# Using Network Motifs to Detect Populist Narratives at the Sub-sentence Level*

Valentina González-Rostani† Bree Bang-Jensen‡ Michael Colaresi§

January 26, 2026

**Abstract**

Populist rhetoric frames politics as a moral conflict between a virtuous "people" and a corrupt "elite," yet existing measures typically label entire speeches or actors as populist, obscuring how such rhetoric is constructed and deployed within texts. We introduce Populist-PULSAR, a fine-grained, interpretable framework that measures populist rhetoric at the sub-sentence level. Building on thin-centered theories of populism, the approach represents narratives as graph-based motifs that capture who acts on whom, with moral valence. Combining human annotation with LLMs, Populist-PULSAR identifies four core populist motifs—victimization, empowerment, resistance, and elite conspiracy—independent of lexical choices. We validate our method using Donald Trump's 2016 campaign speeches, uncovering strategic temporal and geographic variation in populist appeals, notably heightened close to the election, in swing states and economically vulnerable regions. By revealing when and how distinct populist narratives are mobilized, Populist-PULSAR makes it possible to study populism as strategic political behavior rather than as an aggregate trait.

**Keywords**: Populism, text-analysis, network motifs, LLMs.

*We thank participants at MPSA, PolMeth, ISA, NYU's Data Science Frontiers: Society and Politics, Pace Research Institute Oslo, University of Georgia, reviewers at RSF grant, and the 10th Monash-Paris-Warwick-Zurich-CEPR Text-As-Data Workshop for valuable comments and suggestions.

†Department of Political Science and International Relations, University of Southern California. gonzalez.rostani@usc.edu.

‡Department of International Affairs, University of Georgia. Bree.BangJensen@uga.edu.

§Department of Political Science, University of Pittsburgh & Peace Research Institute Oslo mcolaresi@pitt.edu.

# 1 Introduction

Populist rhetoric, which frames politics as a moral struggle between a virtuous "us" and a corrupt "them," has become a defining feature of contemporary politics across democratic systems. This form of rhetoric is not merely symbolic. By casting elites, institutions, or outsiders as adversaries of "the people," populist appeals shape how citizens evaluate political authority, expertise, and social groups, with consequences for democratic trust and polarization (e.g., Gidron and Hall, 2017; Rodrik, 2017; Anelli et al., 2021). Understanding how populist rhetoric is constructed and deployed is central to the study of political behavior and democratic accountability.

Despite its prominence, existing research relies on coarse measures labeling entire politicians, parties, or speeches as populist or not. These approaches obscure political communication's internal heterogeneity: politicians frequently mix populist and non-populist language within speeches or paragraphs. This limitation is consequential given populism's status as a thin-centered ideology—populist frames flexibly attach to diverse issues and adapt to contexts. Without tools capturing this flexibility, we lack insight into when and how populist messages appear.

We address this gap by introducing Populist-PULSAR, a detection pipeline that measures populist rhetoric at the sentence and clause level. Grounded in thin-centered theories of populism, Populist-PULSAR identifies the relational structure of populist narratives—who is acting, who is being judged, and how moral evaluations are directed—rather than relying on fixed vocabularies or document-level labels. The method combines small-scale human annotation with large language model (LLM) assistance in a hybrid workflow. Human coders label core components of populist narratives, and the system uses these annotations to recover recurring narrative structures in text. Populist-PULSAR identifies not only whether a sentence is populist, but also which type of populist appeal it advances and which words and relations activate it. We focus on four representative populist motifs: victimization (the people harmed by elites), empowerment (the people reclaiming power), resistance (defiance against corrupt authorities), and elite conspiracy (elites acting in their own interest).

To demonstrate the analytic payoff of this approach, we apply Populist-PULSAR to a large corpus of U.S. political speech (2000–2025) and analyze all post-primary Trump campaign rallies in

2016. Populist rhetoric intensifies as Election Day approaches and is more prevalent in swing states and economically vulnerable areas. The method also reveals contextual variation in populist narratives: empowerment-oriented, pro-worker appeals dominate in economically vulnerable regions, while elite-conspiracy narratives are most common in highly competitive electoral environments. These findings show that fine-grained text measurement captures not only the intensity of populist rhetoric but also its composition and strategic deployment across time and space.

This paper makes three contributions. First, it advances populist rhetoric measurement by shifting from aggregate labels to internal language structure, aligning text-as-data methods with thin-centered theories and prior dictionary-based approaches (e.g., Rooduijn and Pauwels, 2011a; Di Cocco and Monechi, 2021; Dai and Kustov, 2022a; Meijers and Zaslove, 2021). Second, this measurement strategy yields a substantive contribution by enabling the analysis of populism as a strategic and context-dependent form of political communication rather than a fixed attribute of actors or texts. By identifying when, how, and which populist narratives are mobilized, the framework allows researchers to study geographic and temporal variation in populist rhetoric. Third, we introduce a hybrid human–LLM annotation pipeline that integrates expert judgment with scalable automation, demonstrating how LLMs can expand theoretically grounded analysis without replacing human oversight (e.g., González-Rostani et al., 2025; Le Mens and Gallego, 2025). In this pipeline, LLMs are used both to assist annotation and to augment data by generating semantically equivalent variants of expert-coded sentences, preserving narrative structure while increasing coverage.

The paper proceeds as follows. We first review existing approaches to measuring populism and identify the aggregation problem motivating our analysis (Section 2). We then develop the conceptual framework for fine-grained measurement (Section 3) and introduce Graph Edit Distance (GED) as a novel tool for structural comparison between observed rhetoric and populist motifs (Section 4). Next, we describe the implementation and scaling strategy (Section 5). We then validate the measure and apply it to U.S. campaign rallies from 2016 (Section 6). The conclusion outlines extensions and directions for future research.

## 2 The Case for Fine-grained Measurement

**Populism as discourse and narrative structure.** Populism is a contested concept, variously treated as a strategy of political mobilization, a leadership-centered mode of organization, a communication style, or an "ideational" worldview (e.g., Mudde, 2007a; Hawkins and Rovira Kaltwasser, 2018). We follow the thin-centered/ideational approach in which populism frames politics as a moralized conflict between a virtuous, homogeneous people and a corrupt elite (Mudde, 2007a; Hawkins and Rovira Kaltwasser, 2018). Two implications matter for measurement. First, populism is expressed linguistically and can be measured using political texts. Second, populist appeals need not be stable traits of parties or leaders: the same actor can shift between populist and non-populist framing across topics, audiences, and time. As a result, populist rhetoric is often modular and intermittent rather than pervasive, appearing in specific passages rather than entire speeches. This motivates measurement that can capture within-speech variation and the relational content of claims (who is cast as "the people" or "the elite," who acts on whom, and with what moral valence), rather than by only assigning aggregate labels.

**Why populism rises and why measurement matters.** Across democracies, anti-elite and people-centered appeals mobilize discontent among diverse groups. Sociological accounts emphasize social status, resentment, and isolation (Gidron and Hall, 2017; Hochschild, 2018), while economists stress exposure to globalization, automation, and economic insecurity (Rodrik, 2017; Frey et al., 2017). Political scientists link local economic decline to populist voting and party success (e.g., Anelli et al., 2021; Gonzalez-Rostani, 2025; Milner, 2021), and work in public administration shows how populist incumbents can erode bureaucratic capacity and institutional constraints (e.g., Bauer and Becker, 2020). Yet our understanding of populist rhetoric remains limited; most studies treat populism as a document or speaker-level property, obscuring strategic, episodic deployment. This aggregation obscures when politicians invoke populist narratives, how blame is assigned, and how rhetorical strategies vary across contexts. We address this limitation by developing a fine-grained measure of populist rhetoric that captures the structure of populist narratives at the sentence level.

**Existing approaches to measuring populist rhetoric and the aggregation problem.** Given the theoretical salience of populism, a variety of approaches have emerged to measure populist rhetoric. Existing measures face a recurring trade-off between scale, interpretability, and theoretical fidelity. Expert surveys and expert-coded party scores offer broad coverage but are often coarse and difficult to connect to specific textual evidence (e.g., Meijers and Zaslove, 2021; Norris, 2020). Manual coding of speeches and manifestos can align closely with theory and support rich qualitative validation, but scaling to large multilingual corpora is costly and raises reliability challenges (e.g., Hawkins and Rovira Kaltwasser, 2018; Jenne et al., 2021). Dictionary-based approaches scale easily but are brittle across contexts and struggle with implicit or paraphrased populism (e.g., Rooduijn and Pauwels, 2011b; Pauwels, 2011). Supervised text models improve portability and predictive performance, yet many operationalizations still collapse rhetoric into document- or speaker-level scores, limiting interpretability and obscuring within-text variation (e.g., Di Cocco and Monechi, 2022; Dai and Kustov, 2022b; Ulinskaitė and Pukelis, 2021). A classifier may successfully categorize a speech as populist, but it typically does not indicate which sentences or phrases triggered that classification. This lack of transparency complicates validation and limits substantive interpretation.

Moreover, when training data relies on party-level or speech-level labels, supervised models can inherit what we term the *aggregation problem*: by averaging over entire texts, they may conflate genuine populist narrative elements with unrelated content that merely correlates with populist speakers. As a result, two speeches may receive the same populism label for different reasons—one driven by victimization narratives, another by elite-conspiracy framing—despite sharing little vocabulary. Without finer resolution, such models risk false positives and mask the relational structure central to thin-centered populism.[1]

Across approaches, aggregation conflicts with the thin-centered view of populism. If populist rhetoric is modular, document-level labels necessarily gloss over meaningful variation: a speech might be only partially populist, yet a binary or scalar score captures neither the proportion nor the location of populist content. This limits both theory testing and empirical leverage, making it difficult to study how populist narratives interact with policy content such as nationalism, redistri-

[1] We further discuss this trade-off and document how existing measures differ along these dimensions in subsection A.1 and Table A.1.

bution, or trade. Researchers studying strategic populist communication—such as whether leaders intensify motifs during campaigns or tailor appeals to specific audiences (e.g., Gennaro et al., 2021; Gonzalez-Rostani, 2025)—gain little insight from aggregated scores.

Aggregation also creates interpretive blind spots. Even probabilistic document-level measures obscure how specific textual elements correspond to populist concepts, undermining critique, replication, and cross-context comparability. If populism is inherently context-dependent, research requires sentence- and phrase-level methods to trace how populist narratives emerge, evolve, and intersect with other political discourse. Bridging the gap between scope and granularity is thus central to building cumulative knowledge about populism.

## 3 A Fine-grained Interpretable Measure of Populism

Our project develops a new method to measure populist rhetoric at the sentence and sub-sentence level by extending the PULSAR framework (Park et al., 2018)[2] to identify populist components and structures. PULSAR represents political communication as a graph, with words and phrases as nodes and their syntactic and semantic relationships as edges. This allows us to formalize the core insight that political discourse expresses *who* does what to *whom* with which valence. We extend PULSAR in two ways. First, we introduce populist-specific role labels that distinguish in-group actors and targets (US) from out-group elites (THEM), and identify the judgment holder (speaker vs. non-speaker), a distinction especially relevant for political discourse. Second, we define a set of populist narrative motifs as graph templates that represent the structural essence of thin-centered populism: a moral struggle between "the people" and "the elites."

Rather than treating entire speeches or politicians as uniformly populist, our approach identifies which specific sentences or clauses match these motifs, allowing for fine-grained analysis of when and how populist rhetoric is deployed. Our system can also accommodate multiple narrative structures within a single sentence. Political speech often layers judgments, embeds one narrative within another, or presents competing perspectives. By tracking groupings of related components within sentences, we capture this complexity while maintaining clear links between extracted structures

[2]Parsing Unstructured Language into Sentiment-Aspect Representations (PULSAR) was originally designed to extract structured judgments from human rights reports.

and their textual evidence.

### 3.1 The Narrative Components of Populism

We define a token-level ontology for annotating populist discourse. Below, we outline our conceptualization, and Table 1 lists all labels mapped to core components of thin-centered populism (Mudde, 2007b). Token-level annotation preserves within-sentence structure. **ACTOR** and **TARGET** identify who causes and who receives the judged action, each assigned a role attribute (*US*, *THEM*, or *Residual*). **JUDGMENT** and **ASPECT** capture evaluative stance and the evaluated concept, respectively, each carrying valence attributes (positive, negative, neutral). **JUDGEMENT-HOLDER** distinguishes speaker from non-speaker judgment holders, critical for identifying endorsement versus critique. **GROUP** tags link components belonging to the same evaluative statement within multi-structure sentences.

| **Populist Narrative Component** | **Component Instantiation and Ontology Label** | |
|---|---|---|
| | **Variant A** | **Variant B** |
| Actor (Acting) | Actor–US (People) | Actor–THEM (Elite) |
| Target (Acted upon) | Target–US (People) | Target–THEM (Elite) |
| Judgment (Valence) | Judgment-on-aspect (+) | Judgment-on-aspect (–) |
| Perspective (Judgment Holder) | Judgment-holder–Speaker | Judgment-holder–Nonspeaker |

Table 1: Mapping populist narrative components to Populist-PULSAR ontology labels. *Note:* On the left are populist narrative components, and on the right are the corresponding labels in the PopulistPULSAR system. PopulistPULSAR labels have a main label, like "Actor" as well as specific attributes, for example "Us" as the role attribute for "Actor".

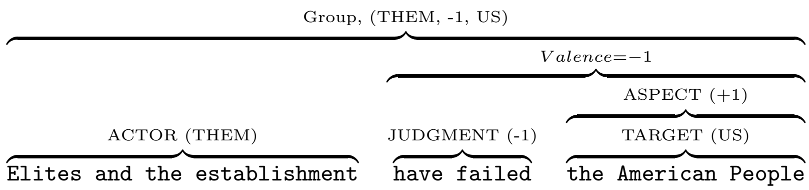
### 3.2 The Structure of Populism

We conceptualize populism as a set of *recurring narrative structures* rather than as a property of entire speeches or speakers. At its core, each populist narrative links an Actor to a Target through an evaluative action, defined by a judgment applied to an aspect. This structure formalizes the idea that populist discourse expresses *who* does *what* to *whom*, and with what *moral valence.*

In our framework, these narrative structures are represented as graph motifs. Nodes correspond to labeled components—actors, targets, and judgment–aspect pairs—while edges encode their relations. Edge direction distinguishes actors from targets, valence captures whether the action is framed positively or negatively, and perspective identifies whose judgment is being expressed (e.g., the speaker or a reported source). When multiple evaluative statements appear within the same sentence, grouping relations determine which components belong to the same narrative instance.

For analytical tractability, we represent each narrative structure in a simplified triplet form: **Judgment-holder(Actor-role, Valence, Target-role)**. For example, the sentence below is encoded as **(THEM, $-1$, US)**. This triplet captures the core narrative in which an out-group actor harms the in-group—what we refer to as an elite-conspiracy motif in populist discourse.

*"Elites and the Establishment have failed the American People."*




## 3.3 Recurrent Narratives: Populist Motifs

This graph-based representation allows us to identify a small set of recurrent and theoretically relevant narratives, which we call **populist motifs**, that capture the structural logic of thin-centered populism. These include narratives of victimization (elites harming the people), empowerment (the people benefiting the people), elite conspiracy (elites benefiting elites), and resistance (the people acting against elites). Each motif corresponds to a distinct configuration of actor role, target role, and valence, independent of the specific vocabulary used. Table 2 summarizes the structure of each motifs and Figure 1 provides illustrative examples.

| **Representative Motif** | **Conceptualization** | | **Operationalization** |
|---|---|---|---|
| | **Narrative Components** | **Narrative Structure** | **Graph Motif** |
| Victimization | Actor (THEM), Target (US), Negative Valence | Elites $\xrightarrow{\text{harm}}$ People | Speaker(THEM, –, US) |
| Empowerment | Actor (US), Target (US), Positive Valence | People $\xrightarrow{\text{help}}$ People | Speaker(US, +, US) |
| Elite Conspiracy | Actor (THEM), Target (THEM), Positive Valence | Elites $\xrightarrow{\text{help}}$ Elites | Speaker(THEM, +, THEM) |
| Resistance | Actor (US), Target (THEM), Negative Valence | People $\xrightarrow{\text{harm}}$ Elites | Speaker(US, –, THEM) |

Table 2: Representative populist motifs and their translation from theoretical narrative components to graph-based operationalization in Populist-PULSAR.

Our graph-based approach moves beyond keyword matching by capturing the structural logic of populist narratives. Sentences with different wording can instantiate the same populist motif when their relational structures align—for example, "We will protect American workers" and "We will bring back our jobs" both express the Empowerment motif—while superficially similar statements

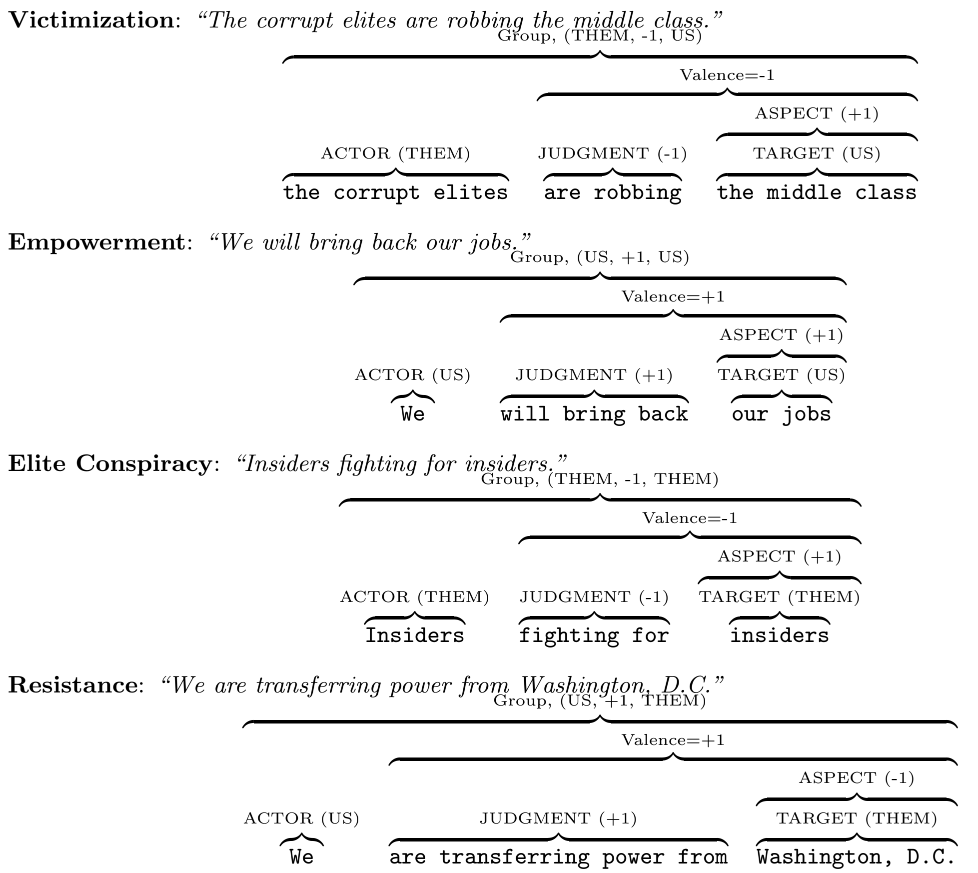



Figure 1: Populist Motifs Illustration through Actor–Judgment–Target decompositions

may reflect distinct narrative frames. By representing populism as graph motifs, Populist-PULSAR enables fine-grained measurement at the sentence and clause level, capturing how populist rhetoric is constructed and deployed in political discourse.

# 4 Comparing Extracted Graphs to Populist Motifs

To assess whether an extracted narrative matches a populist motif, we compute a custom *Graph Edit Distance* (GED) between the sentence-level graph and each motif template. GED yields a scalar measure of structural similarity, comparing narratives by their underlying roles and judgments rather than surface wording. As discussed in the previous section, each sentence is represented as a reduced triplet `Speaker(Actor, Valence, Target)`, and populist motifs correspond to a small set of recurrent relational structures (e.g., Empowerment `(US, +, US)`). Given this shared representation, GED maps the distance between an extracted triplet and each motif to a nonnegative scalar, with lower values indicating closer substantive alignment. Formally, $GED_{ij} : (G_i, G_j) \to d_{ij}$ with $d_{ij} \in \mathbb{Z}_{\geq 0}$. Let $t = (a, v, b)$ denote an extracted triplet and $m = (a^*, v^*, b^*)$ a motif template, where $a, b \in \texttt{US}, \texttt{THEM}, \texttt{Residual}$ and $v \in +, -$. We define an additive distance

$$d(t, m) \;=\; c_{\text{role}}(a, a^*) \;+\; c_{\text{val}}(v, v^*) \;+\; c_{\text{role}}(b, b^*),$$

with costs chosen to encode interpretive "severity":

- **Exact match:** cost 0 when the component matches.
- **Residual ambiguity:** cost 1 when a specified role (`US` or `THEM`) is replaced by `Residual` (or vice versa), reflecting missing or underspecified role information.
- **US/THEM substitution:** cost 2 when switching between `US` and `THEM` in either actor or target, reflecting a substantive reassignment of who is portrayed as virtuous versus culpable.
- **Valence reversal:** cost 4 when flipping polarity ($+ \leftrightarrow -$), reflecting the strongest change because it inverts the moral direction of the statement.

For each extracted triplet $t$, we compute its distance to every motif template $m \in \mathcal{M}$ and take

$$d_{\min}(t) = \min_{m \in \mathcal{M}} d(t, m),$$

using the argmin to identify the nearest motif. A sentence (or triplet) can then be treated as populist when $d_{\min}(t)$ falls at or below a chosen threshold $\tau$. Setting $\tau = 0$ yields the most conservative case (exact structural matches only). Allowing $\tau = 1$ admits minor role ambiguity (via `Residual`), and higher thresholds admit progressively larger departures in actor/target assignment and polarity. This yields a single, interpretable similarity scale that supports both substantive analysis of motif proximity and evaluation of automated extraction against human references. To illustrate, Figure 2 represents populist motifs in a three-dimensional actor–valence–target space induced by the GED penalty structure. Once we have a new sentence, extracted sentence graphs are mapped into this space and assigned to the nearest motif. Table 3 presents examples using an Empowerment populist motif as reference.

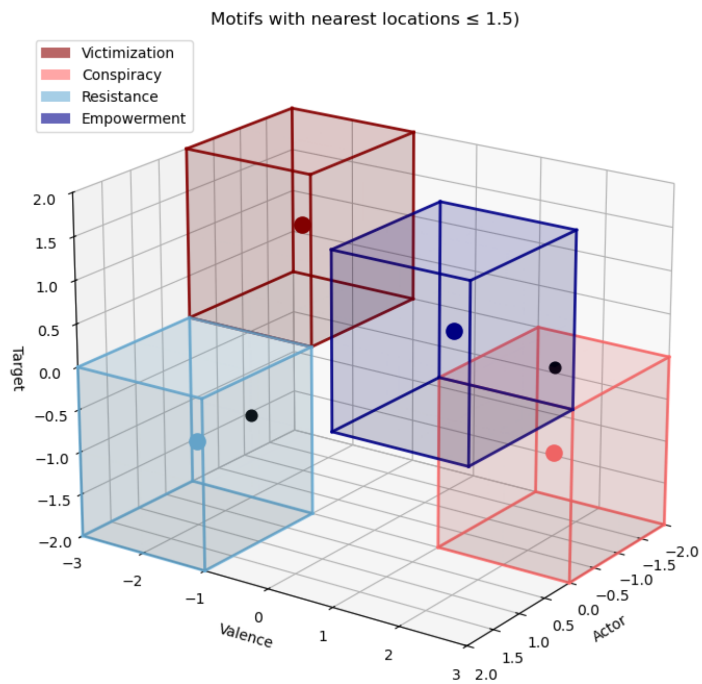



Figure 2: Populist Motif and GED representation

GED serves two roles in our pipeline. First, it supports motif identification by assigning each extracted structure to its closest template (or to "none" if it is too far). Second, it supports evaluation by providing a graded notion of agreement between an automated extraction and a human-annotated reference structure. GED has a long history as an inexact matching tool for graphs (Blumenthal et al., 2021), but here we tailor the edit costs to reflect the meaning of populist actor/target roles and moral polarity.

## 5 Scaling Up - Populist-PULSAR Implementation

Our conceptual contribution—measuring thin-centered populist narratives at the sentence and sub-sentence level—creates a practical bottleneck, as manual annotation is too costly to scale to modern corpora. We address this with a hybrid pipeline that preserves the theoretical precision of expert coding while using computational tools (LLMs and fine-tuned transformers) to expand coverage. Figure 3 summarizes the workflow from raw text to structured motif detection and final outputs, which we apply to U.S. political speech from 2008 to the present across campaign and governing contexts (e.g., rallies, debates, conventions, and formal addresses). The resulting dataset retains

| Sentence Example | Triplet | Penalization | GED to Ref. |
|---|---|---|---|
| **Empowerment Reference:** *"We will bring back our borders"*<br>$\overbrace{\texttt{We}}^{\text{ACTOR (US)}}\quad \overbrace{\texttt{will bring back}}^{\text{JUDGMENT (+1)}}\quad \overbrace{\texttt{our}}^{\text{TARGET (US)}}\quad \overbrace{\texttt{borders}}^{\text{Aspect (+1)}}$ | (US, +, US) | (0,0,0) | 0 |
| **Comparison 1:** *"We will bring back our jobs"*<br>$\overbrace{\texttt{We}}^{\text{ACTOR (US)}}\quad \overbrace{\texttt{will bring back}}^{\text{JUDGMENT (+1)}}\quad \overbrace{\texttt{our}}^{\text{TARGET (US)}}\quad \overbrace{\texttt{jobs}}^{\text{Aspect (+1)}}$ | (US, +, US) | (0,0,0) | 0 |
| **Comparison 2:** *"We will bring back our wealth"*<br>$\overbrace{\texttt{We}}^{\text{ACTOR (US)}}\quad \overbrace{\texttt{will bring back}}^{\text{JUDGMENT (+1)}}\quad \overbrace{\texttt{our}}^{\text{TARGET (US)}}\quad \overbrace{\texttt{wealth}}^{\text{Aspect (+1)}}$ | (US, +, US) | (0,0,0) | 0 |
| **Comparison 3:** *"I know it's all about making America great again for all Americans"*<br>$\overbrace{\texttt{I}}^{\text{ACTOR (Residual)}}\quad \overbrace{\texttt{'s all about making}}^{\text{JUDGMENT (+1)}}\quad \overbrace{\overbrace{\texttt{America}}^{\text{TARGET (US)}}\ \texttt{great Again}}^{\textit{Aspect}(+1)}$ | (Residual, +, US) | (1,0,0) | 1 |
| **Comparison 4:** *"The insiders wrote the rules of the game to keep themselves in power"*<br>$\overbrace{\texttt{The insiders}}^{\text{ACTOR (THEM)}}\quad \overbrace{\texttt{wrote}}^{\text{JUDGMENT (+1)}}\quad \overbrace{\texttt{the rules of the game ...}\ \overbrace{\texttt{themselves}}^{\text{TARGET (THEM)}}\ \texttt{... the money}}^{\textit{Aspect}(+1)}$ | (THEM, +, THEM) | (2,0,2) | 4 |
| **Comparison 5:** *"We have to stop them from cheating in elections."*<br>$\overbrace{\texttt{We}}^{\text{ACTOR (US)}}\quad \overbrace{\texttt{have to stop}}^{\text{JUDGMENT (-1)}}\quad \overbrace{\texttt{them}}^{\text{TARGET (THEM)}}\quad \overbrace{\texttt{from cheating in elections}}^{\text{ASPECT (+1)}}$ | (US, -, THEM) | (0,4,2) | 6 |
| **Comparison 6:** *My opponent does not support tax cuts.*<br>$\overbrace{\texttt{My opponent}}^{\text{ACTOR (Residual)}}\quad \overbrace{\texttt{does not support}}^{\text{JUDGMENT (-1)}}\quad \overbrace{\texttt{job creation}}^{\text{Aspect (+1)}}$ | (Residual, -, Residual) | (1,4,1) | 6 |

Table 3: Examples Illustrating GED Metric for Populist Motifs

*Note:* GED in this example is calculated with respect to the empowerment reference motif (US, +, US), with an example sentence provided. Sentences exactly matching this structure (comparisons 2 and 3) have a GED of 0. Minor deviations, such as ambiguous actor roles, incur a minimal penalty (GED=1). Complete reversal of both roles from US to THEM incurs a penalty of (GED=4), indicating significant rhetorical divergence. The last two comparisons provide examples of how a larger distance can signal a different populist motif (comparison 5 matches the resistance motif) or non-populist narrative structures (comparison 6 has a minimum distance from all populist narrative motifs of 2).

interpretability at scale by storing, for each sentence (and within-sentence group), annotated text spans, a reduced triplet representation, the nearest populist motif (if any), and the GED to that motif—allowing aggregation across speakers and settings while still supporting audits that trace results back to the original text.

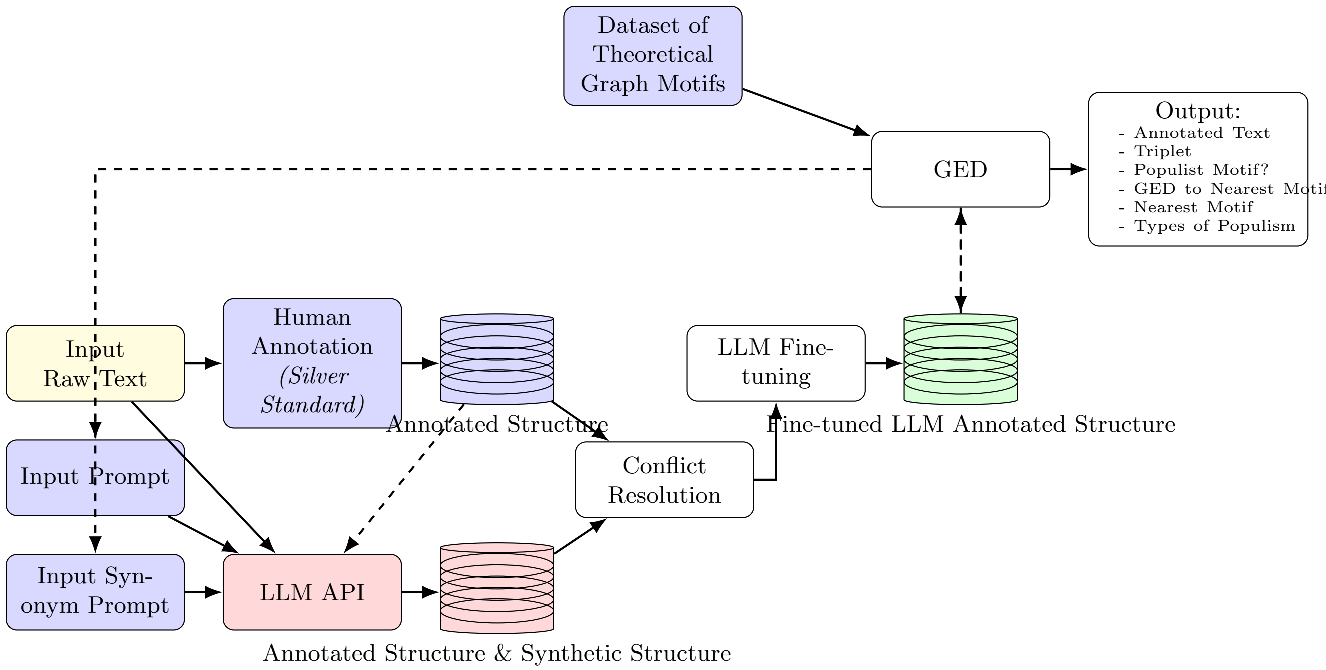



Figure 3: Summary of Our Populist Motif Approach
*Note*: Cylinders represent data stacks; yellow indicates raw inputs, red LLM-generated steps, blue human inputs, and green fine-tuned outcomes.

### 5.1 From Small-Scale Expertise to Large-Scale Annotation

The core scaling idea is staged: a small, carefully curated set of human annotations establishes a reference standard, which is then expanded via LLM-assisted labeling and synthetic augmentation. These labeled data support fine-tuning models that can annotate large corpora automatically.

#### Stage 1: Human Annotations as a Silver Standard

We begin by manually annotating a strategically sampled subset of sentences to establish a high-confidence reference set ("silver standard"). Coders identify the narrative components in the Populist-PULSAR ontology—ACTOR, TARGET, JUDGMENT, ASPECT, and JUDGMENT-HOLDER—and assign group IDs linking components into coherent structures when multiple aspects appear in the same sentence. [3]

[3] Two implementation details matter for downstream modeling: (i) spans are annotated at the character level, allowing labels to capture negation and other polarity-shifting cues; and (ii) groupings need not be contiguous, permitting the system to represent long-distance links between components within a sentence. We use an interactive annotation interface (Prodigy) to support this workflow (see Appendix subsection A.5).

### Stage 2: Prompt-Based LLM Annotation Expansion

We next use LLMs to annotate additional sentences at scale using structured prompting. Prompts include (a) compact label definitions aligned with the Populist-PULSAR ontology, (b) few-shot examples drawn from the silver standard to anchor the task, and (c) modular decomposition of the annotation into subtasks (e.g., identify actors/targets; identify judgment and aspect; assign valence; then group components).[4] We implement LLM calls through the OpenAI API (`gpt-4o-mini-2024-07-18`), setting temperature to zero to improve reproducibility.

Following the logic of Self-Refine prompting (Madaan et al., 2023), we test candidate prompts on held-out silver-standard sentences, diagnose recurrent failure modes (e.g., valence errors; confusing targets vs. aspects; missing group links), and incorporate these cases as targeted counterexamples or clarifications in subsequent prompt versions. This produces LLM-generated annotations that expand the labeled corpus while remaining consistent with the human-coded reference standard.

To evaluate LLM-generated annotations, we compare them to a human-annotated silver standard using the GED metric. We deem a classification acceptable when the total GED is low—specifically 0 or 1, or up to 2–3 when discrepancies reflect minor role substitutions (e.g., `Residual` instead of `US` or `THEM`) without polarity reversal. Overall, 76.2% of LLM annotations meet this criterion. Of these, 27.4% are exact matches (GED = 0), and an additional 40.5% exhibit minimal discrepancy (GED = 1), typically involving a group role versus `Residual`. These results indicate that most LLM–human differences are small and do not meaningfully change the extracted narrative structure.

### Stage 3: Synthetic Data Generation for Linguistic Diversity

Beyond labeling new sentences, we use LLMs to generate semantically equivalent variants of already annotated sentences by substituting context-appropriate synonyms for the ACTOR, TARGET, JUDGMENT, and ASPECT spans while holding the underlying labels fixed. This augmentation increases linguistic variety (e.g., "corrupt elite" $\rightarrow$ "out-of-touch establishment") and helps address class imbalance by multiplying scarce positive instances for particular motifs. If we draw five alternatives for each of four components, a single sentence yields up to $5^4 = 625$ variants. Refer to Figure 4 for an example. Because these synthetic sentences inherit the original labels, they enlarge

[4] Appendix A.5.1 further explains the prompt engineering steps.

the training set without additional human coding.

This approach yields two significant advantages. First, it substantially increases the linguistic diversity of our training corpus (Edunov et al., 2018; Schick and Schütze, 2021). By exposing our models to varied phrasings of the same underlying structure, we improve their ability to generalize beyond the specific vocabulary present in our manually annotated examples. Second, synthetic data generation helps address class imbalance in our training data. Populist sentences, particularly those matching specific motifs like (THEM, +, THEM) or (US, -, THEM), may be less common in our initial sample than non-populist statements. By generating multiple variants of each populist example, we ensure the model receives sufficient positive training instances to learn the structural patterns that define populist discourse.

**Prompt-Based Data Augmentation Example**

*Original sentence:*

"The insiders wrote the rules of the game to keep themselves in power and in the money."

*LLM-extracted components:*

- **Actor**: `elites, establishment, powerful, decision-makers, ruling class`
- **Judgment**: `maintain, preserve, secure, uphold, safeguard`
- **Aspect**: `authority and wealth, influence and financial gain, dominance and prosperity, control and resources, power and riches`
- **Target**: `their grip, their own interests, their position, their status, their benefits`

*Selection of augmented sentences:*

1. "The elites wrote the rules of the game to keep themselves in power and in the money."
2. "The establishment wrote the rules of the game to keep themselves in power and in the money."
3. "The insiders wrote the rules of the game to uphold power and wealth."
4. "The insiders wrote the rules of the game to keep themselves in control and resources."
5. "The insiders wrote the rules of the game to protect their status and wealth."

Figure 4: Example of prompt-based data augmentation. A single sentence is decomposed into structured narrative components and recombined to generate semantically equivalent variants that preserve the underlying narrative structure.

## 5.2 Model Training for Automated Extraction

The hybrid-human LLM annotation pipeline described above is sufficient for direct application to political speech corpora. However, to further enhance scalability and reduce per-sentence annotation costs, we fine-tune transformer-based models for automated label extraction. At the sentence level, we fine-tune DistilBERT in a multi-task setup with parallel heads predicting actor role (US/THEM/Residual), target role (US/THEM/Residual), and judgment/aspect valence. Each head has its own cross-entropy loss function, and we minimize the average loss across all tasks. Joint training allows shared representations to exploit cross-component regularities (e.g., negative judgments co-occurring with THEM targets).[5]

Table 4 reports evaluation metrics for our best-performing sentence-level model on a held-out test set. The model achieves strong performance across most tasks, with F1 scores ranging from 0.86 to 0.96 for Actor and Target role classification, and 0.90 to 0.93 for Judgment valence. Aspect valence proves more challenging, with F1 scores between 0.67 and 0.71, reflecting the difficulty of inferring whether an aspect is inherently good or bad from contextual cues alone. For further details, refer to subsection A.6.

| Label | Attribute | Precision | Recall | F1 |
|---|---|---|---|---|
| Actor | Them | 0.86 | 0.86 | 0.86 |
| | Residual | 0.91 | 0.91 | 0.91 |
| | Us | 0.93 | 0.93 | 0.93 |
| Target | Them | 1.00 | 0.90 | 0.95 |
| | Residual | 1.00 | 0.83 | 0.91 |
| | Us | 0.88 | 1.00 | 0.93 |
| Judgment | -1 | 0.93 | 0.87 | 0.90 |
| | +1 | 0.89 | 0.96 | 0.93 |
| Aspect | -1 | 0.86 | 0.55 | 0.67 |
| | +1 | 0.59 | 0.89 | 0.71 |

Table 4: Sentence-level, test-set evaluation metrics for fine-tuned model with attention pooling.

Sentence-level predictions do not identify the evidence spans. We therefore fine-tune four separate token-level BERT-based sequence tagging models, one for each narrative component (ACTOR, TARGET, JUDGMENT, ASPECT). Each model uses BERT as the encoder to produce contextualized embeddings for every token in the sentence. A token-level classification layer then outputs

[5] Training was conducted using the HuggingFace Transformers framework with the following parameters: 100 epochs, batch size of 8 sentences per GPU, AdamW optimizer with an initial learning rate of $5 \times 10^{-5}$, and a linear learning rate schedule. We monitored performance on a held-out validation set after each epoch, tracking both overall loss and per-task accuracy and F1 scores. We also searched for how many layers to freeze during fine-tuning, with and without attention pooling.

a probability distribution over BIO-style tags for that component.[6][7] Appendix subsection A.7 presents our best-performing token-level fine-tuned model performance.

# 6 Validation

This section evaluates the empirical performance and substantive validity of Populist-PULSAR. We proceed in three steps. First, we assess face and convergent validity using illustrative examples and comparisons with established dictionary-based approaches. Second, we examine construct validity by testing whether populist motifs are deployed strategically as predicted. Third, we show how motif-level disaggregation yields insights that are unavailable to aggregate or document-level measures, while remaining transparent. We conduct our validation on 89 general-election rally speeches by Donald Trump in 2016.

## 6.1 Face Validity: Illustrations

A key face-validity test is whether the measure captures populist language regardless of speaker identity. Table 5 shows well-known populist politicians sometimes use neutral rhetoric, while generally non-populist figures occasionally deploy classic populist motifs, indicating our approach responds to textual content rather than speaker labels.

We further illustrate face validity using a segment from Trump's 2016 Pensacola rally, previously classified as populist in its entirety by Dai and Kustov (2022a). Our sentence-level analysis shows that 10 of 19 sentences contain explicit populist motifs, justifying an overall populist label while revealing substantial internal variation. Two motifs dominate: victimization narratives (elites harming the people) and empowerment narratives (the people reclaiming control). As shown in Figure 5, Trump shifts from policy discussion to populist framing by attributing responsibility to "insiders" or other elite actors. This demonstrates the advantage of sentence-level analysis: Populist-PULSAR identifies not only whether a passage is populist, but where and how populist

[6] For example, the ACTOR model assigns tags indicating whether each token is the beginning of an actor span (B-ACTOR-US, B-ACTOR-THEM), inside an actor span (I-ACTOR-US, I-ACTOR-THEM), or outside any actor mention (O).

[7] We fine-tuned each sequence labeling model for 100 epochs with a batch size of 16, using AdamW optimization (learning rate $2 \times 10^{-5}$, weight decay 0.01). To prevent overfitting, we employed the dropout regularization.

| Triplet | Example (Speaker) |
|---|---|
| **Populist Leaders Using Non-Populist Motifs** | |
| [Residual, +, US] | I will declare a national emergency at our southern border. (Trump)<br>At its most basic, this election is about preserving our democracy. (Sanders)<br>Climate change is real. It is caused by human activity and it is already causing devastating problems in the United States and around the world. (Sanders) |
| [Residual, -, US] | Mothers and children trapped in poverty in our inner cities. (Trump)<br>We're facing the worst public health crisis in 100 years and the worst economic collapse since the great depression. (Sanders) |
| [US, +, Residual] | Here, we memorialize the brave men and women who struggled to sacrifice, and sacrificed so much so that others might live in freedom.(Trump)<br>We are confronting systemic racism and the enormous threat to our planet of climate change. (Sanders)<br>This election is the most important in the modern history of this country. (Sanders) |
| [Residual, +, Residual] | My proudest legacy will be that of a peacemaker and unifier. (Trump)<br>The scientific community has spoken in a virtually unanimous voice. (Sanders) |
| **Non-Populist Leaders Using Populist Motifs** | |
| [US, +, US] | This generation was given the gift of the best education in American history. (Bush) |
| [THEM, -, US] | And we've seen over the last 4 years, the status quo in Washington, they are powerful, and they have fought us every step of the way. (Obama)<br>That's surrender to the same forces of the status quo that has squeezed middle class families for way too long. (Obama) |

Table 5: Examples of Discourse-Speaker Mismatch in Populist Motifs

narratives enter the text.

Comparing methods reinforces this point. Applying the approach of Di Cocco and Monechi (2021) to the same excerpt yields broad agreement that most sentences are populist, but without distinguishing among narrative types (see Table A.9). By contrast, dictionary-based methods flag only a single sentence driven by the explicit keyword "corrupt," missing most populist framing in the passage. This gap reflects a central limitation of keyword-based approaches: they require lexical cues even when populist narratives are expressed implicitly.

## 6.2 Convergent Validity: Comparison with Dictionary Measures

Next, we evaluate how Populist-PULSAR's speech-level scores align with a standard dictionary-based populism measure (Rooduijn and Pauwels, 2011a). We find a significant positive correlation: as Figure 6 shows, speeches with more populist keywords tend to have more populist motif sentences. However, the correlation is far from perfect, as expected. Dictionary methods only flag populism when specific words appear, whereas Populist-PULSAR can identify populist narratives even with-

Change is coming. All the people who've rigged the system for their own personal benefit are trying to stop our change campaign because they know that their gravy train has reached its last stop. It's your turn now. This is your time. The fact that so many encrusted old political insiders oppose our campaign is the best proof you will ever need that we are fighting for real change not partisan change.
We are fighting for all Americans Democrats, Republicans, Independents, Conservatives, Liberals who've been failed by this **corrupt** system. We're fighting for everyone who doesn't have a voice. We're also fighting for every region of this country. For every part of Florida, and every part of America. From Pensacola to Pittsburgh, from Baltimore to Baton Rouge, we are fighting for every last city and every last person in this country. Hillary Clinton is the candidate of the past. Ours is the campaign of the future.
In this future, we are going to pursue new trade policies that put American workers first and that keep jobs in our country. All the people who got NAFTA wrong, and China wrong, and who are trying to give us the Trans-Pacific Partnership are the same failed voices pushing for Hillary Clinton. Our trade deficit with the world is now nearly $800 billion dollars. We've lost one-third of our manufacturing jobs since Bill and Hillary Clinton gave us NAFTA. China is manipulating its currency and taking our jobs. We are going to stop companies from leaving our country and keep those jobs right here in America. The era of economic surrender is over.

Figure 5: Segment from Donald Trump's 2016 Pensacola rally highlighting populist motifs.
*Note*: Magenta represents the Elite Conspiracy Narrative (THEM, +, THEM); orange represents the Victimization Narrative (THEM, –, US); teal represents the Empowerment Narrative (US, +, US); and blue represents the Resistance Narrative (US, –, THEM).

out those exact keywords. Thus, our approach can detect populist rhetoric that a keyword-based method would miss.

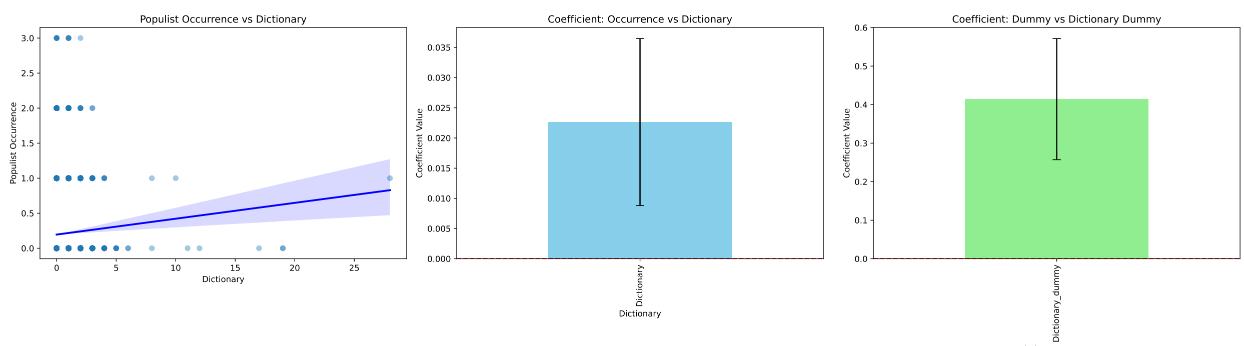



Figure 6: Relationship between Populist Motifs and Dictionary Measures

### 6.3 Construct validity: strategic populism in Trump's 2016 rallies

In addition to comparisons with existing measures, we evaluate whether our approach recovers theoretically predicted patterns. Drawing on 89 general-election rally speeches from Donald Trump's 2016 campaign, we replicate and extend the dictionary-based analysis of Gonzalez-Rostani (2025) to examine when, where, and which populist motifs are deployed. In this application, the annotation pipeline is fully automated, and motif matches are aggregated to the speech level. We test three empirical expectations: populist rhetoric intensifies as Election Day approaches, is more prevalent in competitive (swing-state) contexts, andresponds to local economic distress.

**Temporal Dynamics.** As Election Day approached, Trump's rhetoric became increasingly populist. Regressing the presence of populist motifs on days until the election (controlling for rally characteristics) shows a significant effect: rallies held closer to Election Day have a higher share of populist sentences (Figure 7). This trend supports the idea that populism is a strategic tool for mobilization, rather than a fixed personal style. It also demonstrates that our measure captures within-campaign temporal shifts that a speaker-level average would miss.

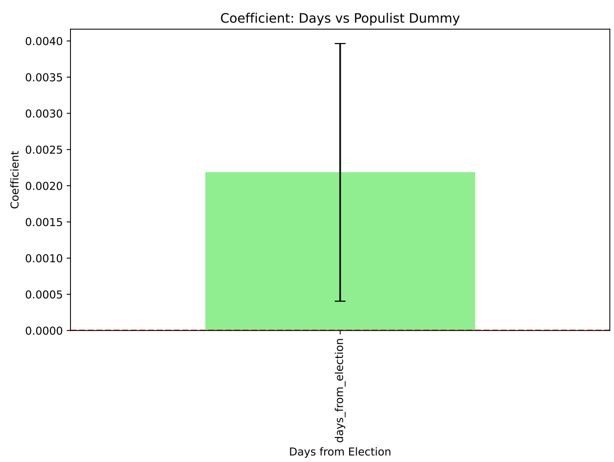
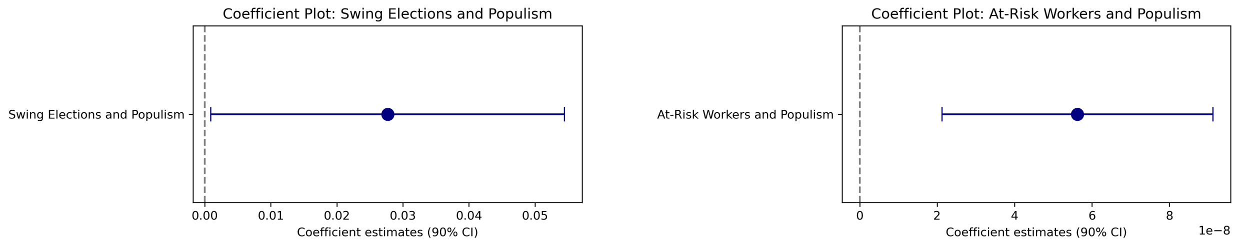



Figure 7: Coefficient plot: Days from Election vs. Populist Dummy

**Geographic Targeting.** Populist rhetoric also varied by location. Trump used more populist language in competitive states: rallies in swing states showed significantly higher populist motif counts than those in safe states (Figure 8a). This is a substantial effect and suggests that our measure is not just capturing the speaker's general style, since the same candidate shifted to a more populist tone in battleground contexts.

Moreover, populist motifs were more common in economically distressed areas. Using the county-level share of jobs at high risk of automation as a proxy for local vulnerability (Muro et al., 2019; Gonzalez-Rostani, 2025), we find that rallies in more exposed counties contained higher levels of populist rhetoric (Figure 8b). This pattern is consistent with theories linking economic insecurity to populist appeal (e.g., Gidron and Hall, 2017; Inglehart and Norris, 2016; Gonzalez-Rostani, 2025; Gennaro et al., 2021).



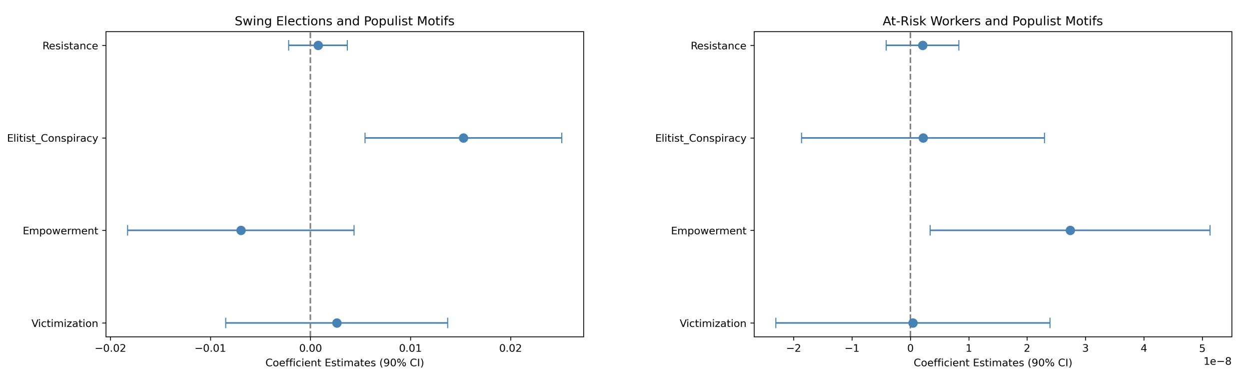

(a) Swing Elections and Populism

(b) At-Risk Workers and Populism

Figure 8: Coefficient plots with 90% confidence intervals

## 6.4 Motif-level heterogeneity and interpretability

Because Populist-PULSAR distinguishes among specific motifs, we can identify which narratives drive these patterns. In swing states, the increase in populist rhetoric comes mainly from elite-conspiracy motifs (Figure 9a). Meanwhile, in economically vulnerable regions, the uptick is driven more by empowerment motifs (Figure 9b), such as pro-worker statements.




(a) Swing Elections and Populist Motifs

(b) At-Risk Workers and Populist Motifs

Figure 9: Coefficient plots for populist motifs with 90% confidence intervals

Overall, Populist-PULSAR performs well across multiple validation exercises. The measure aligns with expert-coded examples and dictionary-based approaches, and it recovers theoretically expected temporal, geographic, and economic patterns in populist rhetoric. At the same time, operating at the sentence and motif level yields insights that aggregate, document-level scores cannot provide.

Rather than treating populism as a single, undifferentiated signal, our results show that populist rhetoric is strategically deployed, with distinct motifs serving different communicative functions across contexts. A key advantage of the approach is its transparency: each classification is directly linked to an explicit narrative structure in the text, making it possible to inspect why a sentence is labeled as populist. This combination of empirical validity, theoretical alignment, and interpretability allows Populist-PULSAR not only to reproduce broad patterns captured by existing methods, but also to reveal how and when specific populist narratives are mobilized, opening the door to richer analyses of populist strategy in political discourse.

# 7 Discussion

This paper introduced Populist-PULSAR, a fine-grained, motif-based approach to measuring populist rhetoric that operates at the sentence and sub-sentence level. By representing populist discourse as structured narrative motifs rather than as coarse document- or speaker-level labels, the approach aligns closely with thin-centered theories of populism and directly addresses the aggregation problem that characterizes much existing work. Instead of asking whether a speech or a politician is "populist," our approach identifies when, how, and in what form populist narratives appear within political text.

Methodologically, the paper introduces structured rhetorical motifs as an intermediate representation between raw text and downstream inference—simpler than full semantic parses but richer than bag-of-words or dictionary-based approaches. Representing rhetoric as graphs and applying graph-edit-distance matching enables the identification of near-equivalent rhetorical structures despite lexical variation. The annotation and evaluation strategy emphasizes interpretability: motif outputs can be inspected, audited, and adjudicated, supporting standard validity claims in political methodology. Combining small-scale human annotation with prompt-based and LLMs reduces annotation costs while preserving theoretical control, demonstrating how structured representations can enhance both transparency and measurement validity.

Empirically, motif-level measurement uncovers variation aggregate indicators obscure, tracing when speakers selectively deploy rhetorical components across contexts. We validate the approach by showing convergent validity with dictionary-based measures and construct validity through theoretically expected patterns. In Donald Trump's 2016 campaign rallies, populist motifs increased as the election approached, appeared more frequently in swing states, and varied with local economic vulnerability. The method reveals motif-level heterogeneity: elite-conspiracy narratives are more common in competitive contexts, while empowerment narratives dominate in economically distressed regions, underscoring populism's flexibility rather than a monolithic style.

While this approach represents a step forward in the measurement of populism, several limitations point to clear directions for future work. The system currently treats sentences as independent units; incorporating co-reference resolution and cross-sentence linking would allow compatible partial motifs to be merged and longer narrative arcs to be recovered. Another promising extension is to connect rhetorical motifs to political mobilization—for example, by combining motif-based text analysis with audio or video data from campaign rallies to assess which narratives elicit stronger audience responses. Finally, although Populist-PULSAR is designed to study populist rhetoric, its underlying ontology and tools are not populism-specific. The hybrid annotation pipeline and graph-based matching strategy are readily extendable to other forms of political rhetoric defined by relational and evaluative structure, including nationalism, blame attribution, and moralized policy discourse.

Fine-grained measurement shapes what researchers can claim about strategy, persuasion, and democratic contestation. By making narrative structure observable at scale, Populist-PULSAR helps move beyond debates over whether an actor "is" populist toward a more precise account of when particular populist moves appear, how they are combined, and how they vary across contexts. This, in turn, enables sharper tests of theories about the conditions under which populist rhetoric is adopted and the consequences it may have for democratic politics. In an era in which populist language is both widespread and heterogeneous, measurement tools that are interpretable, scalable, and structurally grounded are a prerequisite for cumulative empirical progress.

## References


Anelli, M., I. Colantone, and P. Stanig (2021, November). Individual vulnerability to industrial robot adoption increases support for the radical right. *Proceedings of the National Academy of Sciences 118*(47).

Bauer, M. W. and S. Becker (2020, March). Democratic Backsliding, Populism, and Public Administration. *Perspectives on Public Management and Governance 3*(1), 19–31.

Blumenthal, D. B., J. Gamper, S. Bougleux, and L. Brun (2021, June). Upper Bounding the Graph Edit Distance Based on Rings and Machine Learning. *International Journal of Pattern Recognition and Artificial Intelligence 35*(08), 2151008. arXiv:1907.00203 [cs].

Brown, T., B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei (2020). Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, Volume 33, pp. 1877–1901. Curran Associates, Inc.

Casiraghi, M. C. (2021, November). 'You're a populist! No, you are a populist!': The rhetorical analysis of a popular insult in the United Kingdom, 1970–2018. *The British Journal of Politics and International Relations 23*(4), 555–575.

Cleary, D. (2025). PromptHub Blog: 10 Best Practices for Prompt Engineering with Any Model.

Dai, Y. and A. Kustov (2022a, January). When Do Politicians Use Populist Rhetoric? Populism as a Campaign Gamble. *Political Communication*, 1–22.

Dai, Y. and A. Kustov (2022b). When do politicians use populist rhetoric? populism as a campaign gamble. *Political Communication 39*(3), 383–404.

Di Cocco, J. and B. Monechi (2021, October). How Populist are Parties? Measuring Degrees of Populism in Party Manifestos Using Supervised Machine Learning. *Political Analysis*, 1–17.

Di Cocco, J. and B. Monechi (2022). How populist are parties? measuring degrees of populism in party manifestos using supervised machine learning. *Political Analysis 30*(3), 311–327.

Edunov, S., M. Ott, M. Auli, and D. Grangier (2018, October). Understanding Back-Translation at Scale. arXiv:1808.09381 [cs].

Elçi, E. (2019, July). The Rise of Populism in Turkey: A Content Analysis. *Southeast European and Black Sea Studies 19*(3), 387–408.

Frey, C. B., T. Berger, and C. Chen (2017). Political machinery: Automation anxiety and the 2016 US presidential election. *University of Oxford*.

Gennaro, G., G. Lecce, and M. Morelli (2019). Intertemporal evidence on the strategy of populism.

Gennaro, G., G. Lecce, and M. Morelli (2021, January). Mobilization and the Strategy of Populism Theory and Evidence from the United States.

Geroimenko, V. (2025). Key Principles of Good Prompt Design. In V. Geroimenko (Ed.), *The Essential Guide to Prompt Engineering: Key Principles, Techniques, Challenges, and Security Risks*, pp. 17–36. Cham: Springer Nature Switzerland.

Gidron, N. and P. A. Hall (2017, November). The politics of social status: economic and cultural roots of the populist right. *The British Journal of Sociology 68*(S1).

Gonzalez-Rostani, V. (2025). Elections, Right-wing Populism, and Political-Economic Polarization: The Role of Institutions and Political Outsiders. *The Journal of Politics*.

González-Rostani, V., J. Incio, and G. Lezama (2025). Social media versus surveys: A new scalable approach to understanding legislators' discourse. *Legislative Studies Quarterly 50*(2), 258–266. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/lsq.12481.

Greene, K. (2020, September). *Battles Over Perceptions Among Violent Non-state Actors*. University of Pittsburgh ETD. Num Pages: 124 Publisher: University of Pittsburgh.

Gründl, J. (2020, December). Populist ideas on social media: A dictionary-based measurement of populist communication. *New Media & Society*, 1461444820976970.

Hawkins, K. A., R. E. Carlin, L. Littvay, and C. R. Kaltwasser (2019). *The ideational approach to populism: Concept, theory, and analysis*. Routledge.

Hawkins, K. A. and C. Rovira Kaltwasser (2018, April). Measuring populist discourse in the United States and beyond. *Nature Human Behaviour 2*(4), 241–242.

Hochschild, A. R. (2018, February). *Strangers in Their Own Land: Anger and Mourning on the American Right*. The New Press. Google-Books-ID: nJ0uDwAAQBAJ.

Inglehart, R. F. and P. Norris (2016). Trump, Brexit, and the rise of populism: Economic have-nots and cultural backlash. Publisher: HKS Working paper no. RWP16-026.

Jagers, J. and S. Walgrave (2007). Populism as political communication style. An empirical study of political parties' discourse in Belgium.

Jenne, E. K., K. A. Hawkins, and B. C. Silva (2021, June). Mapping Populism and Nationalism in Leader Rhetoric Across North America and Europe. *Studies in Comparative International Development 56*(2), 170–196.

Le Mens, G. and A. Gallego (2025). Positioning political texts with Large Language Models by asking and averaging. *Political Analysis*, 1–9. Publisher: Cambridge University Press.

Madaan, A., N. Tandon, P. Gupta, S. Hallinan, L. Gao, S. Wiegreffe, U. Alon, N. Dziri, S. Prabhumoye, Y. Yang, S. Gupta, B. P. Majumder, K. Hermann, S. Welleck, A. Yazdanbakhsh, and P. Clark (2023, May). Self-Refine: Iterative Refinement with Self-Feedback. arXiv:2303.17651 [cs].

Mede, N. G. and M. S. Schäfer (2020). Science-related populism: Conceptualizing populist demands toward science. *Public Understanding of science 29*(5), 473–491.

Meijers, M. J. and A. Zaslove (2021, February). Measuring Populism in Political Parties: Appraisal of a New Approach. *Comparative Political Studies 54*(2), 372–407.

Milner, H. V. (2021, November). Voting for Populism in Europe: Globalization, Technological Change, and the Extreme Right. *Comparative Political Studies 54*(13), 2286–2320.

Mudde, C. (2007a). Populist radical right parties in europe. *(No Title)*.

Mudde, C. (2007b). *Populist radical right parties in Europe.* Cambridge: Cambridge university press.

Muro, M., R. Maxim, and J. Whiton (2019). Automation and artificial intelligence: How machines are affecting people and places. Publisher: Brookings Institution.

Nishikawa, M. (2021, January). How Populistic were the Populists in 19th Century America?: Analysis by Automated Textual Analysis.

Norris, P. (2020, November). Measuring populism worldwide. *Party Politics 26*(6), 697–717. Publisher: SAGE Publications Ltd.

OpenAI (2025). Prompt engineering - OpenAI API.

Park, B., M. Colaresi, and K. Greene (2018, December). Beyond a Bag of Words: Using PULSAR to Extract Judgments on Specific Human Rights at Scale. *Peace Economics, Peace Science and Public Policy 24*(4), 20180030.

Pauwels, T. (2011, February). Measuring Populism: A Quantitative Text Analysis of Party Literature in Belgium. *Journal of Elections, Public Opinion and Parties 21*(1), 97–119.

Polk, J., J. Rovny, R. Bakker, E. Edwards, L. Hooghe, S. Jolly, J. Koedam, F. Kostelka, G. Marks, and G. Schumacher (2017). Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data. *Research & Politics 4*(1), 2053168016686915.

Reynolds, L. and K. McDonell (2021, February). Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. arXiv:2102.07350 [cs].

Rodrik, D. (2017). Populism and the economics of globalization. Technical report, National Bureau of Economic Research.

Rooduijn, M. (2019). State of the field: How to study populism and adjacent topics? A plea for both more and less focus. *European Journal of Political Research 58*(1), 362–372.

Rooduijn, M., S. L. de Lange, and W. van der Brug (2014, July). A populist Zeitgeist? Programmatic contagion by populist parties in Western Europe. *Party Politics 20*(4), 563–575.

Rooduijn, M. and T. Pauwels (2011a, November). Measuring Populism: Comparing Two Methods of Content Analysis. *West European Politics 34*(6), 1272–1283.

Rooduijn, M. and T. Pauwels (2011b). Measuring populism: Comparing two methods of content analysis. *West European Politics 34*(6), 1272–1283.

Sahoo, P., A. K. Singh, S. Saha, V. Jain, S. Mondal, and A. Chadha (2025, March). A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications. arXiv:2402.07927 [cs].

Schaffer, F. C. (2000, July). *Democracy in Translation: Understanding Politics in an Unfamiliar Culture* (1st edition ed.). Ithaca, NY: Cornell University Press.

Schick, T. and H. Schütze (2021, October). Generating Datasets with Pretrained Language Models. arXiv:2104.07540 [cs].

Schulhoff, S., M. Ilie, N. Balepur, K. Kahadze, A. Liu, C. Si, Y. Li, A. Gupta, H. Han, S. Schulhoff, P. S. Dulepet, S. Vidyadhara, D. Ki, S. Agrawal, C. Pham, G. Kroiz, F. Li, H. Tao, A. Srivastava, H. D. Costa, S. Gupta, M. L. Rogers, I. Goncearenco, G. Sarli, I. Galynker, D. Peskoff, M. Carpuat, J. White, S. Anadkat, A. Hoyle, and P. Resnik (2025, February). The Prompt Report: A Systematic Survey of Prompt Engineering Techniques. arXiv:2406.06608 [cs].

Ulinskaitė, J. and L. Pukelis (2021, June). Identifying Populist Paragraphs in Text: A machine-learning approach. arXiv:2106.03161 [cs].

Zhou, Y., A. I. Muresanu, Z. Han, K. Paster, S. Pitis, H. Chan, and J. Ba (2022, September). Large Language Models are Human-Level Prompt Engineers.

# A Online Appendix

## Contents

## A.1 Problems with Past Measurement

As populists have become increasingly successful at both gaining political power and capturing public imagination in the years following the Great Recession, scholars have developed a range of strategies to classify politicians or political parties as populist. These efforts aim to understand the prevalence of populism, the conditions under which it emerges, and its broader political consequences (Rooduijn, 2019). Researchers have developed a number of approaches to measure populist speech, many of which rely on analysis of political text.

One approach employs expert surveys in which researchers rely on scholars' evaluations of political parties. Notable examples include the 2017 Chapel Hill Expert Survey Polk et al. (2017) and the 2018 Populism and Political Parties Expert Survey (POPPA) (Meijers and Zaslove, 2021), both of which focus on European parties. Norris (2020) expands this method globally through the Global Party Survey (GPS). Other approaches focus directly on textual data, analyzing political content such as party manifestos, legislative speeches, or media coverage. For example, Jagers and Walgrave (2007) developed a dictionary of populist terms and manually classified text based on its presence, an approach also adopted by Hawkins and Rovira Kaltwasser (2018); Hawkins et al. (2019); Jenne et al. (2021) and Casiraghi (2021). While these strategies have meaningfully advanced our understanding of populist discourse, they also exhibit important limitations: namely, a lack of granularity, a limited ability to capture nuance, and high resource costs in terms of human labor and cross-contextual adaptability.

**Coarse Granularity:** Most text-based systems of populism measurement assign a binary label—populist or not populist—to entire political parties based on manifestos or platforms.[8] Other approaches apply these labels to large sections of text, such as multi-paragraph speech segments. This level of granularity may be appropriate for "thicker" definitions of populism, which assume that particular policy positions (e.g., protectionism) inherently signal a populist orientation. However, this approach fails to capture the more nuanced, "thin-centered" definition of populism developed by Mudde (2007b), which focuses on a core worldview that pits a virtuous people against a corrupt elite, without necessarily attaching this dichotomy to a specific ideological or policy agenda.

Political actors and texts often display variation that coarse classification cannot accommodate. A single political party may include both populist and non-populist members. A candidate may employ populist motifs when addressing one audience—such as factory workers—but avoid them in other settings, or may shift their rhetorical strategy over time. Even a single speech may mix populist and non-populist rhetoric, depending on the issues discussed. Certain topics may lend themselves more readily to populist framing than others. Assigning a single label to an entire speech, figure, or party risks mischaracterizing this variation. A thin-centered understanding of populism demands an approach capable of recognizing that not all statements made by a figure with a populist worldview will themselves be populist in nature. Measurement systems must be sensitive to such variation in content and tone.

Moreover, as populism has become a salient political theme, political actors increasingly reference populist narratives—even when they are not endorsing them. Speakers may critique populist arguments, quote populist figures, or invoke populist rhetoric only to reject it. Within one speech, a speaker may critique populist themes used by an opponent while invoking their own. Thus, a measurement system must go beyond detecting the presence of populist motifs, such as the people-versus-elite dichotomy. It must also identify the stance the speaker takes towards that narrative. In

[8] See Appendix A.2 for examples on the problems of these aggregate-level tools.

other words, the system should detect not only whether a populist frame is used, but also whether the speaker adopts, critiques, or neutrally describes that frame.[9]

**Lack of Nuance:** Subject-matter-based or "thick" measures of populism also fail to capture the interplay between populism and other political discourses, such as nationalism, anti-globalism, or technoskepticism. This limits researchers' ability to distinguish, for example, between left-wing and right-wing populism, or to trace how populist rhetoric interacts with different topics across time and context. Incorporating such nuance allows for richer analytical distinctions. For instance, following the COVID-19 pandemic, there has been a growing association between populist discourse and anti-science sentiment, as described by Mede and Schäfer (2020). A robust measurement framework must be capable of identifying not only populist motifs, but also their entanglement with other discursive elements—including those that critique populism itself.

**Human Cost/Effort and Universality Tradeoffs:** Many existing methods rely on manual annotation of political texts, an approach that is both labor-intensive and time-consuming. Coders must be trained extensively, often requiring significant institutional investment. Although this approach allows for tailored annotation schemes developed by domain experts or native speakers—and can be adapted across languages and political systems—it struggles to scale. Some researchers have turned to dictionary-based methods as a more efficient alternative. For example, a number of studies grounded in European politics have used dictionaries to identify key elements of populist discourse—namely "people-centrism" and "anti-elitism"—as operationalized in (Pauwels, 2011; Rooduijn and Pauwels, 2011a; Rooduijn et al., 2014; Gennaro et al., 2021). While widely adopted, such dictionaries are inherently time- and context-specific. They often require substantial adaptation to be valid in new linguistic, political, or cultural contexts (Hawkins et al., 2019; Elçi, 2019; Gründl, 2020; Nishikawa, 2021; Gennaro et al., 2019). Moreover, the meaning of particular words or phrases can shift considerably across settings, echoing the warnings in Schaffer (2000) about the dangers of decontextualized textual analysis.

Table A.1: Existing Measures and the Thin-Centered Conceptual Approach

| **Approach** | **Examples** | **Fine-grained** | **Interpretable** | **Scalable** |
|---|---|---|---|---|
| **Experts Survey** | POPPA (Meijers and Zaslove, 2021); Global Party Survey (Norris, 2020) | | | |
| **Manual Coding** | (Jenne et al., 2021); (Hawkins and Rovira Kaltwasser, 2018) | | | |
| **Dictionary** | (Pauwels, 2011); (Rooduijn and Pauwels, 2011b); (Gennaro et al., 2021) | | | |
| **NLP — supervised** | (Di Cocco and Monechi, 2022); (Ulinskaitė and Pukelis, 2021); (Dai and Kustov, 2022a) | | | |

We need finer granularity to capture *when* and *how* populist narratives appear.

[9]While less common, some utterances may contain more complex structures, such as disagreement with a positive judgment of a populist (or anti-populist) judgment.

## A.2 Examples of the Limitations of Aggregate Party-Level Populism Labels in Prior Work

In this section, we illustrate how prior work using party-level labels to identify populist discourse—such as (Di Cocco and Monechi, 2021)—fails to capture within-party variation in rhetorical styles and political narratives. Specifically, we use examples based on weighted log-odds ratios and feature importance scores from different national contexts to demonstrate the limitations of binary classifications of parties as "populist" or "non-populist."

These figures are generated using the machine learning classified trained following Di Cocco and Monechi (2021), which classifies all speech from parties such as the SNP or UKIP as "populist" and others, such as Labour or Conservatives, as "non-populist." However, when we extract the discriminative features of speech using log-odds ratios and classifier feature importance, we observe two main problems:

1. **False Positives:** Words associated with issue positions (e.g., "government," "Scotland") are labeled as "populist" even if they do not reflect anti-elite, people-centric, or Manichean rhetoric. For instance, Figure A.1 shows that regional identifiers like "scotland" and "uk" are strongly associated with the non-populist class, while "britain" and "our" are associated with the populist class. But these are not inherently populist cues.

2. **False Negatives:** High-valence, moralistic, or collective language (e.g., "we," "freedom," "our country") can appear in parties coded as non-populist, such as mainstream U.S. Republican speeches. Figure A.2 demonstrates this, where terms such as "congress," "republican," and "government" are predictive of the Republican class, despite the party being labeled as populist.

These problems are further illustrated through feature importance scores for the United States and Italy ML models (Figures A.3 and A.4, respectively). For the U.S., terms like "will," "we," "believe," and "democrat" dominate as top features, suggesting that prediction relies more on group identity or future-orientation than on populist content per se. For Italy, technical or procedural language such as "forecast," "abolition," and "access" dominate—again missing the core features of populist rhetoric.

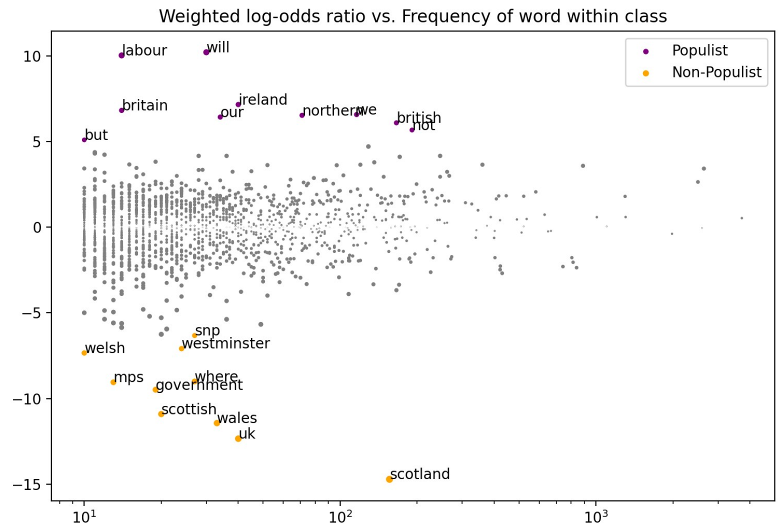



Figure A.1: Weighted log-odds ratio of words used in UK 2017 manifestos. Populist vs. Non-Populist parties.

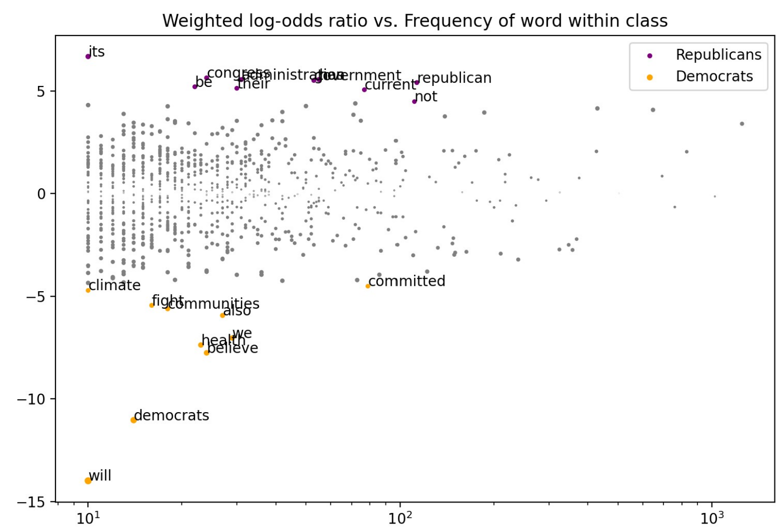



Figure A.2: Weighted log-odds ratio of words used in U.S. 2016 manifestos. Republicans vs. Democrats.

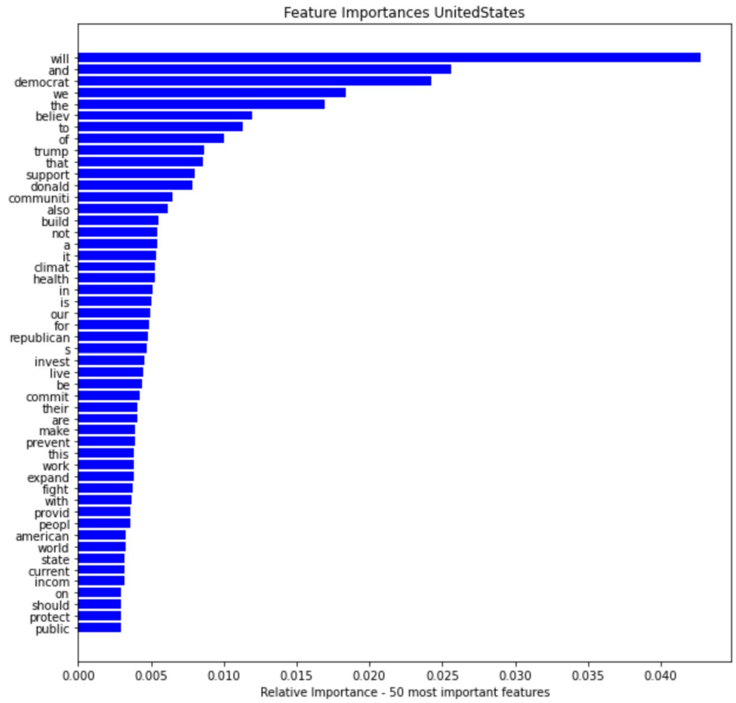



Figure A.3: Top features used by classifier trained to distinguish U.S. parties based on text.

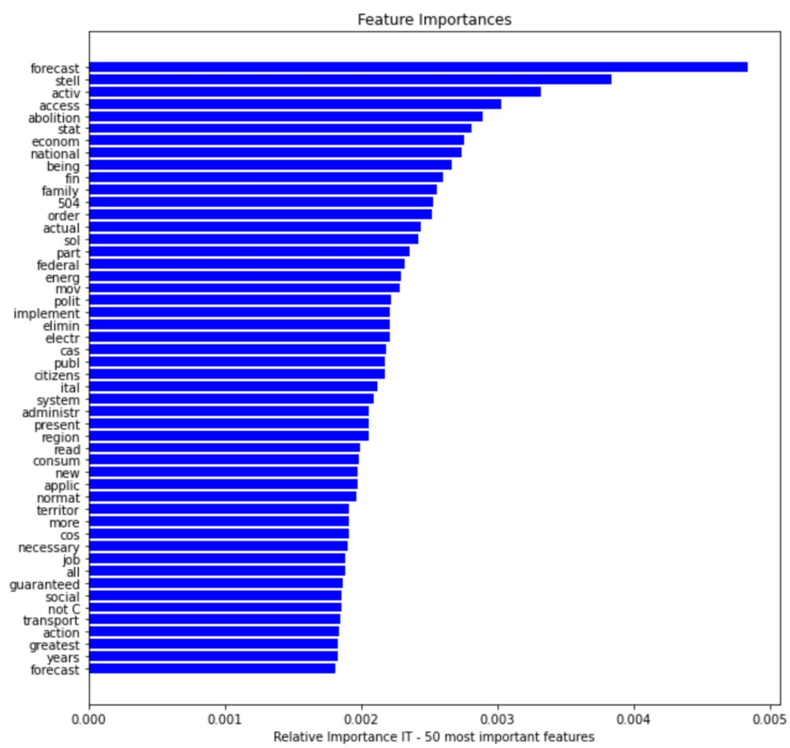



Figure A.4: Top features used by classifier trained on Italian political speech.

These visualizations collectively illustrate that classification based on entire-party labels fails to reflect the structure of populist narratives. As our project emphasizes, more fine-grained approaches—such as the one based on the Populist PULSAR annotation scheme—can more accurately identify the presence of populist communication, accounting for variation within and across speakers, parties, and issues.

## A.3 Benefits of our Approach

This approach offers several key advantages over existing methods, including greater granularity in measurement, the ability to capture the structural relationships within discourse, and improved portability across time periods and geographic contexts.

First, our method enables analysis at the sentence and sub-sentence level, capturing a more granular view of political speech than traditional approaches that classify entire parties, politicians, or speeches as populist or non-populist. This granularity opens up a range of possibilities for researchers interested in the interplay between populist and non-populist discourse. For instance, it enables exploration of how populist motifs co-occur with or diverge from other political themes, subject matters, or rhetorical styles. Right-wing and left-wing populists may invoke populist motifs in relation to different topics or policy areas. Researchers could investigate, for example, how the likelihood of populist framing varies when politicians of different ideological orientations discuss issues such as immigration, trade, finance, artificial intelligence, foreign aid, regulation, climate change, or vaccination. This level of detail also permits analysis of intra-party variation—identifying differences in populist rhetoric among members of the same party, as well as the evolution of an individual politician's use of populist motifs over the course of their career or in response to shifts in audience or political climate.

Second, by leveraging network motifs, our approach captures the structure of relationships between sentence components, rather than simply relying on the presence or absence of key terms. In contrast to dictionary-based approaches, which are anchored in word-level content, our method focuses on the interaction between actors, actions, judgments, and roles. This allows us to isolate the core dynamic of populist discourse: the struggle between the people and the elite. Moreover, this structural focus enables us to distinguish between sincere deployment of populist rhetoric and other uses of populist language, such as quotation, critique, or satire. This distinction is particularly important as politicians increasingly engage in meta-commentary on populism. For instance, in a speech on March 23, 2025, Senator Jon Ossoff stated, "they are literally the elites they pretend to hate"—a sentence that simultaneously references and critiques populist rhetoric. Without analyzing the underlying relational structure, specifically, the judgment implied about the relationship between "they" and "the elites", such a statement might be misclassified by a system relying solely on keywords.

Third, by focusing on abstracted roles, aspects, and judgments, rather than on specific lexical items, our method facilitates application across different languages, time periods, and political contexts. Unlike dictionary-based systems, which must be carefully tailored and revalidated for each new setting, our framework is grounded in rhetorical structure and conceptual roles. This makes it more adaptable for comparative research and better suited to identifying both universal and context-specific trends in populist discourse. As a result, researchers can more easily track the evolution of populist rhetoric over time or examine the global diffusion and localization of populist narratives across different regions and political cultures.

### A.4 Designing a Finer-grained, Interpretable Populist Narrative Measurement System

Here we provide more details of our system, how each part is implemented as well as how the parts fit together.

#### A.4.1 Sampling

Our dataset comprises political speeches delivered by U.S. politicians—including presidents, vice presidents, governors, senators, House members, and candidates—from 2008 onward. Speeches cover general and primary elections, national and state conventions, and formal addresses. We ensured balance between Republican and Democratic speakers, selecting roughly equal numbers of officials previously classified as populist or non-populist. This design supports comparative populist discourse analysis while limiting partisan bias. Future versions of our approach will expand our corpus internationally, facilitating cross-national comparisons of populist rhetoric.

#### A.4.2 PULSAR Ontology and Extension

Importantly, we do not have to build a system to identify populist narrative at the sentence level from scratch. The Parsing Unstructured Language into Sentiment-Aspect Representations (PULSAR) system has been created over the last decade to label allegations and judgments at both high textual resolution, scale to larger sets of documents, and allow for explanations anchored in the text for why specific decisions were made (Park et al., 2018). The PULSAR system operates on tokens at the sentence level – although it can be flexibly augmented with longer strings or limited to phrases. PULSAR assigns labels for tokens such as roles for an ACTOR, TARGET, and what is being judged, known as the ASPECT, and the JUDGMENT being expressed. PULSAR also connects related labels across the sentence – denoting for example that *this* ACTOR is judged to have violated *this* right for *this* TARGET. While PULSAR was developed to automatically annotate human rights reports and press releases (Park et al., 2018), it has previously been extended to encode stance-taking by rebel-government communications within civil conflict (Greene, 2020).

Table **??** compares the existing PULSAR ontology and the Populist-PULSAR extentions to match the components of thin-centered populist narrative components and structures. It also provides some

#### A.4.3 Defining a Database of Populist Narrative Motifs Using Populist PULSAR Ontology

Our first design choice is to identify the conceptual target of inference as the components and structure of thin-centered populist narratives, as described in the main text. We point out that these representations themselves are network motifs – recurrent subgraph structures and that they express the center of thin-centered populist narratives. These structures themselves are latent but emerge from reading texts with a populist-conceptual lens.

The roles include an US (the in-group that represents the people), THEM (the out-group that represents the elites) and who (whether the ACTOR and TARGET are VICTIMS or PERPETRATORS of rights PROTECTIONS or VIOLATIONS). For example, in the sentence, "The corrupt elite have stolen real American's jobs" illustrates a populist motif where a THEM actor ("corrupt elites") is being judged by the speaker to have taken something positive ("jobs") from the US target ("real American's"). We can summarize this motif as a graph-based data object that we define

| Label | Exists | Notes |
|---|---|---|
| ACTOR | ✓ | Causes what is judged |
| ACTOR-US | + | In-group associated with "people" is cause |
| ACTOR-THEM | + | Out-group associated wih "elites" is cause |
| TARGET | ✓ | Effected by what is judged |
| TARGET-US | + | In-group/People effected |
| TARGET-THEM | + | Out-group/Elites effected |
| JUDGMENT-POS | ✓ | Positive judgment of aspect (present, give, agree) |
| JUDGMENT-NEG | ✓ | Negative judgment of aspect (absent, take, oppose) |
| JUDGMENT-NEUTRAL | ✓ | Neutral judgment of aspect |
| ASPECT-POS | ✓ | A positive action, thing, or concept being judged |
| ASPECT-NEG | ✓ | A negative action, thing, or concept being judged |
| ASPECT-NEUTRAL | ✓ | A neural action, thing, or concept being judged |
| JUDGMENT_HOLDER | + | Non-composer/speaker reference who holds judgment |
| GROUP | + | Connection between labels |

Table A.2: Table of labels from PULSAR 3.0 ontology (✓) and those added in this project (+). Attributes for labels, such as US assigned to an ACTOR are treated here as a new label (eg ACTOR-US).

further below in detail as `Speaker(THEM, -, US)`, which is not only the definition of a grievance but an attribution of blame to the out-group. Conversely, a clause in which an US actor performs a negative action toward a THEM actor reflects the motif of the righteous struggle of the deserving people (e.g. $Speaker(US,\ -,\ THEM)$).

### A.4.4 How PULSAR Extracts Narrative Structures from Text

However, the populist network motifs we organize are only useful if we can label specific words and phrases within the text of actual speeches by the constituent parts (such as US, THEM, JUDGMENT) and the semantic relations between them as well as match them to the relevant (if any) populist narrative motif.

In order to explain our extension of PULSAR to code populist narratives, it is important to understand the core of the tool. To identify structured expressions of judgment in natural language, PULSAR has three inter-related parts. First, PULSAR is an ontology of labels. It identifies a finite set of roles that words and phrases can play in communicating a judgment in the text. Thus, instead of supervised learning at the sentence (or coarser) level, PULSAR operates on tokens within sentence to identify the constituent parts of a judgment. We specifically expand on the PULSAR 3.0 ontology which includes a nested tree of roles for judgments. The most basic roles are a) the ASPECT being discussed (typically a noun phrase) and b) JUDGMENT which is similar to the sentiment/valence (often a verb phrase) directed towards the ASPECT. Because of the complexity of language, PULSAR splits the JUDGMENT and ASPECT into components, these are pictured, with several examples in A.5. Crucially, PULSAR includes labels for positive and negative judgments which are represented in language as the presence (positive) or absence (negative) or alternatively the giving (positive) or taking (negative) of an aspect. Similarly, aspects are labeled as being good things, from the point of view of the judgment holder (positive) or harmful things (negative). We also highlight the actor and target in A.5 to continue the example.

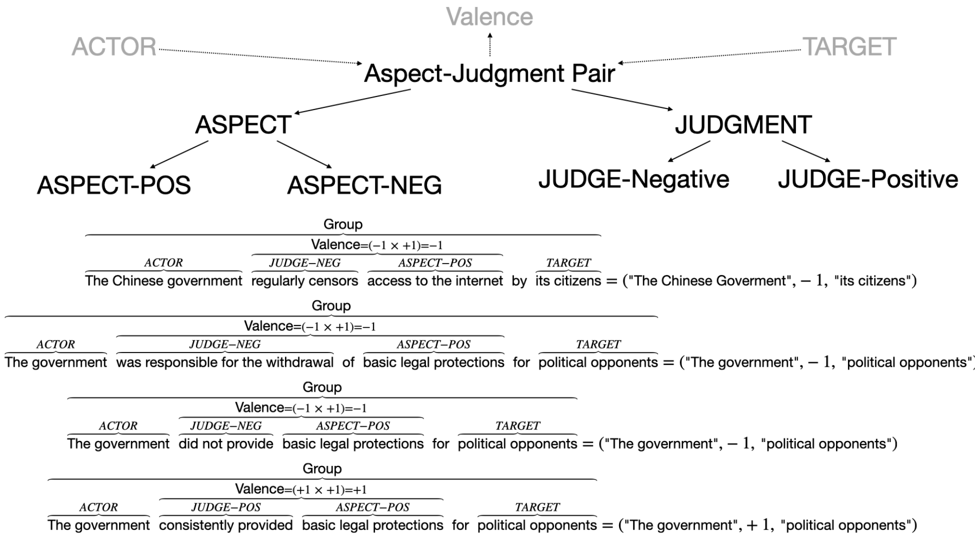



Figure A.5: Nested PULSAR 3.0 ontology.

*Note:* We build upon this ontology; the figure presents four examples from the human rights domain. The core components of the Aspect-Judgment pair are organized in a hierarchy at the top with arrows representing parent-child relations. In light gray we include the ACTOR and TARGET labels which could be linked to domain-specific knowledge graphs. In the lower part of the figure we include 4 examples that assign PULSAR label to specific tokens in the text (just above each example), how those tags are organized first into Aspect-Judgment pairs, then how the valence is calculated, and finally how the ACTOR, VALENCE, and TARGET are grouped together as edges. While these examples connect contiguous tokens and labels this is not required for the system.

To be clear, the system first links aspect and judgments together to create extracted Judgment-on-aspect pairs and valences – where the aspect is what is being judged, and the sentiment/valence reflects the evaluative stance on that aspect. Actor, Target and Judgment-holder are all grouped with their associated Judgment-on-aspect, and attributes.

This ontology can represent a rich array of language components and structures. For example, we can represent that the presence of the taking of a good thing (e.g., "The government was responsible for the withdrawal of basic legal protections for political opponents") or the negation of the presence of the giving of a good thing ("The government did not provide basic legal protections for political opponents") are both negative judgments. Further, the tagging of the tokens identify how and why PULSAR assigns an overall negative valence to a statement and how that is distinct from other positive combinations. The full ontology for the relevant PULSAR labels and those we add in this project is provided in the appendix.

The second component of the previous PULSAR system we update is a supervised learning system that applies the labels in the ontology to specific examples of text. This relies on at least some in-domain, hand-annotated examples and classifiers that can be traditional machine learning models – such as conditional random field (CRF) models or more recent LLM-based approaches. With traditional machine learning models, the text of politically-relevant documents are tokenized, parts of speech are tagged and syntatic dependencies parsed using tools such as the Stanford CoreNLP suite. LLM-based approaches, depending on the details often skip many of these steps and move to either prompt engineering if utilizing an GenAI API or formatting the input text to match available transformer-based architectures. The output of the second component is the sentence-level text automatically annotated with the relevant PULSAR labels. These supply the nodes of the graph that we are eventually going to produce, along with node-attributes (eg the original text, position, etc).[10]

The third component connects the annotations together, identifying which ACTOR, JUDGMENT, ASPECT, and TARGET (and other labels if included, see appendix) are related to each other. The core of this work is identifying the ASPECT-JUDGMENT pair, which JUDGMENT is communicated on which ASPECT. When this connection is made, the system next identifies whether the ACTOR is acting/causing the ASPECT-JUDGMENT pair and the TARGET is on the receiving end. Finally, the system calculate whether the judgment is positive or negative, whether the aspect is negative or positive and whether the joint ASPECT-JUDGMENT pair has an overall valence that is negative or positive. This calculation results from a simple multiplication where a positive label (either JUDGE or ASPECT) is assigned a +1 value and negative label a -1 and then related ASPECT-JUDGMENT pairs which share an edge are multiplied. Sentences can have multiple ASPECTS and JUDGMENTS, which makes the tracking of which judgments are related to which aspects, actors and targets a crucial component of the PULSAR system. The linking of ASPECT-JUDGMENT pairs and ACTORS/TARGETS can be done with rules based on dependency parsing or with supervised learning based on annotations. The most recent PULSAR system uses hand annotation and LLM-based edge identification. This final component supplies the edges that connect nodes of interest across the assigned labels, this can also be seen as grouping the relevant labels and associated tokens into *Speaker*(*Actor, Valence, Target*) triplets.

We provide simple examples in A.5 to build an intuition for how the computational and human processing fit together and provide insights on similarities and difference between sentences. The product of the three components is structured output that captures the speaker's judgment or position toward specific issues or topics. Crucially, PULSAR does not aim to capture all the semantic nuance in text but instead allow the valence of ASPECT-JUDGMENT pairs – what we term Judgment-on-aspect and map to moral good and moral bad components of populist narratives – to be compared across sentences and uses. We provide the resulting parenthetical representation of the structure (eg S$peaker(ACTOR, VALENCE, TARGET)$) for each example in A.5 on the far right. This data object can be understood as a directed graph with the middle information representing the edge attribute of the VALENCE. We also provide the annotation and calculation of the labels and edges in the context of the examples sentences. This illustrates a crucial benefit that we see in PULSAR. If someone wants to understand why one of the examples in A.5 is coded as the government doing a good or bad thing to a particular target, the system supplies the reasons based on the labels assigned, and the relations between then and estimated inferences can be evaluated in the context of the sentence. For example, in the last example, the government is judged to be doing something positive because PULSAR understands that there is a JUDGE-POS (+1) connected to a ASPECT-POS (+1), and more specifically (moving downward), the positive/presence/giving is found in the terms "consistently provided" while the positive aspect is found at the location in the text communicating "basic legal protections" which are deemed a good thing (ASPECT-POS). We label the edges with the *Group* tag to identify the edges in the context of the sentence. Despite using different language, the first three examples all express negative judgments with the the last sentence inferred to have a positive judgment. The system also allows researchers to aggregate and compare actors and targets using knowledge graphs and meta-data. For example, meta-data can identify that "The government" is referring to "The Chinese Government" and a knowledge graph can connect different government agencies and names to a common or related concepts.

[10] A PULSAR pipeline that is based on labeling with LLM models produces hard-classification, labeled as 1 or 0 currently while more traditional supervised learning approaches can output indices that range between 0 and 1 which researchers use to assign labels based on a label-specific threshold. We report on an LLM-based pipeline here.

Despite the fact that PULSAR was originally developed to process human rights documents, its architecture makes it well-suited for annotating political discourse more broadly (Greene, 2020) when interpretability and sub-sentence and sentence-level granularity is of value. In our study, an updated PULSAR system serves as the first stage of analysis, transforming raw political text into a network format that is both specific to the the expressed judgment in the text and general enough to compare meaning across sentences and speakers. This allows for more precise downstream tasks, including the identification of populist narratives within these structures.

### A.4.5 Our Approach

With an understanding of the PULSAR components and how they fit together, in this section we detail how we extend those components to measure populist narratives in general political speech. Our measurement approach closely follows the base PULSAR system in organizing an automated system that maps text into graphs/networks. Where we deviate from previous uses of PULSAR is that our overall task does not end with the measurement of the positive or negative valence of judgments – as in human rights allegations. Instead, we need to match the expressed graph structure of the extracted relations and tags to the structure of populist narratives – which have their own edges, relations, and valences. To accomplish this matching we need to a) represent prototypical, thin-centered populist narratives as networks – which we will treat as network motifs to match against; b) extend PULSAR to label actual text from speeches with the necessary labels and relations that comprise populist network motifs from step a (or do not match for non-populist speeches), and c) measure the degree of discrepancy between each populist network motif and the derived graph from sentences from political speeches.

We organize populist narratives, which are usually expressed as example text, into the specific sets of relations and nodes. These motifs serve as structured representations of populist rhetorical frames, linking together key components of evaluative language: whether an actor (who performs an action or causes what is being judged) represents a populist in-group ("the people" or related terms, many of which will be context-specific) or an out-group ("elites" or related terms, many of which will also be context-specific), whether the target (who is the recipient of the positive or negative actions being judged or is effected) represents the populist in-group or out-group, the aspect (the issue, thing, or concept being evaluated), the judgment (the evaluative expression on that aspect), and the judgment holder (the source of the opinion or judgment which can be distinct from the composer of the utterance). Crucially, core narrative components of thin-centered populism that had not previously been incorporated into PULSAR or other sentence-level annotation systems are the fact that actors or targets can be ascribed in-group or out-group-based roles which we refer to formally as *US* or *THEM* attributes. These optional attributes of ACTOR and TARGET labels capture the moral and identity boundaries that are central to populist rhetoric.

We define four core motifs from the direct speaker perspective (where the speaker is the judgment-holder) that we summarize with the notation: *(US, +, US)*, *(THEM, +, THEM)*, *(US, –, THEM)*, and *(THEM, –, US)*. In the *(US, +, US)* motif, or Empowerment narrative, the righteous or deserving group acts to help or protect itself—such as pledging to preserve American jobs for hardworking Americans. The *(THEM, +, THEM)* motif, or Elite Conspiracy Narrative, depicts the corrupt elite protecting or serving its own interests, as in claims that another politician is shielding special interests. The remaining two motifs capture adversarial dynamics between roles. In the *US, –, THEM,* motif, or Resistance narrative, the people confront or resist the corrupt elite—for instance, by voting, protesting, or pursuing legal accountability. Conversely, the *(THEM, –, US)* motif, or Victimization narrative, highlights scenarios where the elite actively harms the people, such as by sending jobs overseas or enacting self-serving policies at the public's expense. Examples of these motifs and the corresponding annotation logic are illustrated in Figure **??** and discussed further in Section A.4.5.

We also introduce second order populist motifs. These motifs rely on the fact that in political speech, a composer or speaker is often referring to someone else's judgments which they themselves might also be judging with a positive or negative valence. For example, in the sentence "Corrupt elites believes that billionaires should continue to enrich themselves", the inner aspect-judgment pair is captured by the triplet ("*billionaires*", +, "*themselves*"). Our system is tuned to recognize that "billionaires" in this context is an out-group THEM, and themselves is a co-reference to "billionaires", therefore we have $(THEM, +, THEM)$ which would match the first-order elite conspiracy motif. However, that judgment is being ascribed by the speaker to "Corrupt elites" and not themselves. Therefore the system labels "Corrupt elites" as the judgment holder for the extracted triplet and we have a nested second-order populist motif *(THEM, +, (THEM, +, THEM))*. Similar second-order populist motifs are *(THEM, +, (THEM, -, US))*, *(THEM, -, (US, +, US))*, *(THEM, -, (US, -, THEM))*. Similarly, second-order populist motifs also exist for the in-group being the judgment holder, *(US, -, (THEM, +, THEM))*, *(US, -, (THEM, -, US))*, *(US, +, (US, +, US))*, *(US, +, (US, -, THEM))*. We provide a list of the first and second-order populist motifs in the appendix, along with simple examples.[11] These populist motifs comprise a the set of structures that we match parsed empirical sentences against to detect populism.

In order to provide the information necessary to match the parsed actual speeches to the abstract populist narratives, the PULSAR system needed to be updated to measure US and THEM attributes of actors and targets and organize the actor, valence, target triplets by expressed judgment holder.[12] In addition, because the technology of using text as data has improved remarkably since the start of the PULSAR system, we make several additional improvements. First, unlike the original version of PULSAR, we do not rely on fixed grammatical structures or rules to assign roles or label objects. Instead, we train models that are more flexible than previous bag-of-word representations of text and rule-based algorithms.[13] Each sentence is parsed into a network of interrelated elements, guided by a smaller number of hand-annotated examples used to prompt LLMs. As one example of how this change improves performance, an ACTOR is not necessarily the syntactic subject in a sentence, nor is the target always the object, rule-based algorithms using traditional dependency parsing have difficulty unwinding these exceptions.[14] Instead, we extract spans of text tagged as actor, target, aspect, or judgment based on their relationships in the narrative. These are connected via numbered grouping tags, with each element carrying its own attributes, such as

[11]Our notation can be extended to higher order relations – such as structures that resemble "They say that I say that they say ...". For brevity we do not exhaust an n-th order motifs, but stop at third-order judgments.

[12]If no judgment holder is provided, the system ascribed the judgment to the composer/speaker which is included in meta-data. In addition, the use of first person pronouns are used by the system to assign the judgment holder if available.

[13]Put another way, we follow the innovations in LLMs and move away from lexical vectorization of text into computable quantities and towards models based on transformer architectures. Similarly, we also replace the use of a moderate number of known-hand-crafted rules that are applied algorithmically to create edges with LLMs technologies that leverage the relationship between terms and the greater context of the sentence.

[14]For example, in the sentence "Real Americans are suffering from elite greed everyday", "Real Americans" is the subject but our system codes them as the TARGET since something negative happened to them, as "are suffering" is labeled as an ASPECT-JUDGEMENT pair with a negative valence. Further "elite" is labeled as the ACTOR of this group because they are the cause what is being judged. For completeness, we note that the system would also code "elite" as the ACTOR for "greed" as another ASPECT-NEG. This highlights the fact that our system is multi-label and multi-group and thus one phrase or token can play multiple roles across different groupings/graphs.

valence or group role, based on surrounding lexical cues.[15]

### A.4.6 Handling Complex Cases

**Nested Aspects:** Our system also captures nested narratives within judgment-on-aspects, which are distinct from second-order judgment holders. Nested judgments-on-aspects within aspects allow us to identify when a speaker evaluates not only an action, but also the motivations or effects behind that action. Take, for instance:

*"I am fighting against billionaires who raid our country."*

**Outer Judgment**

$$\overbrace{\overbrace{\texttt{I}}^{\text{ACTOR (US)}}\ \overbrace{\overbrace{\texttt{am fighting against}}^{\text{JUDGMENT (-1)}}\ \overbrace{\texttt{billionaires who raid}\ \overbrace{\texttt{our country}}^{\text{TARGET (US)}}}^{\text{ASPECT (-1), Group 2}}}^{\text{valence=+1}}}^{\text{Group 1:}(US,+,US)}$$

**Inner Judgment**

$$\overbrace{\overbrace{\texttt{billionaires}}^{\text{ACTOR (THEM)}}\ \overbrace{\overbrace{\texttt{who raid}}^{\text{JUDGMENT (-1)}}\ \overbrace{\texttt{our country}}^{\text{ASPECT (+1) — TARGET (US)}}}^{\text{valence=-1}}}^{\text{Group 2: (THEM, -1, US)}}$$

In this example, the outer layer evaluates the speaker's opposition to harmful actors, while the inner layer attributes harm to those actors. The nesting allows the system to represent how the speaker (*US*) positions themselves against *THEM* (the billionaires) by condemning their actions toward a valued object (our country). The structure supports the identification of populist themes such as moral polarity and group-based conflict.

**Non-Speaker Judgment-holders:** Our approach effectively captures judgments attributed to individuals other than the speaker, a critical feature for identifying indirect populist assertions or rebuttals. Political discourse frequently involves rhetorical structures such as "my opponent says... but I say...," where speakers reference or critique judgments voiced by others. Ignoring this distinction could misattribute sentiments or misinterpret narrative nuances central to populist rhetoric. Consider the following example from Bernie Sanders:

"*My opponent says that as a senator, she told bankers to 'cut it out' and end their destructive behavior, but, in my view, establishment politicians are the ones who need to 'cut it out.' The reality is that Congress doesn't regulate Wall Street. Wall Street and their lobbyists regulate Congress.* "

$$\overbrace{\texttt{My opponent}}^{\text{ACTOR (Residual), JUDGMENT HOLDER (Non-speaker)}}\qquad \texttt{...}$$

$$\overbrace{\overbrace{\texttt{bankers}}^{\text{TARGET (THEM)}}\ \overbrace{\texttt{to `cut it out' and end}}^{\text{JUDGMENT (-1)}}\ \overbrace{\texttt{their destructive behavior}}^{\text{ASPECT (1)}}}^{\text{Valence=-1}}$$

## A.5 Annnotation Process

We use the Prodigy system to label sentences, review the draft annotations, and reconcile any disagreements. Currently, in the prototype system we have only one coder per sentence which is then reviewed by a team of the three authors. Figure A.6 provides a screenshot of the Prodigy

[15] As we detail below, the dependencies between labeled tokens/phrases are implemented as group-labels that can be non-contiguous.

interface we use to annotate a sentence. After annotation, we have a json-new-line (jsonl)-formatted dataset that can be used for supervised learning and in prompts to LLMs.

### A.5.1 Prompt Engineering for Populist Motif Detection

This study adopts a structured prompt[16] engineering approach to extract narrative motifs from political speech using LLMs. Our goal is to identify patterns characteristic of populist discourse—particularly those portraying in-group actors ("US") as victims of actions taken by out-group actors ("THEM"). To this end, we design prompts that instruct the model to annotate each sentence according to a predefined schema that captures core elements of moralized political judgments. The model is also instructed to assign a valence to both the judgment and the aspect, enabling us to infer the valence of the action, and to attribute a role (US, THEM, or NA) to both actor and target.

Our prompt engineering strategy aligns with current best practices. First, we employed few-shot prompting (Brown et al., 2020), providing curated input-output examples to help the model internalize the annotation task.[17] These examples were selected to cover common syntactic structures and sources of ambiguity, improving generalization. Second, we followed the principle of modular decomposition (Zhou et al., 2022; Sahoo et al., 2025; OpenAI, 2025), breaking down the annotation task into subtasks such as identifying grammatical subjects, distinguishing aspects from targets, and independently assigning valence. Rather than instructing the model to label long, undifferentiated spans of text, we guided it to extract specific evaluative components.

To enforce consistent outputs, we instructed the model to return only labeled content. We explicitly permitted overlap between aspect and target spans—for instance, in "They attacked us," where "us" functions as both the evaluated entity and the affected group. These constraints reflect emerging conventions in instruction formatting and schema-based prompting (Geroimenko, 2025; Cleary, 2025).

To operationalize this prompt-based annotation process, we interfaced with the OpenAI API using a Python script that submitted each sentence to the `gpt-4o-mini-2024-07-18` chat endpoint. Each API call included a system message defining the model's role, followed by a user message comprising the full prompt, including task instructions, definitions, and few-shot examples. We set the temperature parameter to zero to approach deterministic, reproducible outputs, acknowledging that AI-generated responses are inherently probabilistic.

The development of the prompt followed an iterative and hybrid self-refinement process, inspired by recent work on Self-Refine prompting (Madaan et al., 2023). In our implementation, the LLM was used not only to generate labeled outputs but also to support prompt revision. After testing early versions on a silver-standard annotated dataset, we analyzed recurring errors—such as misassignments of valence, incorrect segmentation of targets, or confusion between modifiers and intrinsic aspect valence. These were reformulated into diagnostic cases and presented to the model via meta-prompts to support the refinement of definitions, examples, and instructions. This strategy mirrors what Reynolds and McDonell (2021) describe as metaprompt programming—using the model itself to iteratively revise the prompts that structure its behavior. While critique was guided by a human researcher, the model played an active role in suggesting revised prompt components, functioning as both annotator and co-designer.

[16] A prompt serves as an input provided to a generative AI model to guide the generation of its output (Schulhoff et al., 2025; Brown et al., 2020).

[17] As Reynolds and McDonell (2021) explain, few-shot examples activate pre-trained capabilities rather than teaching new tasks.

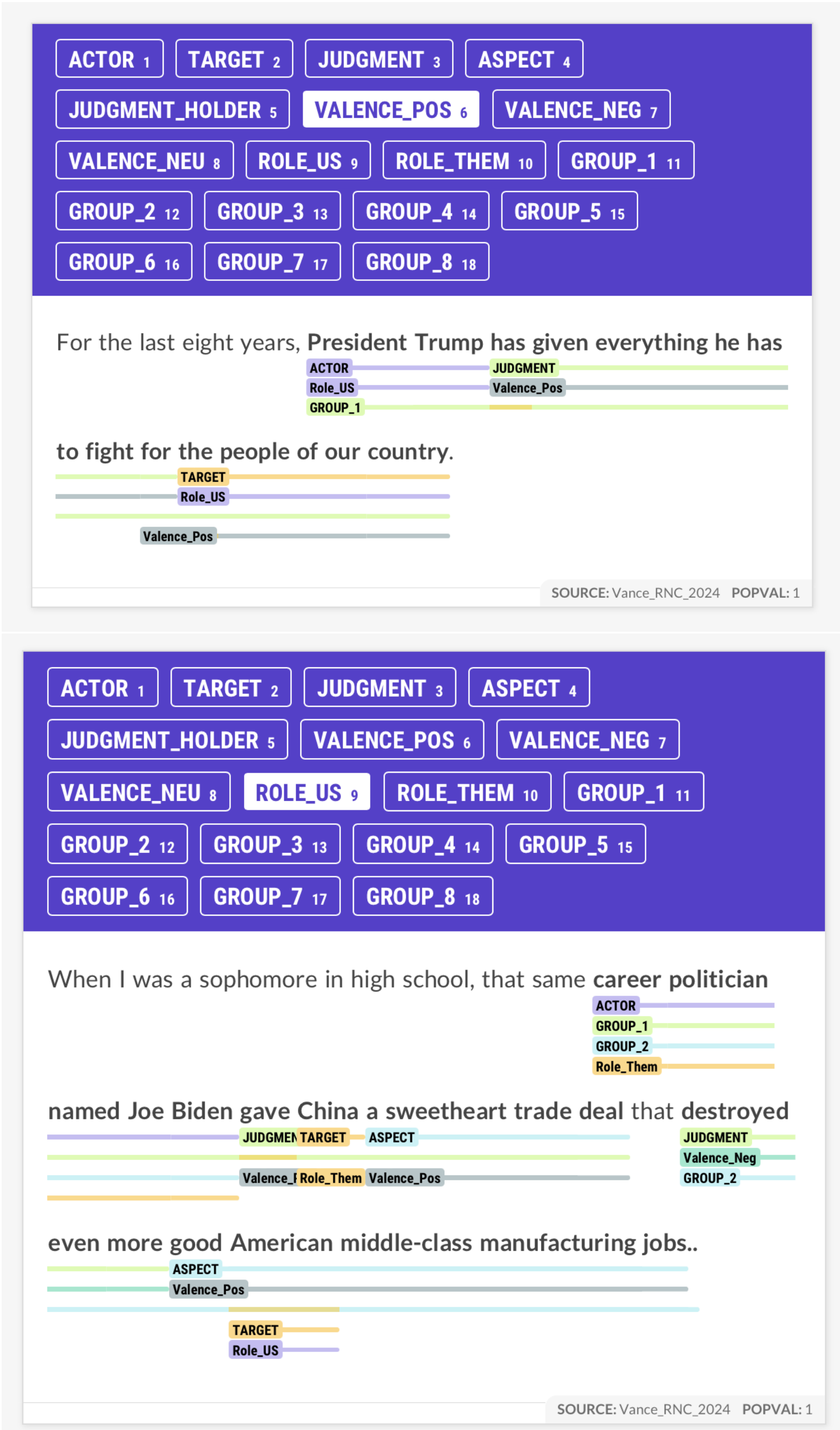



Figure A.6: Screenshots from human annotation interface for populist PULSAR labeling tasks and grouping illustrating two examples. The second example includes non-contiguous groupings.

To evaluate whether prompt iterations improved model performance, we implemented a GED–based evaluation metric that compared model-generated annotations with human-annotated graphs (see subsubsection A.5.3). This enabled us to measure whether changes in the prompt brought model outputs into closer alignment with our annotation framework, offering a more robust performance indicator than token-level accuracy. This evaluation approach reflects recent calls for tailored, task-specific benchmarks in prompt engineering research (Schulhoff et al., 2025).

In sum, our approach combines principles of few-shot prompting, modular decomposition, and iterative prompt refinement within a clearly defined task context. This methodology not only improved the reliability of model outputs but also illustrates how domain-specific applications can be advanced by integrating best practices from the evolving field of prompt engineering.

### A.5.2 Synthetic Data

| **Synthetic Variants of Political Speech Sentence** |
|---|
| 1. Every real American is entitled to liberty. |
| 2. Each true American is entitled to liberty. |
| 3. All genuine Americans are entitled to liberty. |
| 4. Every authentic American has the right to liberty. |
| 5. Each legitimate American deserves liberty. |
| 6. All real patriots have a right to liberty. |

Table A.3: Synthetic variants preserving the meaning of a sentence while introducing syntactic and lexical diversity.
Note: Original sentence was "*Every real American deserves freedom.*"

### A.5.3 Matching Extracted Narrative Structures to Motifs Using a Graph Edit Distance

As noted in the main text, we then use Graph Edit Distance to make comparisons and match extracted specific structures to the general populist narrative motifs.

Traditional analyses of political speech often rely on exact matching—searching for specific keywords or fixed patterns—which fails to account for the flexible and often implicit nature of populist rhetoric. Populist motifs prioritize relational structures (e.g., people vs. elites) over precise lexical forms, and the same underlying logic can surface through varied or indirect expressions. To capture these patterns, we implement a custom Graph Edit Distance (GED) metric that quantifies structural similarity between annotated motifs, enabling principled comparisons even when wording or syntax differs.

Each sentence is reduced to a triplet form: `[ACTOR role, valence score, TARGET role]`, where valence is computed as the product of the judgment and aspect valence. We refer to this triplet as reduce graph representation. Roles are numerically encoded (e.g., `US = 0`, `THEM = 2`, `Residual = 1`). GED is then computed as the minimum cost of transforming one triplet into another, with penalties designed to reflect rhetorical severity. For example, reversing polarity (positive to negative) incurs a maximum penalty of 4; substituting an explicit actor with an ambiguous one incurs a minimal penalty of 1. Switching from US to THEM is assigned a penalty of 2. GED thus mirrors human judgment, recognizing rhetorical similarity even when surface forms differ. Table A.4 summarizes GED penalities.

Table A.4: Conceptual Explanation of GED Penalization System

| Penalty Value | Explanation |
|---|---|
| 0 | No change: Actor, valence, and target match exactly. |
| 1 | Minor ambiguity: Explicit actor replaced by ambiguous actor (Residual). |
| 2 | Actor/Target substitution: Switching between US and THEM roles. |
| 4 | Valence reversal: Changing polarity from positive to negative or vice versa. |

This structure-aware method enables us to go beyond superficial similarity, identifying when two sentences convey substantively aligned motifs. It also supports evaluation of annotation quality by quantifying how closely model-generated labels match human-coded structures. GED has a long-standing role in pattern recognition and computational linguistics for inexact graph matching (Blumenthal et al., 2021), and here it serves both analytical and evaluative functions: identifying clustering and divergence in populist discourse and assessing the reliability of automated motif extraction.

To illustrate, consider three syntactically distinct but structurally aligned sentences: "We will bring back our borders," "We will bring back our jobs," and "We will bring back our wealth." All share the same reduced graph: `[0,1,0]`, indicating a US actor, positive valence, and US target. The GED between them is zero (and it is also zero by definition to the "empowerment" motif). In contrast, "I know it's all about making America great again for all Americans" conveys a similar message but with less explicit actor framing. We encode this as `[1,1,0]`, with a minimal penalty of 1 reflecting the ambiguous actor role. This shows how our metric captures semantic proximity despite variation.

Now consider: "The insiders wrote the rules of the game to keep themselves in power and in the money." This sentence reflects a THEM-supporting-THEM motif: `[2,1,2]`. Compared to the previous US-supporting-US examples, the GED is 4—capturing a full actor and target role reversal. This higher distance reflects a deeper rhetorical divergence between the two structures, demonstrating how our approach distinguishes narrative similarity from ideological opposition.

## A.6 Fine-tuning Details

### A.6.1 Sentence-Level Multi-Task Classification

**Model Architecture** We fine-tune a transformer-based DistilBERT model for multi-task sentence classification. The architecture employs a shared DistilBERT encoder that generates contextual representations of sentences. The output of the encoder, specifically the hidden representation of the [CLS] token, is fed into four parallel classification heads. Each head consists of a single linear layer producing logits corresponding to four distinct labeling tasks:

- **Actor Role**: Identifies whether the sentence's actor is in-group (US), out-group (THEM), or residual.
- **Target Role**: Identifies whether the sentence's actor is in-group (US), out-group (THEM), or residual.
- **Judgment Valence**: Determines if the judgment expressed is positive, negative, or neutral.

- **Aspect Valence**: Evaluates if the mentioned aspect or value is framed positively, negatively, or neutrally.

**Training Procedure** The model is jointly trained across all four tasks using multi-task learning. During fine-tuning, we minimize the average cross-entropy loss from the four classification heads simultaneously, allowing the shared encoder to develop comprehensive sentence representations. Training is conducted with the HuggingFace Trainer API using standard hyperparameters: 3 epochs, batch size of approximately 8, and a learning rate of 2e-5. Model performance is assessed using accuracy and F1-score metrics on a held-out validation set for each classification task.

Inference At inference, the model outputs logits for each classification head, from which we select the highest-scoring class (via argmax) to obtain the final labels. The resulting classification tuple for a sentence includes labels for actor role, target role, judgment valence, and aspect valence, providing a concise characterization of the political statement.

### A.6.2 Token-Level Sequence Labeling

**Objective and Tagging Scheme** To achieve granular text analysis, we perform token-level sequence labeling, akin to Named Entity Recognition (NER). Tokens within sentences are annotated with category-specific labels or an "O" tag if irrelevant. Specifically, the tagging categories include:

- **Actor Identification**: Tokens tagged as ACTOR-US or ACTOR-THEM.
- **Target Identification**: Tokens tagged as TARGET-US or TARGET-THEM.
- **Judgment Identification**: Tokens reflecting sentiment tagged as JUDGMENT-POS or JUDGMENT-NEG.
- **Aspect Identification**: Tokens describing values or aspects tagged as ASPECT-POS or ASPECT-NEG.

Non-relevant tokens receive an "O" tag.

**Data Processing** Sentences are tokenized with the BERT tokenizer, converting words into subword tokens. Labels align with subword tokens, assigning original labels to the initial subword token and marking subsequent subword tokens with a special label (-100) to avoid inflated loss calculations. Sequences are padded or truncated to a fixed length (128 tokens), and padding tokens are assigned the -100 label to ensure proper loss calculation.

**Model Training and Configuration** We fine-tune separate BERT-based token classification models for each category. Each model uses a BERT encoder with a linear classification head predicting token-level labels. Models initialize from pretrained bert-base-uncased weights, with randomly initialized linear layers. Training typically involves 3 epochs, batch size around 16, and a learning rate of 2e-5. The cross-entropy loss function excludes tokens labeled as -100.

**Inference and Evaluation** During inference, each token-level model predicts labels for input tokens. Extracted spans are identified by grouping contiguous tokens tagged within the same category (e.g., ACTOR-US). Model performance is evaluated using precision, recall, and F1 metrics from the seqeval library, comparing model outputs against annotated silver-standard data.

### A.7 Fine-Tuning Metrics

Preliminary results for exact token-level evaluation in a held-out test set for the fine-tuning model is provided in table Table A.5. The overall precision and recall are lower for the token-level prediction tasks because these are much more difficult tasks. We continue to improve the system and add more training data and run experiments.

| Label | Attribute | Precision | Recall | F1 |
|---|---|---|---|---|
| Actor | Them | 0.46 | 0.46 | 0.46 |
| | Residual | 0.20 | 0.27 | 0.23 |
| | Us | 0.67 | 0.72 | 0.70 |
| Target | Them | 0.27 | 0.33 | 0.30 |
| | Residual | 0.00 | 0.00 | 0.00 |
| | Us | 0.33 | 0.30 | 0.31 |
| Judgment | -1 | 0.05 | 0.04 | 0.05 |
| | +1 | 0.15 | 0.18 | 0.16 |
| Aspect | -1 | 0.08 | 0.11 | 0.10 |
| | +1 | 0.13 | 0.16 | 0.14 |

Table A.5: Token-level, test-set evaluation metrics for fine-tuned model.

### A.8 Examples of Our Approach

To further illustrate our labeling system in action, we apply it to two paragraphs from Donald Trump's December 2023 speech in Reno, Nevada. Each sentence is annotated using our Populist Annotation Scheme, generating a reduced graph that encodes the group alignment of the actor and target (US, THEM, or unspecified) and the valence of the judgment. A sentence is classified as populist if its reduced form corresponds to one of the core populist motifs. A sentence-by-sentence breakdown of the reduced graphs for these paragraphs is available in Appendix Table A.6–A.7.

The first paragraph, which focuses on immigration, foreign policy, and domestic priorities, features a series of declarative judgments that consistently portray the out-group (e.g., "Crooked Joe Biden," "environmental maniacs") as prioritizing others over "our workers," "our families," or the country itself. These sentences frequently align with the (THEM, -, US,) or (US, - THEM,) motifs, both central to populist discourse. Out of 13 sentences, 10 are classified as populist according to our annotation criteria. This segment exemplifies what we refer to as victimization and resistance narratives—frames in which the in-group is depicted as under attack or neglected by corrupt or hostile elites.

The second paragraph is structured around an anecdote involving pollsters and strategic disinformation. Although less programmatic than the first, it contains references to manipulative elites attempting to suppress voter participation. These statements also map onto the (THEM, -, US) motif, with the in-group implied to be Trump and his supporters, and the out-group consisting of media or polling institutions. In this case, 4 out of 9 sentences are classified as populist (blue color).

> I will immediately restore and expand the Trump travel ban on an entry from terror plague countries, and I will implement strong ideological screening for all immigrants. If you hate America, if you want to abolish Israel, if you sympathize with Jihadists, then

> we don't want you in our country. We don't want you. Crooked Joe Biden puts China first, puts Asia first, puts Ukraine first, puts illegal aliens first, environmental maniacs first. Everyone else first. He puts everybody first. That's bad. But he puts America last. He puts Nevada last. He puts our workers last. He frankly puts unions last. Look what he did with the car. They're going to make all electric cars in our country, but not when I'm in there. I'll end that the first day. He puts our families last. (...)
>
> I asked John McLaughlin the great pollster, I said, "Why would they do that? I'm winning in Wisconsin." He said, "They do that because that way people don't vote." I said, "What do you mean? Well, why don't they do it like at 5:00? Because they do get killed when they lose." He said, "They don't care about that. At 17 people won't vote. They'll say, we love Trump, but we're not going to vote for him. At 5:00 they say, he has a chance. We'll vote." I said, "Wow, that's really sick, huh?" "That's what it is. That's how bad they are that we won Wisconsin." "They had us down 17 and we won Wisconsin, and then we did, frankly, better the second time where we did better in most places the second time." People say, "How'd you do the second time?" I say, "We did much better."

Turning our attention to a leader typically labeled as non-populist by previous classification methods, we find that Barack Obama's August 2024 speech in Chicago nonetheless contains a high proportion of populist motifs: 7 out of 10 sentences follow one of the defined populist structures (see Appendix Table A.8). Most prominently, Obama relies on the (US, +, US,) frame, which emphasizes in-group moral strength and unity. Although his tone remains less adversarial than that found in more traditional populist rhetoric, the narrative structure—highlighting shared struggle, virtuous leadership, and the need to defend national values—demonstrates how non-populist leaders can strategically draw on populist elements to mobilize support and establish moral contrast.

> At a time when millions of our fellow citizens were sick and dying, we needed a leader with the character to put politics aside and do what was right. At a time when our economy was reeling, we needed a leader with the determination to drive what would become the world's strongest recovery—15 million jobs, higher wages, lower health care costs. And at a time when the other party had turned into a cult of personality, we needed a leader who was steady, and brought people together, and was selfless enough to do the rarest thing there is in politics: putting his own ambition aside for the sake of the country.
>
> History will remember Joe Biden as an outstanding president who defended democracy at a moment of great danger. I am proud to call him my president, but I am even prouder to call him my friend.
>
> Now the torch has been passed. Now it's up to all of us to fight for the America we believe in. And make no mistake: it will be a fight. For all the incredible energy we've been able to generate over the last few weeks, for all the rallies and the memes, this will still be a tight race in a closely divided country—a country where too many Americans are still struggling. Where a lot of Americans don't believe government can help.

### A.8.1 Trump 12/2023, Reno

I will immediately restore and expand the Trump travel ban on an entry from terror plague countries, and I will implement strong ideological screening for all immigrants. If you hate America, if you want to abolish Israel, if you sympathize with Jihadists, then we don't want you in our country. We don't want you. Crooked Joe Biden puts China first, puts Asia first, puts Ukraine first, puts illegal aliens first, environmental maniacs first. Everyone else first. He puts everybody first. That's bad. But he puts America last. He puts Nevada last. He puts our workers last. He frankly puts unions last. Look what he did with the car. They're going to make all electric cars in our country, but not when I'm in there. I'll end that the first day. He puts our families last.

| Sentence | Reduced Graph | Populist Motif | Notes |
|---|---|---|---|
| I will immediately restore and expand the Trump travel ban on entry from terror plague countries, and I will implement strong ideological screening for all immigrants. | [1, 1, 1] | 0 | No explicit US/THEM roles or oppositional framing. Policy stance, but not populist in structure. |
| If you hate America, if you want to abolish Israel, if you sympathize with Jihadists, then we don't want you in our country. | [0, -1, 2] | 1 | (US, -, THEM,); clear populist oppositional structure. |
| We don't want you. | [0, -1, 2] | 1 | Same motif as above, direct rejection of out-group. |
| Crooked Joe Biden puts China first, puts Asia first, puts Ukraine first, puts illegal aliens first, environmental maniacs first. | [2, 1, 2] | 1 | (THEM, +, THEM,); classic elite conspiracy motif. |
| Everyone else first. He puts everybody first. That's bad. | [2, 1, 2] | 1 | Reinforces (THEM, +, THEM,) with negative valence. |
| But he puts America last. | [2, -1, 0] | 1 | (THEM, -, US,); populist victimization. |
| He puts Nevada last. | [2, -1, 0] | 1 | Local in-group harmed. |
| He puts our workers last. | [2, -1, 0] | 1 | (US, -, THEM,) |
| He frankly puts unions last. | [2, -1, 0] | 1 | Again, (THEM, -, US,). |
| Look what he did with the car. | [2, -1, 1] | 0 | Possible criticism, but ambiguous target—role is unspecified. |
| They're going to make all electric cars in our country, but not when I'm in there. | [2, -1, 0] | 1 | Speaker positions (THEM, -, US,); promises reversal. |
| I'll end that the first day. | [0, 1, 1] | 0 | Supportive US action, but no THEM present. Not inherently populist. |
| He puts our families last. | [2, -1, 0] | 1 | Again, (THEM, -, US,) (families). |

Table A.6: Sentence-Level Annotation of Populist Motifs in Trump Speech

### A.8.2 Trump 12/2023, Reno

I asked John McLaughlin the great pollster, I said, "Why would they do that? I'm winning in Wisconsin." He said, "They do that because that way people don't vote." I said, "What do you mean? Well, why don't they do it like at 5:00? Because they do get killed when they lose." He said, "They don't care about that. At 17 people won't vote. They'll say, we love Trump, but we're not going to vote for him. At 5:00 they say, he has a chance. We'll vote." I said, "Wow, that's really sick, huh?" "That's what it is. That's how bad they are that we won Wisconsin." "They had us down 17 and we won Wisconsin, and then we did, frankly, better the second time where we did better in most places the second time." People say, "How'd you do the second time?" I say, "We did much better."

| Sentence | Reduced Graph | Populist Motif | Notes |
|---|---|---|---|
| I asked John McLaughlin the great pollster, I said, "Why would they do that? I'm winning in Wisconsin." | [1, 1, 1] | 0 | No actor is clearly marked as US or THEM → *non-populist.* |
| **He said, "They do that because that way people don't vote."** | **[2, -1, 0]** | **1** | **THEM (media or political elites) are harming US (Trump voters) by discouraging participation.** |
| He said, "What do you mean? Well, why don't they do it like at 5:00? Because they do get killed when they lose." | [2, -1, 2] | 0 | Actor and target are both THEM, judgment is negative — *not a populist motif.* |
| **He said, "They don't care about that. At 17 people won't vote. They'll say, we love Trump, but we're not going to vote for him."** | **[2, -1, 0]** | **1** | **THEM (pollsters/media) are again undermining US (Trump supporters) — matches the victimization motif.** |
| At 5:00 they say, he has a chance. We'll vote. | [0, 1, 1] | 0 | Actor is US, supporting something good, but not [0, 1, 0], so *not a Pop. motif.* |
| I said, "Wow, that's really sick, huh?" | [1, -1, 2] | 0 | Actor is ambiguous, not explicitly US or THEM, so *not a Pop. motif.* |
| **That's what it is. That's how bad they are that we won Wisconsin.** | **[2, -1, 0]** | **1** | **THEM (media) tried to suppress US (Trump voters), but failed — matches THEM harming US motif.** |
| **They had us down 17 and we won Wisconsin, and then we did, frankly better the second time where we did better in most places the second time.** | **[2, -1, 0]** | **1** | **Again, THEM (pollsters/media) misrepresented US (Trump) → populist motif.** |
| People say, "How'd you do the second time?" I say, "We did much better." | [1, 1, 0] | 0 | Speaker praises in-group, but not in [0, 1, 0] format — *non-populist.* |

Table A.7: Populist Motif Analysis of Trump's Speech Segment

### A.8.3 Obama 8/2024, Chicago

At a time when millions of our fellow citizens were sick and dying, we needed a leader with the character to put politics aside and do what was right. At a time when our economy was reeling, we needed a leader with the determination to drive what would become the world's strongest recovery—15 million jobs, higher wages, lower health care costs. And at a time when the other party had turned into a cult of personality, we needed a leader who was steady, and brought people together, and was selfless enough to do the rarest thing there is in politics: putting his own ambition aside for the sake of the country. History will remember Joe Biden as an outstanding president who defended democracy at a moment of great danger. I am proud to call him my president, but I am even prouder to call him my friend. Now the torch has been passed. Now it's up to all of us to fight for the America we believe in. And make no mistake: it will be a fight. For all the incredible energy we've been able to generate over the last few weeks, for all the rallies and the memes, this will still be a tight race in a closely divided country—a country where too many Americans are still struggling. Where a lot of Americans don't believe government can help.

### A.8.4 Trump 9/2016, Florida

| Sentence | Reduced Graph | Populist Motif | Notes |
|---|---|---|---|
| At a time when millions of our fellow citizens were sick and dying, we needed a leader with the character to put politics aside and do what was right. | [0, 1, 0] | 1 | US (we) supporting US (a good leader); expresses moral evaluation and positive self-reference. |
| At a time when our economy was reeling, we needed a leader with the determination to drive what would become the world's strongest recovery -15 million jobs, higher wages, lower health care costs. | [0, 1, 0] | 1 | Positive framing of US leadership in a crisis; in-group affirming message. |
| And at a time when the other party had turned into a cult of personality, we needed a leader who was steady, and brought people together, and was selfless enough to do the rarest thing there is in politics: putting his own ambition aside for the sake of the country. | [2, -1, 0] | 1 | THEM (other party) framed as harmful; in contrast, US represented by Biden is moral and selfless. |
| History will remember Joe Biden as an outstanding president who defended democracy at a moment of great danger. | [0, 1, 0] | 1 | US leader defending core values; positive valence and in-group affirmation. |
| I am proud to call him my president, but I am even prouder to call him my friend. | [0, 1, 0] | 1 | Personal endorsement of in-group figure; supportive framing of US actor. |
| Now the torch has been passed. | [1, 0, 0] | 0 | Metaphorical; lacks explicit judgment or clear group alignment for actor/target. |
| Now it's up to all of us to fight for the America we believe in. | [0, -1, 2] | 1 | US actor in opposition to THEM (implied threat to American values); a call to defend the in-group. |
| And make no mistake: it will be a fight. | [0, -1, 2] | 1 | Reinforces the previous adversarial frame; threat-oriented language. |
| For all the incredible energy we've been able to generate over the last few weeks, for all the rallies and the memes, this will still be a tight race in a closely divided country — a country where too many Americans are still struggling. | [0, -1, 0] | 0 | US is struggling, but no clear external adversary is identified; lacks THEM framing. |
| Where a lot of Americans don't believe government can help. | [1, -1, 1] | 0 | Critical of belief or sentiment, not directly assigning blame; judgment not clearly attached to a THEM actor. |

Table A.8: Sentence-level Populist Analysis of Barack Obama Speech

| Sentence | Reduced Graph | Our Approach | Dictionary | Dai and Kustov (2022a) | Di Cocco and Monechi (2021) | Comments |
|---|---|---|---|---|---|---|
| Change is coming. | [1, 1, 1] | 0 | 0 | 1 | 1 | Metaphorical; lacks explicit adversarial framing. |
| All the people who've rigged the system for their own personal benefit are trying to stop our change campaign because they know that their gravy train has reached its last stop. | [2, 1, 2], [2, -1, 0] | 1 | 0 | 1 | 1 | THEM (corrupt elite) vs. US (campaign for change); adversarial framing. |
| It's your turn now. | [1, 1, 1] | 0 | 0 | 1 | 1 | Encouraging but lacks clear adversarial or populist framing. |
| This is your time. | [1, 1, 1] | 0 | 0 | 1 | 1 | Affirmative but lacks populist motif or adversarial element. |
| The fact that so many encrusted old political insiders oppose our campaign is the best proof you will ever need that we are fighting for real change not partisan change. | [2, -1, 0], [0, 1, 1] | 1 | 0 | 1 | 1 | THEM (political insiders) negatively framed; US positively framed as fighting for real change. |
| We are fighting for all Americans Democrats, Republicans, Independents, Conservatives, Liberals who've been failed by this corrupt system. | [0, 1, 0], [2, -1, 0] | 1 | 1 | 1 | 1 | THEM (corrupt system) harming US (all Americans); inclusive populist framing. |
| We're fighting for everyone who doesn't have a voice. | [0, 1, 0] | 1 | 0 | 1 | 0 | US championing marginalized groups; clear populist framing. |
| We're also fighting for every region of this country. | [0, 1, 0] | 1 | 0 | 1 | 0 | US advocating inclusively for entire nation; populist unity motif. |
| For every part of Florida, and every part of America. | [1, 1, 1] | 0 | 0 | 1 | 1 | Broad unity but lacks adversarial or explicit populist frame. |
| From Pensacola to Pittsburgh, from Baltimore to Baton Rouge, we are fighting for every last city and every last person in this country. | [1, 1, 1] | 0 | 0 | 1 | 1 | Populist inclusiveness and national unity; US group positively affirmed. |
| Hillary Clinton is the candidate of the past. | [2, -1, 1] | 0 | 0 | 1 | 1 | Critical framing but lacks broader populist appeal or group alignment. |
| Ours is the campaign of the future. | [0, 1, 1] | 0 | 0 | 1 | 1 | Affirmative future-oriented framing without clear populist motif. |
| In this future, we are going to pursue new trade policies that put American workers first and that keep jobs in our country. | [0, 1, 0], [0, 1, 0] | 1 | 0 | 1 | 0 | US group positively framed; defending economic interests against implicit adversary. |
| All the people who got NAFTA wrong, and China wrong, and who are trying to give us the Trans-Pacific Partnership are the same failed voices pushing for Hillary Clinton. | [2, 1, 0] | 1 | 0 | 1 | 1 | THEM (failed elites) negatively framed against US interests. |
| Our trade deficit with the world is now nearly $800 billion dollars. | [1, 1, 1] | 0 | 0 | 1 | 1 | Economic critique without explicit populist framing. |
| We've lost one-third of our manufacturing jobs since Bill and Hillary Clinton gave us NAFTA. | [0, 1, 0], [2, -1, 0] | 1 | 0 | 1 | 1 | THEM (elite policy makers) blamed for negative economic outcomes; populist framing. |
| China is manipulating its currency and taking our jobs. | [2, -1, 0], [2, -1, 0] | 1 | 0 | 1 | 0 | THEM (foreign adversary) harming US economic interests; populist framing. |
| We are going to stop companies from leaving our country and keep those jobs right here in America. | [0, -1, 2], [0, 1, 0] | 1 | 0 | 1 | 0 | US group positively advocating against corporate adversaries; populist economic nationalism. |
| The era of economic surrender is over. | [1, 1, 1] | 0 | 0 | 1 | 1 | Strong rhetoric but lacks clear populist adversary. |

Table A.9: Sentence-level Populist Analysis with Comparative Approaches Donald Trump
Note: *Remarks at a Rally at the Pensacola Bay Center in Pensacola, Florida. September 9th, 2016.*