ the t-distributed Stochastic Neighbor Embedding (t-SNE) to reduce the dimensionality of the embeddings (Yao et al., 2018). The goal of this task extends beyond merely tracing the evolution of Chavez's anti-elitist rhetoric; it also validates that the keywords utilized to gauge intensity consistently retain the intended meaning throughout the period studied.

In Figure 2, we observe the temporal evolution of the terms 'elite' and 'opposition.' The three closest neighbors exhibit stability during specific years but significant variations in others—for the word elite, the initial years presented a less confrontational tone, with Chavez referring to the goal of eradicating elitist culture upon assuming the presidency. However, starting in 2001, the word elite began to gravitate closer to terms associated with the political class, such as caste, cupula, oligarchy, and corrupts. In 2007, we also witnessed the emergence of terms related to the wealthy and the economic elite, notably the bourgeoisie, mafia, and bankers.

A similar trend unfolded with the term opposition. In its early usage, the word had vague connotations, but after the second year, opposition became interchangeable with opposing political parties. Following the elections in the 2000s, terms like defeated were employed, and after the failed coup in 2002, 'gone crazy' entered the lexicon. Meanwhile, other political terms, such as coup mongers, leadership, the right, and candidate, remained closely associated. However, from 2008 to 2010, bourgeoisie began to appear as the primary term, as words related to the political elites dwindled in prominence.

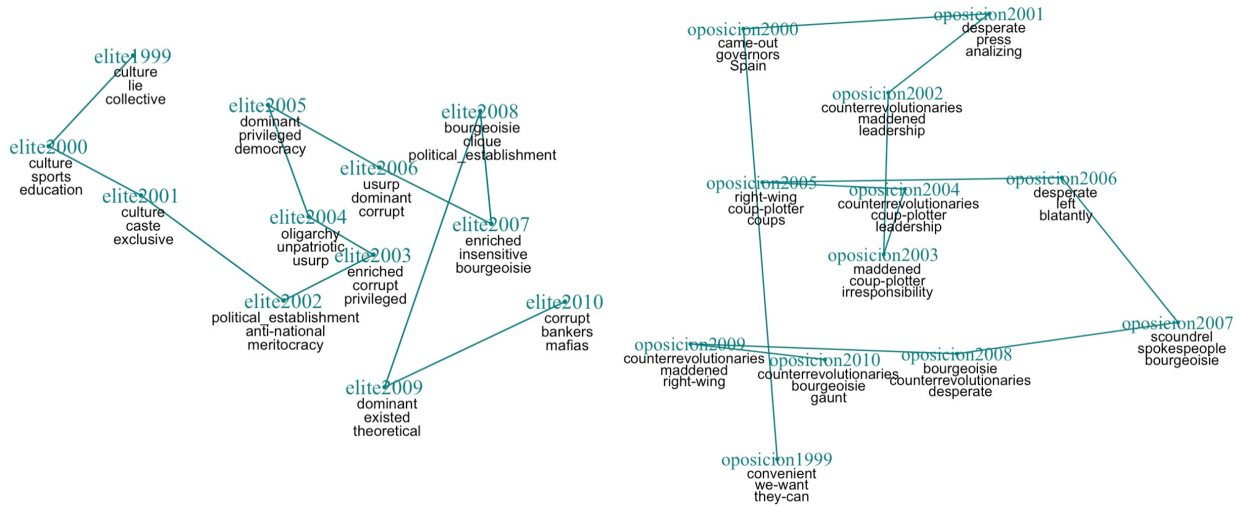


Figure 2. 2-D t-SNE Projection of Temporal Embeddings of 'Elite' and 'Opposition'

Despite the increase in the frequency of the international anti-elitism index, as demonstrated in the previous section, it is noteworthy that keywords associated with international elites do not feature prominently in the evolving semantics of 'elite' and 'opposition.' To explore this further, I examine whether words like 'imperialism' and 'yanquis' indeed carry the negative connotations and anti-elitist rhetoric that I hypothesized.

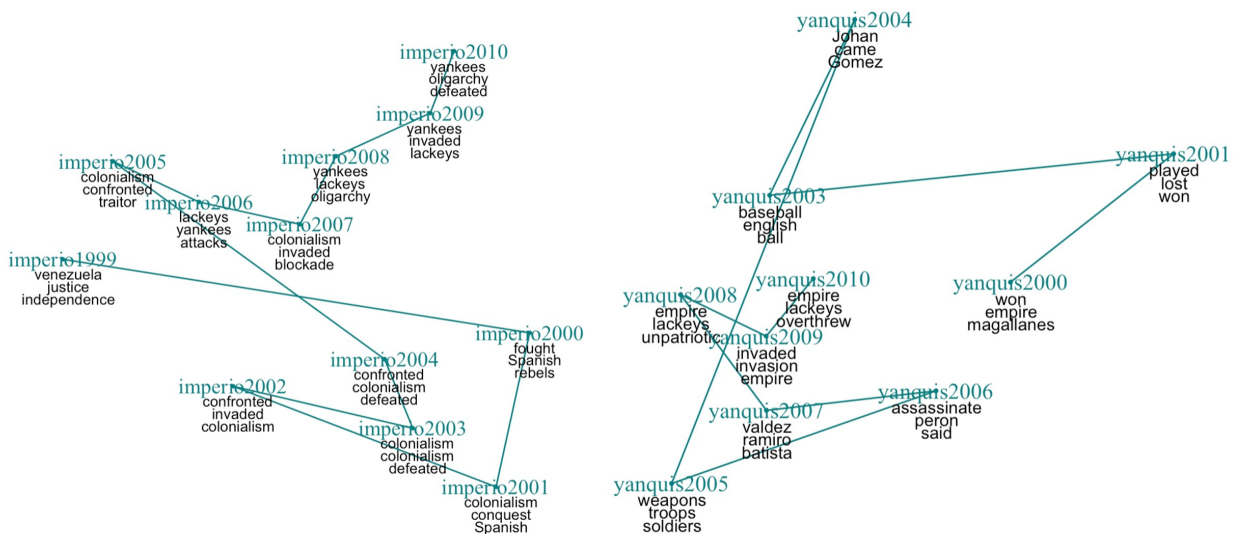


Figure 3. 2-D t-SNE Projection of Temporal Embeddings of 'Empire' and 'Yanquis'

As shown in Figure 3, both words, 'empire' and 'yanquis,' exhibited markedly different meanings at the outset of the period but underwent significant shifts between 2005 and 2006, coming to denote what we initially associate with international anti-elitism. In the initial years, when Chavez used the term 'empire,' he primarily referred to the Spanish empire and the colonial era, explicitly discussing it within Venezuela's historical context. However, from 2006 onward, words like 'yanquis' and 'lackeys' gained prominence. Interestingly, 'oligarchy' also appeared to have closer associations with 'empire' between

2008 and 2010, as Chavez intended to associate the local elites with its ties to imperialism. Likewise, the term 'yanquis' initially had connotations related to baseball, using words like 'won,' 'lost,' and 'ball.' However, a notable transformation occurred in 2005 when military terminology began to take prominence, followed by the emergence of imperialist terms in 2008.

Discussion

Studies exploring Hugo Chavez's rise generally agree that the Venezuelan president accessed power by making salient an existing social cleavage. This cleavage was not a product of class relationships, as it is often thought, but a product of the exclusion of an important share of the population, that of rural workers and urban informal workers, from the existing political system during the Punto-Fijo Pact (Hawkins, 2010; Mazzuca, 2013). Hence, Chavez's route to the presidency was marked by targeting the political class rather than the wealthy elites (Hawkins, 2010). Chavez placed the social cleavage in a Manichean dichotomy, characteristic of populist discourse. This highly divisive construct did not allow any nuance or dissent:

*“Everyone here has to know which side of the battle they are on. We are in times of apocalypse (...), where **there are no intermediate positions**: either you are a patriot, or you are against the country. Either you are with Yes or you are with No. Or you are with God, and whoever is with God is with Yes.”* Hugo Chavez in *Aló Presidente*, November 1999, in Weyland (2001).

The findings of the analysis on the intensity and meaning of anti-elitism dictionaries throughout Chavez's tenure support my theoretical expectations. Populist discourse is not static. Upon taking office, Chavez widened the scope of his criticism beyond the political elite—initially the adversaries of his primary voter base. Progressively, Chavez's rhetoric transitioned from a political anti-elitism to an anti-imperialist stance and later to an economic anti-elitism, indicating a versatile adaptation of his populist discourse.

Prior qualitative research on Chavez's populism has already highlighted the shifting portrayal of the antagonistic minority throughout his presidency. Hawkins (2010) stresses that Chavez initially targeted the old political elite represented by the two traditional parties and that after the tumultuous events of the 2002 coup attempt and subsequent national strike, his discourse shifted to include the idea of an international conspiracy led by the United States as the driving force behind these groups.

This study contributes to the existing literature in two significant ways. Firstly, it provides empirical evidence of a previously underexplored transition in Chavez's rhetoric from anti-imperialism to targeting domestic capitalism. Secondly, the study sheds light on the timing of Chavez's rhetorical shifts, offering a deeper understanding of his strategic versus programmatic motivations (Roberts, 2022).

If anti-elitism shifts were a function of ideological motivations, we should expect the rhetorical shifts to align with either new threats to the Chavismo voter base or to correspond with any concrete policy initiatives. In our case study, we would expect Chavez to adopt an anti-imperialist stance immediately following the 2002 coup attempt, as Hawkins claims. Yet, it wasn't until 2005, almost three years later, that such a stance became prominent. This is perceptible in the increased usage of terms associated with international elites and in the shifting meanings of words like 'empire' and 'yanquis.'

Relatedly, Chavez's use of class war language and his targeting of the bourgeoisie did not peak until 2010, contradicting the expectation that after the 2006 reelections his anti-elitism would reflect the proclaimed pivot to the "21st-century socialism." This ideological shift towards the radical left was symbolically marked by the addition of 'socialism' to the revolutionary slogan "homeland or death," which evolved into "*Patria, Socialismo, o Muerte.*" The updated motto even translated into tangible policy actions in 2007, such as nationalizing key industries, including oil, and the shutdown of RCTV, the oldest and largest private television network (Carroll, 2014).

This pattern suggests that Chavez's anti-elitism was not simply a veneer for programmatic changes rooted in ideological considerations. Instead, they were an integral part of a strategic approach, dynamically intertwined with the political and ideological landscape he was navigating. Thus reinforcing the conception of populism as a strategic tool rather than an ideological commitment.

Theoretical Propositions

In pondering the discursive changes occurring around 2005 and again in 2009, I propose theories, yet untested, that offer a perspective on Chavez's strategic shifts.

Initially, the foreign involvement in the 2002 coup appeared as an opportunity for Chavez to shift the narrative towards imperialism. But it was not until Chavez won the 2004 referendum by 59 percent of votes that Chavismo consolidated as the established political class in Venezuela, leaving a feeble opposition in the domestic arena, that had already tried and failed every possible avenue, illegal and legal, to stop Chavez. In short, it became clear that Chavez could not continue to target the old political class because Chavismo became synonymous with the political class.

Initially, foreign intervention in the 2002 coup presented an opportunity for Chavez to shift the narrative, emphasizing imperialism. However, the consolidation of Chavismo as Venezuela's dominant political force only materialized post-2004 referendum, with Chavez securing 59 percent of votes. This victory rendered the political opposition feeble, having exhausted all means, both legal and illicit, to thwart Chavez (Carroll, 2014). In summary, Chavez could not continue to target the old political class because Chavismo became synonymous with the political class.

On the international side, the price of oil began to surge in late 2003 to its peak in 2008. The fact that the United States relied on Venezuela's oil supply allowed Chavez to present himself as a "champion of the poor" against an evil representation of the United States without significant material consequences. Moreover, the rise of left-leaning leaders across South America, known as the Pink Tide movement, allowed Chavez to foster a sense of regional unity while reshaping his supporters' social identity, enclaved in a shared regional struggle against 'the north.' In January 2007, Chavez invited the Minister of Information of the Syrian Arab Republic to the show *Aló Presidente*:

"We will always be with Syria and with the struggles of the Arab people, and we do it out of a moral commitment to human beings, and to the just struggles of human beings, of the brotherly peoples of the world, against imperialism, against capitalism, against the hegemony of the

imperialist powers and their lackeys and their imperialist arms, in the military, in the economic, in the territorial. And it is a fight that unites us and will unite us forever.”

In 2009-2010, the second shift seems to stem from an economic crisis triggered by the dramatic fall in oil prices from over \$147 per barrel in July 2008 to \$33 in February 2009, and an energy crisis due to severe droughts affecting hydroelectric power (Carroll, 2014). Confronted with escalating public dissatisfaction and impending elections, Chavez pivoted his anti-elitism rhetoric to blame capitalists. More importantly, he aimed to increase the members of his coalition to all workers (see appendix B for the frequency of the word workers). In an *Aló Presidente* episode in January 2010, he stressed that, even after 11 years in power, the revolution was just beginning:

*“The rich got used to making full use and taking advantage of the resources of the people (...) they managed and took control of the country's main resources, including water, energy, and money. Capitalism is alive and well here. This is just the beginning, comrades. (...) Socialism is being born, while **capitalism is a monster with tentacles and networks everywhere**. But we will defeat capitalism and establish Bolivarian socialism in Venezuela.”*

These proposed theories on the causes for Chavez’ evolving anti-elitism intend to mirror the dynamic and strategic essence of populist framing, illustrating that while populists may capitalize on genuine threats to their base, they predominantly use these as opportunities for furthering their political agenda, employing populism instrumentally.

Limitations and Future Research

While the temporal approach to deconstructing and measuring the evolution of populism marks a significant advancement, it is important to acknowledge this paper's limitations in constructing a causality test to examine how shifts in populist discourse affect voter perceptions. Although this research is able to capture the strategic nature of populist framing in populist regimes, we still need to research if and when such a strategy is successful. Hence, further research is needed to establish a direct causal link between the change in populist rhetoric and its impact on the electorate's views and attitudes.

This is especially true in light of recent concerns regarding democratic backsliding. Critical questions emerge to be answered, such as the interplay between populist discourse and voters' perception of populists as legitimate representations of "the will of the people" and how this relates to voters' attitudes towards horizontal vis-a-vis vertical accountability. For instance, why would voters seek to constrain a populist leader they perceived to represent their group? And if minorities are viewed as the malign elite, what would be the incentive to advocate for a more pluralistic society? To unravel these complexities, a more sophisticated methodological framework, including additional data sources, is necessary to analyze populist broader democratic implications.

Finally, there is a clear need for additional research employing a temporal approach that compares multiple cases of populist regimes. Such comparative analysis would allow for a broader understanding of the generalizability and distinct characteristics of populist discourses over time. It would also shed light

on these regimes' varying impacts on democratic processes and structures across different political and cultural contexts. This expanded scope of study could provide valuable insights into the mechanisms of populist rhetoric and its effects on the political landscape.

Conclusion

In this paper, I highlight the need to make a distinction between the role of populism to access power from its role in the exercise of such power. In particular, in deconstructing populist discourse by looking at anti-elitism rhetoric. To my knowledge, existing empirical studies on populism were missing an efficient and cost-effective approach to deconstructing and tracing the evolution of populist discourse. Here, word embeddings, the latest innovation in natural language processing, offer a solution for this analytical task. This approach enables researchers to construct specialized dictionaries with minimum input while facilitating the measurement of rhetorical intensity and the analysis of semantic shifts across keywords. This approach promises a more nuanced understanding of the dynamic nature of populism.

The second contribution of this paper is substantive. It advances the understanding of populism, providing insights into the strategic nature of discursive shifts, specifically focusing on how populists redefine social identities and cleavages to retain power amidst shifting political challenges. This insight challenges the simplified view of populism as a byproduct of social and economic processes, emphasizing instead the deliberate and strategic construction of social narratives by populist leaders.

References

- Abts, K., & Rummens, S. (2007). Populism versus democracy. *Political studies*, 55(2), 405-424.
- Anderson, C., Hildreth, J. A. D., & Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological bulletin*, 141(3), 574.
- Arora, S., Li, Y., Liang, Y., Ma, T., & Risteski, A. (2015). Random walks on context spaces: Towards an explanation of the mysteries of semantic word embeddings. arXiv preprint arXiv:1502.03520, 385-399.
- Barr, R. R. (2009). Populists, Outsiders and Anti-Establishment Politics. *Party Politics*, Vol. 15 (1), pp. 29–48.
- Bellodi, L. (2023). A dynamic measure of bureaucratic reputation: New data for new theory. *American Journal of Political Science*, 67(4), 880-897.
- Bonikowski, B., & Gidron, N. (2016). Multiple traditions in populism research: Toward a theoretical synthesis. *APSA Comparative Politics Newsletter*, 26(12), 7-14.
- Bornschieer, S. (2010). *Cleavage politics and the populist right. The new cultural conflict in Western Europe*. Temple University Press.
- Canovan, M. (1999). Trust the people! Populism and the two faces of democracy. *Political studies*, 47(1), 2-16.
- Carroll, R. (2014). *Comandante: Hugo Chávez's Venezuela*. Penguin Books.
- Chester, P. J. (2023). *Embedded Lexica*.
- Chesterley, N., & Roberti, P. (2018). Populism and institutional capture. *European Journal of Political Economy*, 53, 1-12.
- De la Torre, C. (2017). *Populism in Latin America* (pp. 195-213). Oxford: Oxford University Press.
- Deegan-Krause, K., & Haughton, T. (2009). Toward a more useful conceptualization of populism: Types and degrees of populist appeals in the case of Slovakia. *Politics & Policy*, 37(4), 821-841.
- Di Cocco, J., & Monechi, B. (2022). How populist are parties? Measuring degrees of populism in party manifestos using supervised machine learning. *Political Analysis*, 30(3), 311-327.
- Galston, W. A. (2019). The populist challenge to liberal democracy. *Journal of Democracy*, 29 (2), 5-19.
- Gidron, N., & Bonikowski, B. (2013). Varieties of populism: Literature review and research agenda.
- Haggard, S., & Kaufman, R. R. (2016). *Dictators and democrats. masses, elites and regime change*. Princeton, NJ: Princeton University Press.
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. arXiv preprint arXiv:1605.09096.

- Hawkins, K. A. (2010). *Venezuela's Chavismo and populism in comparative perspective*. Cambridge University Press.
- Hawkins, K. A., Carlin, R. E., Littvay, L., & Kaltwasser, C. R. (Eds.). (2018). *The ideational approach to populism: Concept, theory, and analysis*. Routledge.
- Houle, C., & Kenny, P. D. (2018). The political and economic consequences of populist rule in Latin America. *Government and Opposition*, 53 (2), 256–287.
- Howard, J., & Ruder, S. (2018). Universal language model fine-tuning for text classification. arXiv preprint arXiv:1801.06146.
- Kazin, M. (1995). The Agony and Romance of the American Left. *The American Historical Review*, 1488-1512.
- Khodak, M., Saunshi, N., Liang, Y., Ma, T., Stewart, B., & Arora, S. (2018). A La Carte Embedding: Cheap but Effective Induction of Semantic Feature Vectors. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, eds. Iryna Gurevych and Yusuke Miya, 12–22. Cedarville, OH: Association for Computational Linguistics.
- Kozłowski, A. C., Taddy M, & Evans J. (2018). The Geometry of Culture: Analyzing Meaning through Word Embeddings. arXiv preprint arXiv:1803.09288.
- Kulkarni, Vivek, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. (2015). Statistically Significant Detection of Linguistic Change. In *Proceedings of the 24th International Conference on World Wide Web, gen. chairs*. Aldo Gangemi, Stefano Leonardi, and Alessandro Panconesi, 625–35. Geneva: International World Wide Web Conferences Steering Committee and Republic and Canton of Geneva, Switzerland.
- Laclau, Ernesto (2006). On Populist Reason. *Tijdschrift Voor Filosofie* 68 (4):832-835.
- Ruth-Lovell, S. P., Lührmann, A., & Grahn, S. (2019) Democracy and Populism: Testing a Contentious Relationship. V-Dem Working Paper 2019:91.
- Mazucca, S. L. (2013). The rise of rentier populism. *J. Democracy*, 24, 108.
- Mols, F., & Jetten, J. (2016). Explaining the appeal of populist right-wing parties in times of economic prosperity. *Political Psychology*, 37(2), 275-292.
- Mudde, C. (2004). The populist zeitgeist. *Government and opposition*, 39(4), 541-563.
- Mudde, C., & Kaltwasser, C. R. (2017). *Populism: A very short introduction*. Oxford University Press.
- Pappas, T. S. (2019). *Populism and liberal democracy: A comparative and theoretical analysis*. Oxford University Press.
- Pop-Eleches, G. (2010). Throwing out the bums: Protest voting and unorthodox parties after communism. *World Politics*, 62(2), 221-260.

- Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science*, 54(1), 209-228.
- Reicher, S., & Haslam, S. A. (2017). The politics of hope: Donald Trump as an entrepreneur of identity. Why irrational politics appeals: Understanding the allure of Trump, 25-40.
- Rheault, L., & Cochrane, C. (2020). Word embeddings for the analysis of ideological placement in parliamentary corpora. *Political Analysis*, 28(1), 112-133.
- Rice, D. R., & Zorn, C. (2021). Corpus-based dictionaries for sentiment analysis of specialized vocabularies. *Political Science Research and Methods* 9 (1): 20–35. issn: 20498489.
- Roberts, K. M. (2022). Populism and polarization in comparative perspective: Constitutive, spatial and institutional dimensions. *Government and Opposition*, 57(4), 680-702.
- Rodman, E. (2020). A timely intervention: Tracking the changing meanings of political concepts with word vectors. *Political Analysis*, 28(1), 87-111.
- Rodriguez, P. L., Spirling, A., & Stewart, B. M. (2023). Embedding regression: Models for context-specific description and inference. *American Political Science Review*, 1-20.
- Rooduijn, M., & Pauwels, T. (2011). Measuring populism: comparing two methods of content analysis. *West Eur.Polit.* 34(6), 1272–1283 (2011)
- Rydgren, J. (2007) ‘The Sociology of the Radical Right’. *Annual Review of Sociology*, Vol. 33, No. 1, pp. 241–62.
- Ruth, S. P. (2018). Populism and the erosion of horizontal accountability in Latin America. *Political Studies*, 66 (2), 356-375.
- Stavrakakis, Y., Kioupiolis, A., Katsambekis, G., Nikisianis, N., & Siomos, T. (2016). Contemporary Left-wing Populism in Latin America: Leadership, Horizontalism, and Postdemocracy in Chávez's Venezuela. *Latin American Politics and Society*, Fall 2016, Vol. 58, No. 3 (Fall 2016), pp. 51-76
- Turney, P., & Pantel, P. (2010). From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research* 37:141–188.
- Weyland, K. (2001). Clarifying a contested concept: Populism in the study of Latin American politics. *Comparative Politics*, 34, 1-22.
- Yao, Z., Sun, Y., Ding, W., Rao, N., & Xiong, H. (2018). Dynamic Word Embeddings for Evolving Semantic Discovery. *WSDM 2018*, February 5–9, 2018, Marina Del Rey, CA, USA.