

Which Populism, Where? Measuring Narrative Strategy at the Sentence Level*

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

April 5, 2026

Abstract

We introduce Populist-PULSAR, an interpretable and scalable computational system for measuring populist rhetoric at the sub-sentence level. Building on thin-centered theories, the method represents populist discourse as graph-based motifs that capture who acts on whom and with what moral valence. Populist-PULSAR integrates human annotation, LLMs, network-based representation, and synthetic data augmentation to identify four recurrent motifs: victimization, empowerment, resistance, and elite conspiracy. Fine-grained measurement allows researchers to study the intensity of populism, and the substantive issues associated with motifs, such as corruption and jobs. Applied to Trump’s 2016 general-election speeches, the method shows that populist rhetoric is more prevalent in competitive and economically vulnerable contexts. Elite-conspiracy narratives dominate in competitive states, whereas empowerment narratives are more prevalent in economically distressed areas. Cross-national analyses of the United Kingdom and Spain reveal similar strategic deployment, but also systematic differences in the motifs emphasized by right-wing and left-wing populists.

Keywords: Populism, text-analysis, network motifs, LLMs, synthetic data augmentation.

*We thank participants at PolMeth, NYU’s Data Science Frontiers: Society and Politics, University of Southern California, Pace Research Institute Oslo, MPSA, ISA, reviewers at RSF grant, University of Southern California, the 10th Monash-Paris-Warwick-Zurich-CEPR Text-As-Data Workshop, and Claudia Salas Gimenez, 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 minority groups 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](#); [Gonzalez-Rostani, 2025a](#)). Understanding how populist rhetoric is constructed and deployed is central to the study of political behavior and democratic accountability.

We argue that the literature implies that populism is best understood as a rhetorical strategy rather than a fixed attribute of politicians, parties, or texts. Populist discourse is composed of modular rhetorical moves that can be activated, combined with other appeals, and withdrawn across issues and contexts. In practice, these theories imply that politicians will alternate, even within a single paragraph, among populist attacks, nationalist claims, policy argument, technocratic appeals, and ordinary campaign language. This strategic view of populism directs attention to when and how populist rhetoric appears, how much of a text it occupies, which narrative structures it invokes, what other rhetorical strategies accompany it, and how these patterns vary over time.

Existing applied and methodological research, however, typically relies on coarse measures that classify entire politicians, parties, or speeches as populist or not, obscuring the internal heterogeneity of political communication. This necessarily masks the expected within and cross-speech variation of references to different groups, actors, targets or grievances across time and space. These limitations are consequential to our understanding of deployment of populism.

We address this gap by introducing Populist-PULSAR, a measurement pipeline that detects

populist rhetoric at the sentence and clause level. The goal is to recover the rhetorical moves that compose populist discourse. This level of granularity is useful because actor-target moral relations often become visible only within sub-sentence units of political text. Grounded in thin-centered theories of populism, Populist-PULSAR identifies who acts, who is 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 annotate core components of populist narratives, and the system uses these annotations to recover narrative structures and expand the labeled sample through synthetic data augmentation. LLMs are optional in our system, and direct human annotation can be substituted for these when necessary or desired.

We focus on four representative populist motifs: victimization (the people harmed by elites), empowerment (the people reclaiming power), resistance (the people showing defiance against corrupt authorities), and elite conspiracy (elites acting in their own interest). These motifs are micro-level building blocks rather than individually sufficient expressions of full populist discourse. Victimization and resistance directly encode conflict between people and elites, whereas empowerment isolates people-centered appeals and elite conspiracy isolates elite-centered appeals that often become fully populist when paired with the missing relational element in nearby rhetoric. To compare these rhetorical structures systematically across texts, we introduce Graph Edit Distance (GED), which measures how closely an extracted sentence-level graph matches each populist motif. This makes it possible to compare statements by their relational structure rather than by shared vocabulary alone. Within these motifs, the system also identifies who is referred as the “elite”, “the people”, and what appeal is articulated in the text.

To demonstrate the analytic payoff of this approach, we apply Populist-PULSAR to a large corpus of U.S. political speech (2008–2025) and analyze all post-primary Trump campaign rallies in 2016. Our application demonstrates that populist rhetoric 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 anti-elite narratives are most common in highly competitive electoral environments. Our fine-grained approach also makes it possible to identify the substantive issues associated with different motifs. For example, elite conspiracy is linked primarily to corruption, whereas empowerment is associated more with social groups, community, and job provision. We also show that the strategic use of populism extends beyond the United States. In the United Kingdom, right-wing populists increasingly emphasize victimization as elections approach, whereas in Spain, radical-left populism also rises near elections but centers more on empowerment. 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.

The paper makes several contributions. Most centrally, it bridges a persistent gap between theories that treat populism as a modular rhetorical strategy and empirical measures that treat it as a fixed attribute of speakers or texts. By recovering the narrative structure of populist discourse at the sentence level, our approach aligns measurement with the thin-centered theoretical tradition and extends prior dictionary-based and scaling methods (e.g., [Rooduijn and Pauwels, 2011](#); [Di Cocco and Monechi, 2022](#); [Dai and Kustov, 2022](#); [Meijers and Zaslove, 2021](#)). This granularity makes it possible to ask not just whether a speech is populist, but which populist narratives it deploys, on which issues, and in which contexts — questions that aggregate measures cannot address. To achieve this at scale, we introduce a hybrid human–LLM annotation pipeline that integrates expert judgment with scalable automation (e.g., [González-Rostani et al., 2025](#); [Le Mens and Gallego, 2025](#)), using LLMs both to expand labeled data and to generate semantically equivalent variants that preserve narrative structure while increasing linguistic coverage. We pair this pipeline with Graph Edit Distance (GED), a transparent structural similarity measure that identifies when differently worded statements instantiate the same underlying populist relation — enabling comparison

across speakers, languages, and time periods without reliance on fixed vocabularies.

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 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 of Thin Centered Populism

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, 2007](#); [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, 2007](#); [Hawkins and Rovira Kaltwasser, 2018](#)). Two implications follow for measurement. First, populism is expressed linguistically and can be measured using political texts.¹

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

¹This does not imply that populism is only expressed linguistically. Images, non-textual symbols, and in-person demonstrations are additional forms of expression. However, each is often or nearly always, accompanied by some textual information. Future work can expand our pipeline to include multi-modal communications.

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. Sociology, political science, and economics all agree that populism is central to understanding a host of social outcomes, yet contradictions persist within and between disciplines. Sociological accounts emphasize social status, resentment, and isolation ([Gidron and Hall, 2017](#); [Hochschild, 2018](#)), while economists stress exposure to globalization, automation, and economic insecurity ([Roddrik, 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, 2025a,b](#)), and work in public administration shows how populist incumbents can erode bureaucratic capacity and institutional constraints (e.g., [Bauer and Becker, 2020](#)).

It is not surprising that these contradictions linger; 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 including which grievances, elites, and people were emphasized. 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, 2011](#); [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, 2022](#); [Ulinskaite 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.² This lack of transparency complicates validation, limits substantive interpretation, and directly contributes to ambiguous evidence related to the drivers of populism.

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.³

Perhaps the most critical limitation of current measurement approaches is that 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 within the text. This limits both theory testing and empirical leverage, making it difficult to study how populist narratives interact with policy content such as na-

²In addition, feature importance on document level classification is still an aggregate, coarse quantity of interest.

³We further discuss this trade-off and document how existing measures differ along these dimensions in [subsection A.1](#) and [Table A.1](#).

tionalism, redistribution, 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., 2024](#); [Gonzalez-Rostani, 2025a](#)), or which aspects of populism motivate voters ([Neuner and Wratil, 2022](#); [Neuner, 2025](#); [Silva et al., 2023](#)), gain limited insight from aggregated scores.⁴

A related concern arises from the multidimensional and non-compensatory nature of populist rhetoric. Scholars argue that core features such as people-centrism, anti-elitism, and anti-pluralism cannot be traded off, so high scores on one dimension cannot compensate for absence on another, making additive measures theoretically problematic ([Wuttke et al., 2020](#)). Populist-PULSAR addresses this concern in two ways. First, it measures structural components separately at the motif level rather than collapsing them into a single score, preserving distinctions that aggregation would erase. Second, by comparing extracted rhetorical graphs to theoretical motifs using GED, it allows researchers to set stricter or looser classification thresholds. At the most conservative threshold, a sentence must exactly match the target structure, yielding a non-compensatory classification rule. Researchers can further require co-occurrence of multiple features within larger textual units, such as speeches or paragraphs, thereby recovering a multi-feature definition while preserving within-text heterogeneity.

A further question concerns the appropriate unit of analysis. Recent work argues that key features of populism may not be observable within a single sentence and may require paragraph-level aggregation for people-centrism and anti-elitism to co-occur (e.g., [Ulinskaitė and Pukelis, 2021](#)). We agree that aggregation matters: our validation operates at the speech level for this reason. Sentence-level detection supports this insight by providing the

⁴Aggregation 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.

granularity needed to observe it empirically. By identifying which sentences contain specific populist components, Populist-PULSAR allows targeted aggregation: researchers can require that anti-elitist and people-centric motifs co-occur within the same speech, paragraph, or text window. This provides a more informative and flexible starting point for researchers than paragraph- or document-level labels, which do not indicate which narrative components are present.

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)⁵ 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 ‘people’ (US) from out-group ‘elites’ (THEM), the actor and target of purported actions, and identify the judgment holder (speaker, 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.”

Clause-level parsing serves as the analytic primitive that can then be aggregated upward to paragraphs, speeches, and speakers. 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

⁵Parsing Unstructured Language into Sentiment-Aspect Representations (PULSAR) was originally designed to extract structured judgments from human rights reports.

links between extracted structures and their textual evidence.⁶

3.1 The Narrative Components of Populism

We define a token-level ontology for annotating populist discourse, with all labels mapped to core components of thin-centered populism in Table 1 (Mudde, 2007). **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). **JUDGMENT-HOLDER** distinguishes speaker from non-speaker sources, critical for separating endorsement from critique. **GROUP** tags link components belonging to the same evaluative statement within multi-structure sentences. These labels are not individually sufficient to define populism; they are component parts that allow us to reconstruct higher-order antagonistic structures transparently.

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 operationalize 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

⁶Our methodology can be used with other documents and party-level systems as well.

with what *moral valence*. The sentence in Figure 1 is an example of our labeling of populist narrative components.

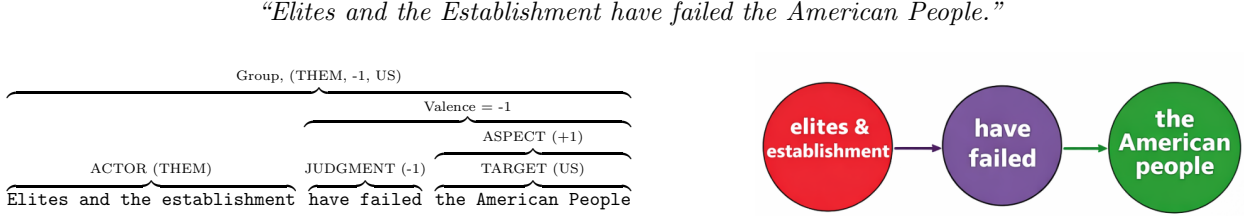


Figure 1: Example of rhetorical structure and graphical representation. See Figure A.3 for more examples.

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

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 judgement-holder identifies whose judgment is being expressed (e.g., the speaker or a reported source). Grouping relations determine which components belong to the same narrative instance, defines edges and the graph G_I . At the same time, we do not assume that every micro-level structure is independently sufficient to constitute full populist discourse. Instead, these structures are the smallest observable rhetorical units from which broader populist passages are assembled

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 rather than a fixed vocabulary. We treat them as representative micro-level rhetorical moves, not as interchangeable dimensions of a single compensatory score. [Table 2](#) summarizes the structure of each motif and [Figure 2](#) provides illustrative examples.

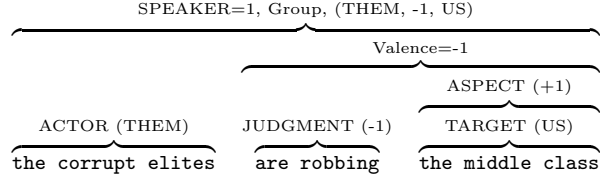
The four motifs map directly onto the canonical dimensions of thin-centered populism in the ideational literature ([Mudde, 2004](#); [Hawkins et al., 2019](#)). Victimization and resistance are the most directly antagonistic because they explicitly encode conflict between the people and the elites. Empowerment and elite conspiracy capture supporting moves: they center the people or the elites respectively, and often become fully populist when paired with antagonistic contrast in nearby rhetoric. The motif inventory is thus not an ad hoc typology but a structural decomposition of the relational properties that thin-centered theory identifies as constitutive of populism. Measuring each motif separately preserves the theoretical distinctions the ideational approach was designed to capture.

Representative Motif	Conceptualization		Operationalization of Graph Motif
	Narrative Components	Narrative Structure	
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)

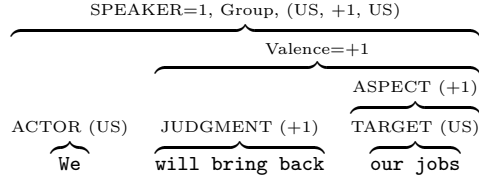
Table 2: Representative populist motifs and their translation from theoretical narrative components to graph-based operationalization in Populist-PULSAR.

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. Populist-PULSAR is therefore not reducible to fixed keywords, speaker labels, or whole-document scores.

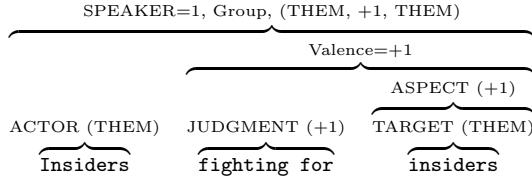
Victimization: “The corrupt elites are robbing the middle class.”



Empowerment: “We will bring back our jobs.”



Elite Conspiracy: “Insiders fighting for insiders.”



Resistance: “We are transferring power from Washington, D.C.”

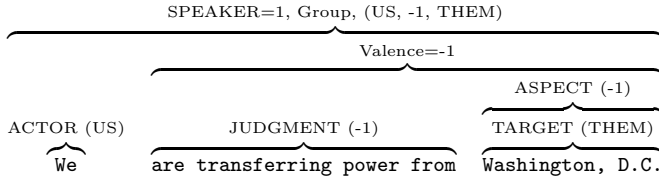


Figure 2: Populist Motifs Illustration through Actor-Judgment-Target decompositions

4 Comparing Extracted Graphs to Populist Motifs

To flexibly assess whether an extracted narrative matches a populist motif, we compute a custom *Graph Edit Distance* (GED) between the sentence-level extracted triplet $\hat{G}_i$ and each motif graph G_j , for all $j \in \{1, 2, 3, 4\}$. GED yields an integer 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, $GED_{ij} : (\hat{G}_i, G_j) \rightarrow d_{ij}$ with $d_{ij} \in \mathbb{Z}_{\geq 0}$. Let $\hat{G}_i = (a, v, b)$ denote an extracted triplet and $G_j = (a^*, v^*, b^*)$ a motif template, where $a, b \in \{\text{US}, \text{THEM}, \text{Residual}\}$ and $v \in \{+, -\}$. We define an additive distance

$$d(\hat{G}_i, G_j) = c_{\text{role}}(a, a^*) + c_{\text{val}}(v, v^*) + c_{\text{role}}(b, b^*),$$

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 $\hat{G}_i$, we compute its distance to every motif graph $G_j \in \mathcal{G}$, where $\mathcal{G}$ is the set of all motif graphs, and take

$$d_{\min}(\hat{G}_i) = \min_{G_j \in \mathcal{G}} d(\hat{G}_i, G_j),$$

using the argmin to identify the nearest motif. A sentence (or triplet) can then be treated as populist when $d_{\min}(\hat{G}_i)$ falls at or below a chosen threshold τ . 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 3](#) represents populist motifs in a

three-dimensional actor–valence–target space induced by the GED penalty structure, shown here for the case $\tau = 1$. Each colored voxel is associated with a motif whose centroid marks the location of an exact match, while the bounding boxes around those centroids are defined by the threshold τ . The two black dots illustrate ambiguity about the actor (on left) and target (on right); each corresponds to an observation with a **Residual** category in that slot. Because $\tau = 1$, both cases still fall within the corresponding motif voxel and would therefore be classified as matches. 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.

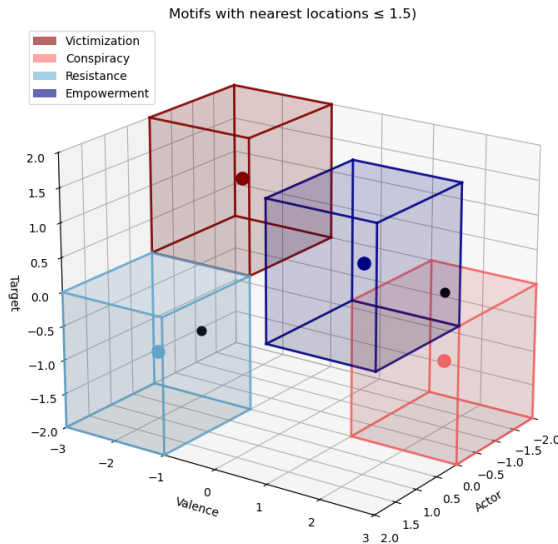


Figure 3: Populist motifs in actor–valence–target space under GED

Notes: The figure illustrates the actor–valence–target space induced by the GED penalty structure for $\tau = 1$. Colored voxels denote motif regions; each centroid is an exact match, and the surrounding box contains sentence graphs with minimum GED at or below τ . The black dots represent cases with a **Residual** actor (left) or **Residual** target (right). Because $\tau = 1$, both are still classified as matches. New sentence graphs are mapped into this space and assigned to the nearest motif by minimizing GED.

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

Sentence Example	Triplet	Empowerments Penalization	GED to Empowerment
Empowerment Reference: “We will bring back our borders”	(US, +, US)	(0,0,0)	0
ACTOR (US) JUDGMENT (+1) TARGET (US) Aspect (+1) We will bring back our borders			
Comparison 1: “We will bring back our jobs”	(US, +, US)	(0,0,0)	0
ACTOR (US) JUDGMENT (+1) TARGET (US) Aspect (+1) We will bring back our jobs			
Comparison 2: “We will bring back our wealth”	(US, +, US)	(0,0,0)	0
ACTOR (US) JUDGMENT (+1) TARGET (US) Aspect (+1) We will bring back our wealth			
Comparison 3: “I know it’s all about making America great again for all Americans”	(Residual, +, US)	(1,0,0)	1
ACTOR (Residual) JUDGMENT (+1) TARGET (US) Aspect (+1) I ’s all about making America great Again			
Comparison 4: “The insiders wrote the rules of the game to keep themselves in power”	(THEM, +, THEM)	(2,0,2)	4
ACTOR (THEM) JUDGMENT (+1) TARGET (THEM) Aspect (+1) The insiders wrote the rules of the game ... themselves ... the money			
Comparison 5: “We have to stop them from cheating in elections.”	(US, -, THEM)	(0,4,2)	6
ACTOR (US) JUDGMENT (-1) TARGET (THEM) ASPECT (+1) We have to stop them from cheating in elections			
Comparison 6: My opponent does not support tax cuts.	(Residual, -, Residual)	(1,4,1)	6
ACTOR (Residual) JUDGMENT (-1) Aspect (+1) My opponent does not support job creation			

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).

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 potential practical bottleneck, as manual annotation is too costly to scale to the modern availability of text 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 4](#) 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). Regardless of the upstream annotator, all substantive results rely on the same final step: extracted sentence structures are mapped into a common triplet representation and matched to theoretical motifs using GED. The resulting dataset retains 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.

The scaling strategy is staged: a small, carefully curated set of human annotations establishes a reference standard (Stage 1), which is then expanded through LLM-assisted labeling and synthetic augmentation (Stages 2 and 3). Researchers could stop there and rely on LLM-based populism annotation, or use these labeled data to fine-tune a classifier for large-scale automated annotation (Stage 4).

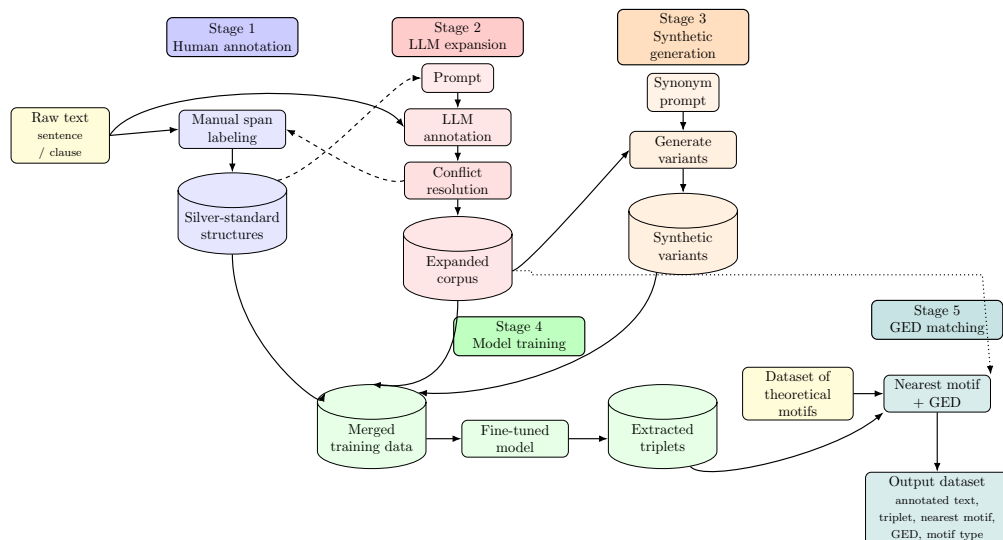


Figure 4: Workflow for scaling annotation and motif classification in Populist-PULSAR.

Note: Human annotation first establishes a silver-standard corpus, which is then expanded through prompt-based LLM annotation and synthetic data generation. Disagreements identified during conflict resolution feed back into human annotation, allowing refinement of span labels and annotation decisions. These data support training of the automated extractor. In the final step, extracted triplets are matched to a dataset of theoretical motifs using graph edit distance (GED). The dotted path indicates that LLM-expanded annotations can also be linked directly to motifs without relying on the fine-tuned model.

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”).⁷ 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.⁸

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).⁹ We implement LLM calls through the OpenAI API (gpt-4o-mini-2024-07-18), setting the 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

⁷This is described as the silver standard because at this point in the pipeline the hand-annotations have not gone through conflict resolution.

⁸Two 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).

⁹Appendix A.5.1 further explains the prompt engineering steps.

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 ($\text{GED} = 0$), and an additional 40.5% exhibit minimal discrepancy ($\text{GED} = 1$), 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.¹⁰ This augmentation increases lexical 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 5 for an example. Because these synthetic sentences inherit the original labels, they enlarge 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

¹⁰For more details, refer to [subsubsection A.5.2](#).

positive training instances to learn the structural patterns that define populist discourse.

Original sentence:

Group (THEM, +1, THEM)
Valence = +1

ACTOR (THEM)
JUDGMENT (+1)
TARGET (THEM)
ASPECT (+1)

Synthetic candidates by role (slot-preserving):

Actor	Judgment	Target	Aspect
the elites	wrote the rules to keep	themselves	in power and in wealth
the establishment	designed the rules to keep	their own position	in control and prosperity
the ruling class	set up the rules to preserve	their status	in authority and riches

Examples of augmented sentences:

1. “The elites wrote the rules of the game to keep themselves in power and in the money.”
2. “The establishment designed the rules to keep their own position in control and prosperity.”
3. “The ruling class set up the rules to preserve their status in authority and riches.”

Figure 5: Prompt-Based Data Augmentation Example.

Note: The original sentence is decomposed into narrative roles. Synthetic variants are generated by replacing each role with semantically compatible alternatives while preserving the underlying narrative structure: Group (THEM, +1, THEM). No separate grammar-correction step is applied.

Stage 4: Model Training for Extended Scalability

The hybrid human-LLM annotation pipeline described above already provides a workable annotation strategy for political speech corpora. Put differently, researchers could stop at the LLM-based stage and use those annotations directly, without the additional fine-tuning step we introduce below. We include that step because it further improves scalability and reduces marginal annotation costs that would otherwise accumulate through API calls. To do so, we fine-tune an open-weight, transformer-based model for automated label extraction on our own compute infrastructure at substantially lower cost. Specifically, we train a sentence-level DistilBERT model 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, and the model minimizes the average loss

across tasks.¹¹

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 0.91 to 0.96 for Actor and Target role classification, and 0.94 to 0.96 for Judgment valence. Aspect valence proves more challenging, with F1 scores between 0.64 and 0.81, 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	0.90	0.92	0.91
	Residual	0.91	0.94	0.92
	Us	0.96	0.91	0.94
Target	Them	0.97	0.88	0.92
	Residual	0.93	0.88	0.90
	Us	0.88	0.95	0.91
Judgment	-1	0.98	0.91	0.94
	+1	0.94	0.99	0.96
Aspect	-1	0.65	0.63	0.64
	+1	0.80	0.82	0.81

Table 4: Sentence-level test-set evaluation metrics.

Note: Notes: Results are from a multi-task transformer fine-tuned for 100 epochs. The model uses attention pooling over token representations, unfreezes the last eight transformer layers, and predicts actor role, target role, judgment valence, and aspect valence with separate task-specific classification heads. The reported loss is the unweighted average of the four cross-entropy losses.

6 Validation

This section evaluates the validity of Populist-PULSAR through four types of evidence. First, we assess face validity by examining whether the system’s sentence-level codings match what close reading would suggest, including variation within a single speech. Second, we evaluate divergent validity by showing that our measure of populism is not reducible to speaker

¹¹Joint training allows shared representations to capture cross-component regularities, such as negative judgments co-occurring with THEM targets. 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 5×10^{-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.

identity: the same individuals produce both populist and non-populist sentences. Third, we assess convergent validity by confirming that speakers and speeches independently identified as more populist by existing measures also receive higher populist sentence counts from our system. Finally, because divergent and convergent validity are subtypes of construct validity, we turn to a broader assessment of construct validity by testing whether the measure recovers theoretically expected empirical patterns in Trump’s 2016 general election rally speeches.

6.1 Face Validity

We assess face validity using a segment from Trump’s 2016 Pensacola rally, previously classified as wholly populist by Dai and Kustov (2022). Our sentence-level analysis shows that 10 of 19 sentences contain explicit populist motifs, which supports an overall populist label while also revealing substantial internal variation. The dominant patterns are victimization narratives, in which elites harm the people, and empowerment narratives, in which the people reclaim control. As shown in Figure 6, Trump shifts from policy discussion into this rhetorical register by assigning blame to “insiders” and other elite actors. Dictionary-based methods perform worse, flagging only the sentence containing the word “corrupt” and missing most of the framing conveyed through structure rather than explicit keywords. This illustrates the main advantage of Populist-PULSAR: it identifies not only whether a passage is populist, but also where those narratives appear and which type is being invoked.

6.2 Divergent Validity

A key validity test is whether sentence-level populist rhetoric varies independently of speaker identity. Figure 7 illustrates this divergent validity with a two-by-two comparison: the rows separate populist from non-populist sentences, while the columns separate non-populist speakers (Obama, Bush) from populist speakers (Sanders, Trump). Dotted lines connect sentences with identical narrative triplet structures across columns. Non-populist speakers

Dai and Kustov (2022)

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.

Populist-PULSAR annotations

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 6: Side-by-side comparison of prior scholar annotations and Populist-PULSAR motif-specific annotations for a segment from Donald Trump’s 2016 Pensacola rally.

Note: Left panel uses yellow highlighting to identify populism coded by previous work. Right panel uses motif-specific highlights: pink = Elite Conspiracy Narrative; orange = Victimization Narrative; green = Empowerment Narrative; blue = Resistance Narrative.

occasionally deploy classic populist motifs. Obama uses a victimization frame and Bush an empowerment frame that structurally match Trump and Sanders sentences. Conversely, populist speakers frequently produce non-populist rhetoric. This confirms that Populist-PULSAR responds to textual content rather than speaker identity or party affiliation.

A further illustration comes from the Bush passage in Figure 8. Although Dai and Kustov (2022) do not classify that paragraph as populist, Populist-PULSAR identifies several sentence-level populist motifs within it, including empowerment, victimization, and resistance frames. This contrast shows that our method can recover embedded populist rhetoric even in passages that prior work codes as non-populist at a more aggregated level. In that sense, sentence-level analysis not only separates rhetoric from speaker identity, but also reveals within-passage variation that paragraph-level labels can miss.

6.3 Convergent Validity

Next, we assess how Populist-PULSAR’s scores compare with a standard dictionary-based measure of populism applied to Trump sentences (Rooduijn and Pauwels, 2011), and with

Non-Populist Populist	Non-Populist Speakers (Obama, Bush)		Populist Speakers (Sanders, Trump)	
	[THEM, -, US] “... they have fought us every step of the way.” — Obama		[THEM, -, US] “The corrupt elites are robbing the middle class.” — Trump	
	[US, +, US] “This generation was given the best education...” — Bush		[US, +, US] “We are going to create an economy that works for all...” — Sanders	
	[Res., +, US] “Our union can be perfected.” — Obama		[Res., +, US] “I will declare a national emergency...” — Trump	
	[Res., +, Res.] “A decent education is the gateway...” — Bush		[Res., +, Res.] “The scientific community has spoken...” — Sanders	

Figure 7: Divergent validity across speaker types.

Dai and Kustov (2022)

Thank you, Dow Williams. And thank you, Gainesville. What a fantastic rally. I am grateful to you, and it makes me think we are going to win this election. Let me just say how pleased I am to be accompanied on this trip by two of America's greats: Governor Campbell, the governor of South Carolina, and Senator Strom Thurmond, the great United States senator. And you know, everywhere I go I see signs that say, "Clean House!" Let me say let's also clean Senate and send Paul Coverdell to the United States Senate. May I thank Dow Williams, who's our master of ceremonies, and everybody responsible for this fantastic rally. And let me say this: One left me in Norcross. I expect I will see her in the White House Thursday night, but I am very proud of our First Lady, Barbara Bush, who sends you her very best. It is great to be in the land of the free and the home of the Braves. You know—you got it. This is what we're going to do to Governor Clinton, Clinton-Gore. You know, baseball's exactly like politics. So you forget about all these crazy polls. Don't let these newscasters tell you what's happening or how to vote. On Election day we're going to show America it ain't over 'til Cabrera swings, and that's exactly the way it's going to be. Anybody see that debate last night? Well, it seems to me we had a chance to lay out the difference before the American people. I was talking about a difference in experience, a difference in philosophy, and, yes, a difference in character. And I think all three of those things matter. The scariest moment of that debate was when Governor Clinton said he wanted to do for the United States what he's done to Arkansas. That scares me. We cannot have that. Arkansas is the fiftieth out of fifty States in environmental initiatives; forty-ninth in high school diplomas; forty-fifth in well-being of children; and incomes and jobs and wages lag the entire nation. We do not need that for the greatest country on the face of the earth. You know, I agreed with one thing that Ross Perot said. He said, the grocery store is no preparation for Wal-Mart, and I think he's right. This is the big leagues. You can't have it both ways on every issue. You've got to say what you think, admit your mistakes, and lead and not waffle. I am that kind of a leader, and the other man is the waffler. I had a chance to point out last night—he didn't like it a bit—that Governor Clinton has already said he wants to raise your taxes by \$150 billion. He wants to spend—I'm sorry. He wants to spend \$220 billion. He says he'll take it all from the rich. All you cab drivers, all you nurses, all you boys'-club workers, all you watch out, because he's coming right after your wallet, coming right after it. We are not going to let him tax the American people any more.

Populist-PULSAR annotations

Thank you, Dow Williams. And thank you, Gainesville. What a fantastic rally. I am grateful to you, and it makes me think we are going to win this election. Let me just say how pleased I am to be accompanied on this trip by two of America's greats: Governor Campbell, the governor of South Carolina, and Senator Strom Thurmond, the great United States senator. And you know, everywhere I go I see signs that say, "Clean House!" Let me say let's also clean Senate and send Paul Coverdell to the United States Senate. May I thank Dow Williams, who's our master of ceremonies, and everybody responsible for this fantastic rally. And let me say this: One left me in Norcross. I expect I will see her in the White House Thursday night, but I am very proud of our First Lady, Barbara Bush, who sends you her very best. It is great to be in the land of the free and the home of the Braves. You know—you got it. This is what we're going to do to Governor Clinton, Clinton-Gore. You know, baseball's exactly like politics. So you forget about all these crazy polls. Don't let these newscasters tell you what's happening or how to vote. On Election day we're going to show America it ain't over 'til Cabrera swings, and that's exactly the way it's going to be. Anybody see that debate last night? Well, it seems to me we had a chance to lay out the difference before the American people. I was talking about a difference in experience, a difference in philosophy, and, yes, a difference in character. And I think all three of those things matter. The scariest moment of that debate was when Governor Clinton said he wanted to do for the United States what he's done to Arkansas. That scares me. We cannot have that. Arkansas is the fiftieth out of fifty States in environmental initiatives; forty-ninth in high school diplomas; forty-fifth in well-being of children; and incomes and jobs and wages lag the entire nation. We do not need that for the greatest country on the face of the earth. You know, I agreed with one thing that Ross Perot said. He said, the grocery store is no preparation for Wal-Mart, and I think he's right. This is the big leagues. You can't have it both ways on every issue. You've got to say what you think, admit your mistakes, and lead and not waffle. I am that kind of a leader, and the other man is the waffler. I had a chance to point out last night—he didn't like it a bit—that Governor Clinton has already said he wants to raise your taxes by \$150 billion. He wants to spend—I'm sorry. He wants to spend \$220 billion. He says he'll take it all from the rich. All you cab drivers, all you nurses, all you boys'-club workers, all you watch out, because he's coming right after your wallet, coming right after it. We are not going to let him tax the American people any more.

Figure 8: Side-by-side comparison of prior scholar annotations and Populist-PULSAR motif-specific annotations for a segment from George Bush's 1992 rally.

Note: Left panel uses yellow highlighting to identify populism coded by previous work. Right panel uses motif-specific highlights: pink = Elite Conspiracy Narrative; orange = Victimization Narrative; green = Empowerment Narrative; blue = Resistance Narrative. Note that the sentence "He says he will take it from the rich" demonstrates the resistance narrative and a motif advanced by a judgment holder other than the speaker.

annotated chunks from Dai and Kustov (2022). We find a significant positive correlation: as Figure 9 shows, speeches containing more populist keywords also tend to contain more sentences classified as populist motifs. The correlation is far from perfect, which is consistent with the different logic of these approaches. For example, dictionary methods only

flag populism when specific words appear, whereas Populist-PULSAR can identify populist narratives even when those exact keywords are absent.

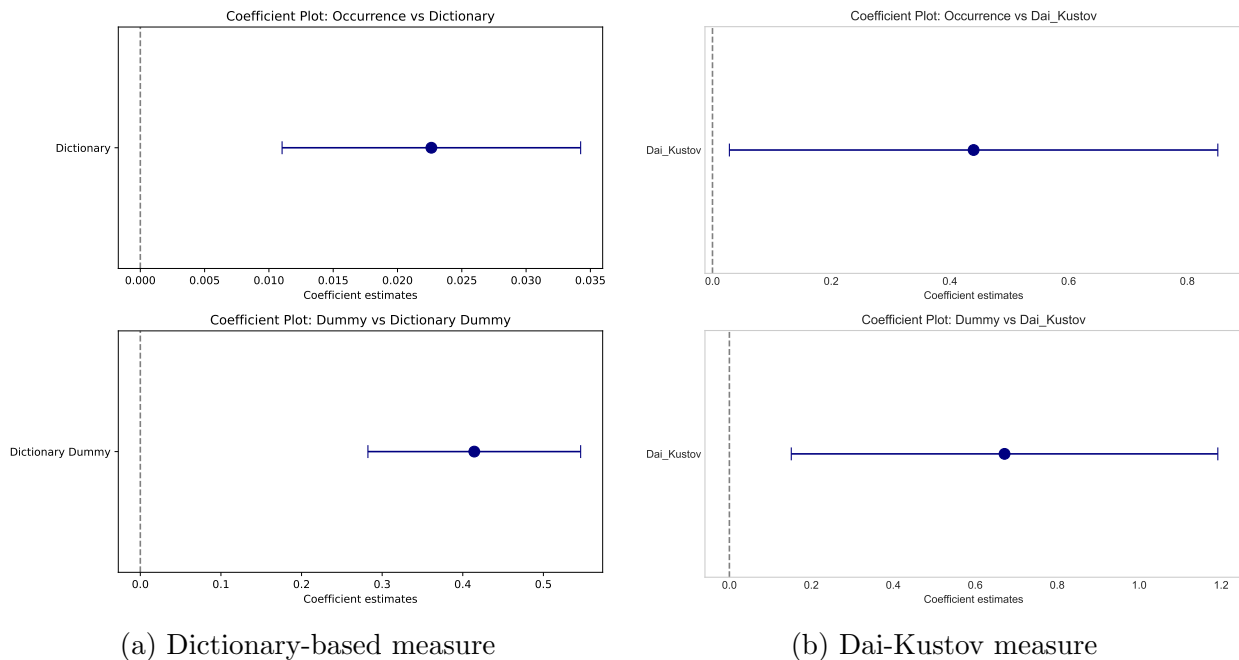


Figure 9: Populist Motifs and Alternative Measures of Populist Rhetoric

Notes: Each panel reports coefficient estimates from models where the dependent variable is the number of populist motifs extracted from each text unit. The left panel uses the dictionary-based measure of populist rhetoric applied to Trump 2016 (N=5331 sentences); the right panel uses the Dai-Kustov measure applied to their replication data (N=587 chunks/paragraphs of text). Horizontal lines indicate confidence intervals. Table A.3 reports the full regression results for the dictionary-based measure, and Table A.4 reports the corresponding results for the Dai-Kustov measure.

6.4 General Construct Validity

In addition to comparisons with existing measures, we evaluate whether our approach recovers theoretically predicted patterns linking thin-centered populism to political competition in a case widely understood as a paradigmatic example of populist rhetoric and studied extensively by previous scholars: Donald Trump’s 2016 general-election campaign. Using rally speeches with complete covariate data, we replicate and extend the dictionary-based analysis of [Gonzalez-Rostani \(2025a\)](#) to examine when, where, and which motifs are deployed. In this application, the annotation pipeline is fully automated using LLMs, and motif matches are aggregated to the speech level using exact matching ($GED = 0$). We test two empirical expectations: first, that populist rhetoric is targeted geographically in competitive swing-state

contexts; and second, that it is more prevalent in regions with a greater share of workers exposed to economic distress.

Our evidence confirms our expectations, and suggests that Trump used more populist language in competitive states: rallies in swing states showed significantly higher populist motif counts than rallies in safe states (Figure 10a). This is a substantial effect and suggests that our measure is not merely capturing the speaker’s general style, since the same candidate shifted toward a more populist tone in competitive 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, 2025a), we find that rallies in more exposed counties contained higher levels of populist rhetoric (Figure 10b). This pattern is consistent with theories linking economic insecurity to populist appeal (e.g., Gidron and Hall, 2017; Inglehart and Norris, 2016; Gonzalez-Rostani, 2025a; Gennaro et al., 2024).

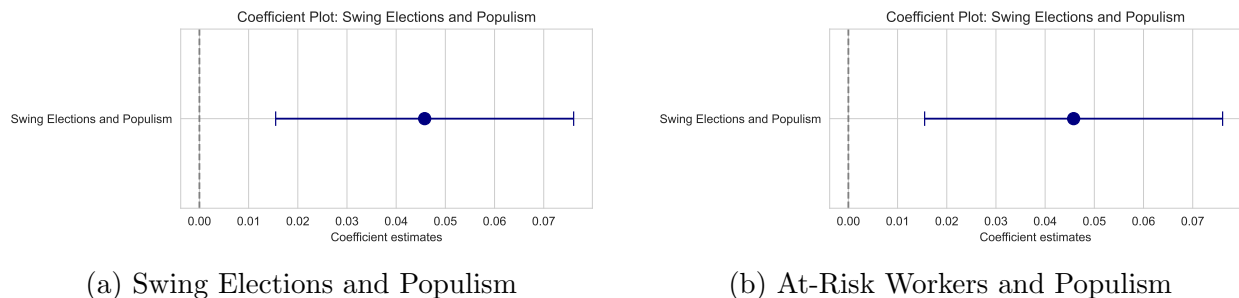


Figure 10: Geography of Populism Deployment

Notes: The dependent variable is a dummy equal to one when a text contains a populist motif. Panel (a) shows the estimated coefficient for *Swing Elections*, defined as a dummy equal to one when the 2012 presidential election in a state was decided by less than 10 percentage points. Panel (b) reports the coefficient for exposure to at-risk workers. The full regression results corresponding to these coefficient plots are reported in Table A.5 in the Appendix. The unit of analysis is speeches $N = 70$.¹²

The preceding analyses validate that Populist-PULSAR recovers expected aggregate patterns. But the method’s central payoff is what it reveals beneath those aggregates and cannot be uncovered at scale through other approaches. Because Populist-PULSAR distinguishes among specific motifs, it shows that the same overall increase in populist rhetoric in two different contexts is produced by fundamentally different narrative strategies.

In swing states, the increase is driven primarily by elite-conspiracy, victimization, and resistance motifs; narratives that cast elites as the central source of threat (Figure 11a). In economically vulnerable regions, by contrast, the increase is driven by empowerment motifs, which focus less on blaming elites and more on affirming “the people” through pro-worker claims (Figure 11b). An aggregate populism score would assign similar levels to both contexts; only motif-level disaggregation reveals that they reflect rhetorically distinct campaign strategies, with direct implications for how researchers interpret the conditions under which populist appeals succeed.

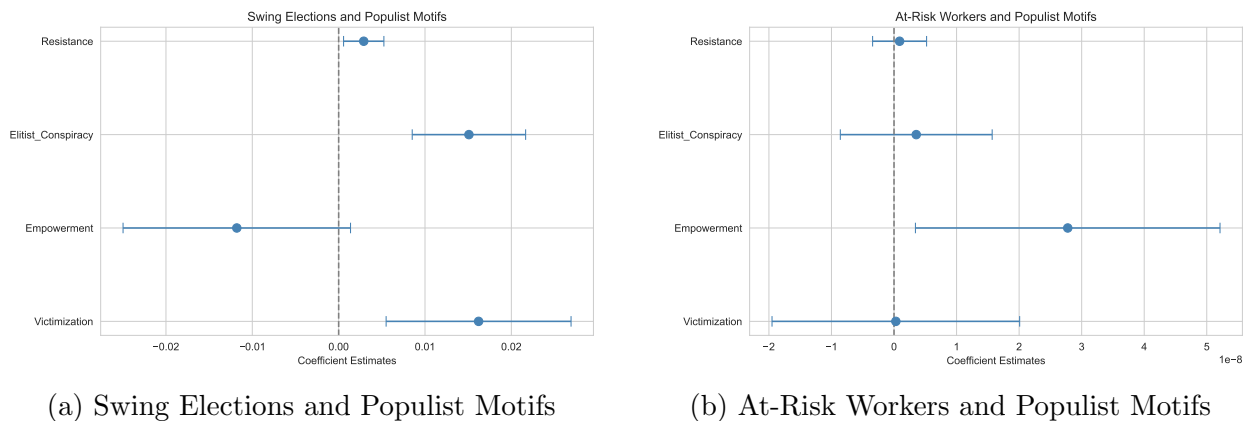


Figure 11: Geography of Type of Populism Deployment

Notes: The dependent variable is a dummy equal to one when a text contains a populist motif. Panel (a) reports the estimated relationship between populist motifs and *swing elections*, defined as a dummy equal to one when the 2012 presidential election in a state was decided by less than 10 percentage points. Panel (b) reports the relationship between exposure to at-risk workers and populist motifs. The full regression results corresponding to these coefficients are reported in Table A.6 in the Appendix. Unit of analysis is speech, $N = 70$.

These motifs are not only statistically separable but substantively interpretable, which is a direct consequence of the method’s inspectability. A fighting-words analysis (Monroe et al., 2017) shows that each motif corresponds to a distinct lexical field (Figure 12). Empowerment is characterized by inclusive, future-oriented language (*American, our, going*), whereas elite-conspiracy rhetoric is anchored in named antagonists and institutional power (*Hillary, Clinton, powerful*). Resistance is marked by conflictual economic language (*them, executives, corporations*), while victimization combines named opponents with threat-oriented terms (*they, she, us*). These differences indicate that the motifs capture distinct ways of structuring political conflict: empowerment constructs a collective political subject, resis-

tance emphasizes contestation, elite conspiracy locates threat in captured institutions, and victimization frames politics through injury imposed by identifiable opponents.

The more structured breakdown in [Figure 13](#), which classifies actors, targets, judgments, and aspects within each motif, reinforces this conclusion.¹³ Empowerment disproportionately centers on Trump, his campaign, citizens, workers, and voters, paired with judgments of creation/change and topics such as economy/jobs and social groups/community. Elite conspiracy clusters around Clinton/Democrats, corruption/elites, and government administration. Victimization casts Trump campaign allies and citizens as targets rather than actors, tied to political opponents and social groups/community. Resistance occupies yet another position, emphasizing conflict/restriction and communication/knowledge. Read together, these figures show that the motifs differ not only in frequency but in the actors, issues, and vocabularies through which they communicate political claims.

Finally, to show that the strategic use of populism extends beyond the U.S. case, we present appendix evidence from party YouTube videos in the United Kingdom (2014–2024) and Spain (2014–2023). Outsider parties in both countries intensify populist rhetoric as elections approach, but the motifs they emphasize vary systematically by ideological orientation. In Spain, outsider left parties increase empowerment motifs centered on pro-people and pro-worker narratives.¹⁴ In the United Kingdom, outsider right parties increase victimization and elite-conspiracy motifs that frame politics as conflict with powerful elites.¹⁵ This left-right asymmetry mirrors the U.S. pattern: grievance-oriented motifs are more common where electoral competition is the main driver, whereas affirmative, people-centered motifs are more common where economic vulnerability is more salient. The consistency across the three countries suggests that motif-level variation is not simply an artifact of Trump’s idiosyncratic

¹³We use LLM for annotation into categories; for further details refer to [subsection A.10](#).

¹⁴The cross-national analyses use official party YouTube videos, which include a broader mix of content, such as policy announcements, candidate introductions, and issue advocacy, than the rally speeches used in the U.S. analysis. This makes the YouTube corpus a harder test for detecting strategic motif deployment, since populist rhetoric is diluted by non-campaign material. See [Figure A.8](#).

¹⁵See [Figure A.8](#).

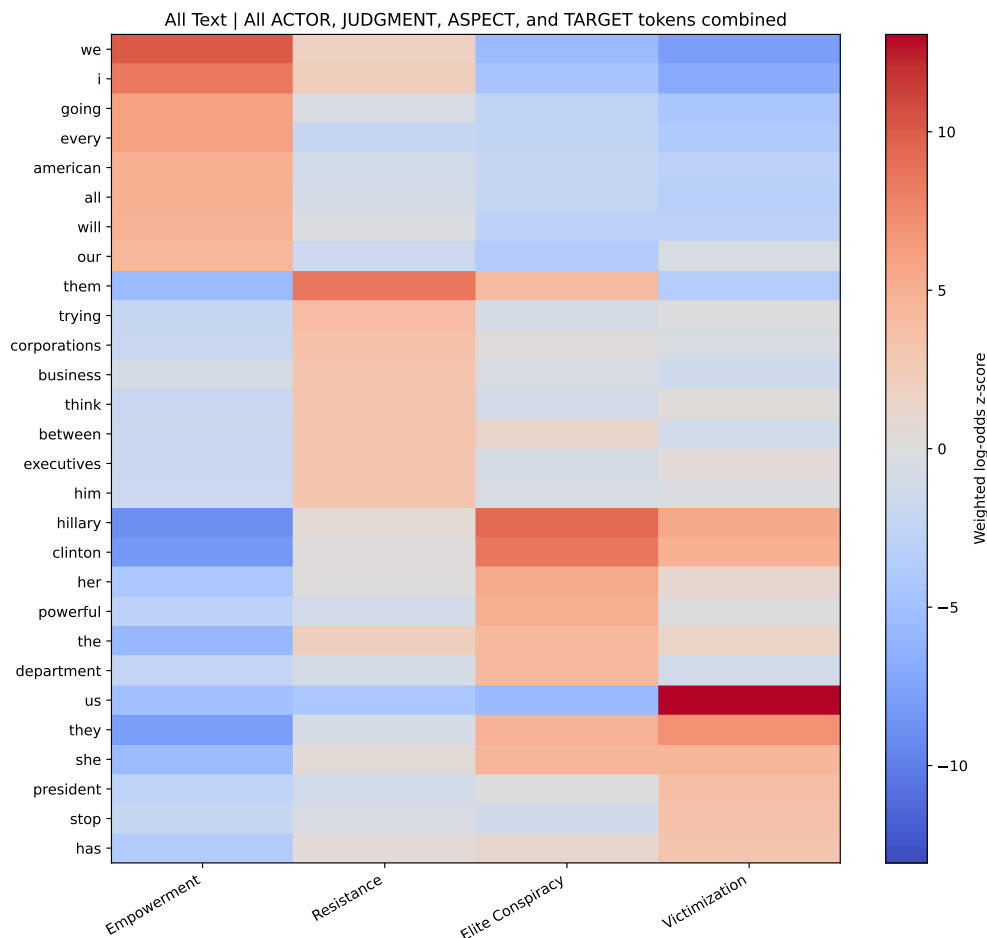


Figure 12: Fighting words across all text for extracted populist motifs

Notes: The figure shows the words most strongly associated with each extracted populist motif using the full text and combining all ACTOR, JUDGMENT, ASPECT, and TARGET tokens. Cells report weighted log-odds z-scores from the “fighting words.” Positive values indicate that a word appears disproportionately more often in texts associated with that motif relative to the rest of the corpus, while negative values indicate the opposite. Colors reflect the magnitude of the association (red = more characteristic of the motif; blue = less characteristic). Appendix A.9 provides details on the calculation.

style, but a broader feature of how populist rhetoric is strategically assembled.

Overall, Populist-PULSAR performs well across multiple validation exercises. Face validity confirms that sentence-level codings track close readings of speech text. Divergent validity shows that the measure is not reducible to speaker identity. Convergent validity demonstrates alignment with existing approaches. General construct validity shows that the measure captures the strategic use of populism. Populist-PULSAR’s distinctive contribution also lies in revealing what underlies aggregate patterns by identifying the actors, targets, and aspects that structure rhetoric across contexts. Populist rhetoric is not a single phenomenon

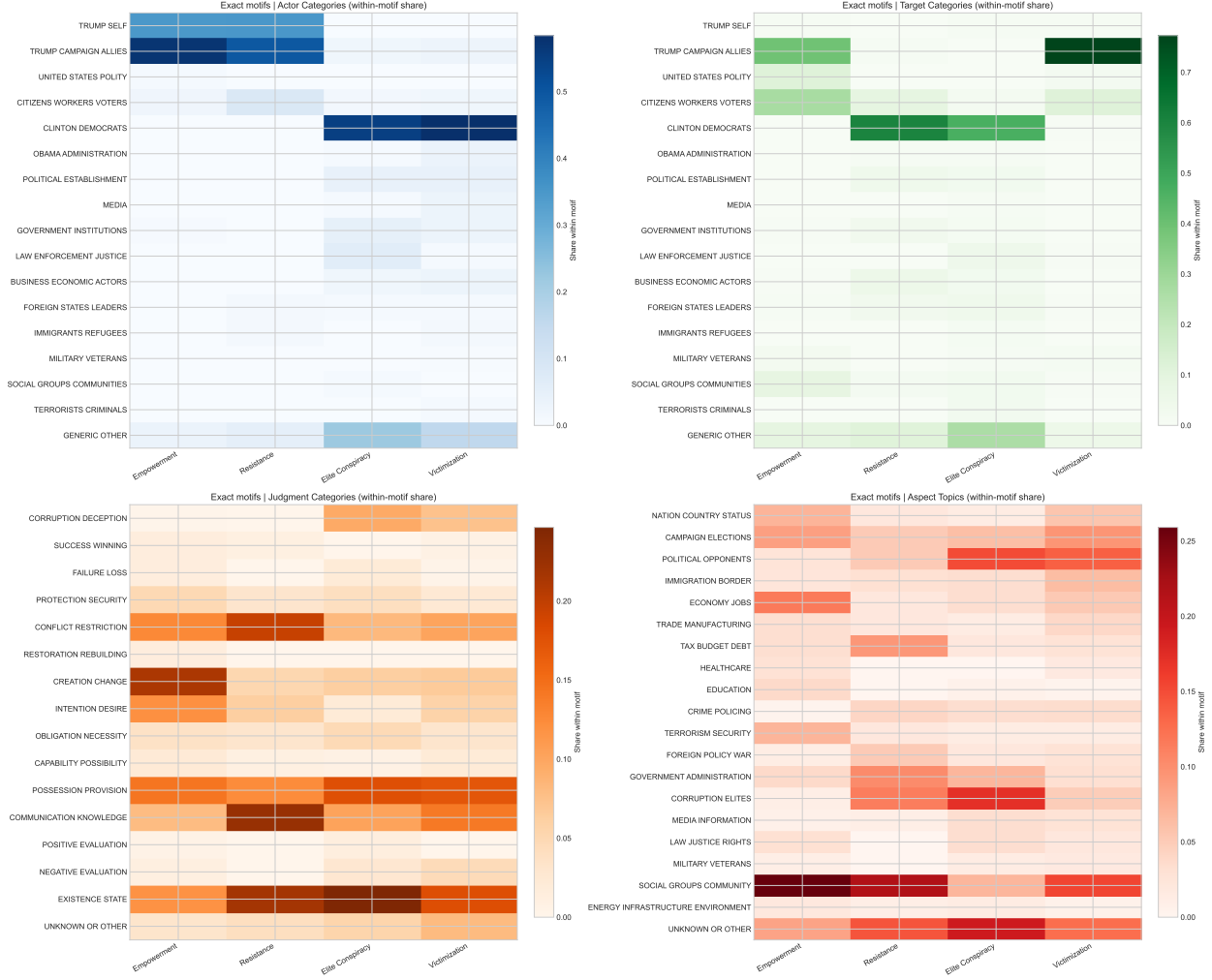


Figure 13: Within-motif shares across actor, target, judgment, and issue categories
Notes: This figure reports heatmaps for the motifs Empowerment, Resistance, Elite Conspiracy, and Victimization. Each cell shows the share of observations within a motif assigned to a given classification. Darker cells indicate higher within-motif shares. Categories were created through LLM-based annotation of the content of the texts and then grouped into substantively interpretable classes; full details are in Appendix A.10. For the corresponding bar-chart version including the overall populism, see Figure A.9.

deployed in varying amounts, but a repertoire of narratives selected to match electoral and economic contexts.

7 Discussion

This paper introduced Populist-PULSAR, a motif-based approach to measuring populist rhetoric at the sentence and sub-sentence level. By representing populist discourse as structured narrative motifs rather than document- or speaker-level labels, the approach aligns

measurement with thin-centered theories of populism and directly addresses the aggregation problem in existing work. Our pipeline is flexible enough to be used with or without LLMs, either closed or open-weight versions, as well as with combinations — as we have done here.

The central empirical finding is that aggregate populism scores conceal strategically distinct rhetorical choices. In Trump’s 2016 campaign, the same overall increase in populist rhetoric was produced by different narrative strategies depending on context: elite-conspiracy and victimization motifs in competitive states, empowerment motifs in economically distressed regions. Cross-national evidence from Spain and the United Kingdom suggests this is not idiosyncratic to the US but reflects a more general pattern in which the flexibility of populist rhetoric operates at the level of motif type. This has consequences for how researchers interpret populist campaign strategy. Theories that treat populism as a single dimension of political communication — more or less of it — miss the compositional choices that motif-level measurement makes visible. Understanding why politicians reach for grievance narratives in one context and affirmative, people-centered narratives in another requires the kind of disaggregation this approach provides.

These results also open several directions for future research. A first priority is to connect rhetorical structure to political mobilization. Combining motif-based text analysis with audio or video data from campaign events would allow researchers to assess which narrative forms elicit stronger audience responses and whether frequently deployed motifs are also the most effective. Experimental designs could extend this agenda by testing whether exposure to different motifs affects outcomes such as willingness to vote, political participation, or other forms of engagement. More broadly, linking motif-level measurement to models of persuasion could help identify whether different narrative structures activate distinct cognitive or affective responses across audiences, shedding light on heterogeneous effects.

On the methodological side, the underlying pipeline is 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 moves 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.

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 (2022, January). When Do Politicians Use Populist Rhetoric? Populism as a Campaign Gamble. *Political Communication*, 1–22.
- 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 (2024). Intertemporal evidence on the strategy of populism in the united states. *Electoral Studies* 92, 102853.
- 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. (2025a). Elections, Right-wing Populism, and Political-Economic Polarization: The Role of Institutions and Political Outsiders. *The Journal of Politics*.
- Gonzalez-Rostani, V. G. (2025b). From Automation to “Make It Great Again”: The Cultural Path to Right-Wing Populism.

- 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>.
- 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. *European Journal of Political Research* 46(3), 319–345.
- 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. Wiegrefe, U. Alon, N. Dziri, S. Prabhumoye, Y. Yang, S. Gupta, B. P. Majumder, K. Hermann, S. Welleck, A. Yazdan-

- bakhsh, 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.
- Monroe, B. L., M. P. Colaresi, and K. M. Quinn (2017, January). Fightin’ Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. *Political Analysis* 16(4), 372–403.
- Mudde, C. (2004). The Populist Zeitgeist. *Government and Opposition* 39(4), 541–563.
- Mudde, C. (2007). *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.
- Neuner, F. G. (2025, September). Can (Thin) Populism be Manipulated without Manipulating Host Ideology? Evidence from a Conjoint Validation Approach. *Journal of Experimental Political Science*, 1–15.
- Neuner, F. G. and C. Wratil (2022, June). The Populist Marketplace: Unpacking the Role of “Thin” and “Thick” Ideology. *Political Behavior* 44(2), 551–574.
- 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., 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 (2011). 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. Goncarenko, 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].
- Silva, B. C., F. G. Neuner, and C. Wratil (2023, November). Populism and Candidate Support in the US: The Effects of “Thin” and “Host” Ideology. *Journal of Experimental Political Science* 10(3), 438–447.
- Ulinskaitė, J. and L. Pukelis (2021, June). Identifying Populist Paragraphs in Text: A machine-learning approach. arXiv:2106.03161 [cs].
- Wuttke, A., C. Schimpf, and H. Schoen (2020). When the whole is greater than the sum of its parts: On the conceptualization and measurement of populist attitudes and other multidimensional constructs. *American Political Science Review* 114(2), 356–374.
- 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	1
A.1.1	Coarse Granularity	1
A.1.2	Limited Nuance	1
A.1.3	Scalability and Cross-Context Portability	1
A.2	Examples of the Limitations of Aggregate Party-Level Populism Labels in Prior Work	1
A.3	Benefits of our Approach	3
A.4	Designing a Finer-grained, Interpretable Populist Narrative Measurement System	4
A.4.1	Sampling	4
A.4.2	Handling Complex Cases	4
A.5	Annotation Process	6
A.5.1	Prompt Engineering for Populist Motif Detection	6
A.5.2	Synthetic Data Generation	10
A.5.3	Matching Extracted Narrative Structures to Motifs Using a Graph Edit Distance	10
A.6	Fine-tuning Details	11
A.6.1	Sentence-Level Multi-Task Classification	11
A.7	Examples of Our Approach	12
A.8	Tables and Figures	14
A.9	Identifying Distinctive Words Across Motifs	16
A.10	LLM-Based Topic Classification	17

A.1 Problems with Past Measurement

2 introduces the measurement challenges motivating Populist-PULSAR. Here we organize these challenges into three categories: granularity, nuance, and scalability. We then link each to specific limitations of existing approaches. Table A.1 summarizes how prominent methods fare on each dimension.

A variety of strategies have been developed to classify politicians or political parties as populist, including expert surveys (Polk et al., 2017; Meijers and Zaslove, 2021; Norris, 2020), manual coding of speeches and manifestos (Hawkins et al., 2019; Jenne et al., 2021; Casiraghi, 2021), dictionary-based methods (JAGERS and WALGRAVE, 2007; Pauwels, 2011; Rooduijn and Pauwels, 2011), and supervised text classification (Di Cocco and Monechi, 2022; Dai and Kustov, 2022; Ulinskaitė and Pukelis, 2021). While these approaches have advanced the study of populist discourse, they share three recurring limitations.

A.1.1 Coarse Granularity

Most existing systems assign labels at the level of parties, speakers, or multi-paragraph text segments¹⁶. This level of resolution is adequate for "thick" definitions that equate populism with specific policy positions, but it cannot accommodate the within-text variation implied by thin-centered theory. A single speech may alternate between populist and non-populist framing depending on the topic; a single party may include members who differ sharply in their use of populist rhetoric. Beyond this variation, existing approaches generally cannot distinguish the stance a speaker takes toward populist narratives. As populism has become a salient political theme, speakers increasingly reference populist frames without endorsing them — critiquing an opponent’s people-versus-elite rhetoric, for instance, while deploying their own.¹⁷

A.1.2 Limited Nuance

"Thick" or subject-matter-based measures also obscure the interplay between populism and adjacent discourses such as nationalism, anti-globalism, or techno-skepticism. This makes it difficult to distinguish left-wing from right-wing variants of populism, or to trace how populist rhetoric intersects with specific policy domains. The growing association between populist discourse and anti-science sentiment following the COVID-19 pandemic (Mede and Schäfer, 2020) illustrates the analytical cost: a framework that cannot separate populist motifs from the issue content they attach to will conflate these distinct dimensions.

A.1.3 Scalability and Cross-Context Portability

Manual annotation produces theoretically precise labels but is costly to scale, particularly across languages and political systems. Dictionary-based methods offer efficiency but are brittle across contexts. Dictionaries developed for European populism (Pauwels, 2011; Rooduijn and Pauwels, 2011; Rooduijn et al., 2014; Gennaro et al., 2024) require substantial adaptation for new settings (Hawkins et al., 2019; Elçi, 2019; Gründl, 2020; Nishikawa, 2021; Gennaro et al., 2024), and the meaning of key terms can shift across contexts in ways that undermine validity (Schaffer, 2000).

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, 2022)—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

¹⁶See A.2 for examples of how aggregate party-level labels can misclassify political speech

¹⁷Some utterances contain more complex structures, such as a speaker disagreeing with another actor’s positive evaluation of a populist claim.

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, 2011); (Gennaro et al., 2024)	✗	✓	✓
NLP — supervised	(Di Cocco and Monechi, 2022); (Ulinskaitė and Pukelis, 2021); (Dai and Kustov, 2022)	✗	?	✓

Notes: The table summarizes common approaches used to measure populism in political texts. The columns indicate whether each approach allows fine-grained measurement of populist narratives, produces interpretable classifications, and can scale to large text corpora.

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 classifier trained following Di Cocco and Monechi (2022), 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, such as “we,” “freedom,” “our country”, can appear in speech from any party. Figure A.1 shows that the top features predictive of the Republican class in the U.S. model include terms like “congress,” “republican,” and “government,” which are not inherently populist despite the party’s populist coding. This suggests the classifier is learning partisan identity rather than populist narrative structure.

These problems are further illustrated through feature importance scores for the United States and Italy ML models (Figure A.2). 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.

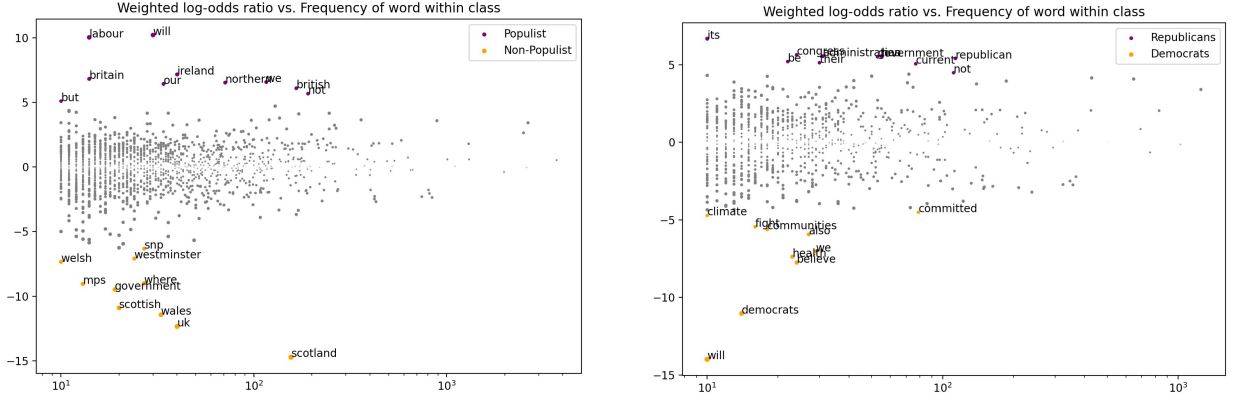


Figure A.1: Weighted log-odds ratio of words used in manifestos: UK 2017 (Populist vs. Non-Populist parties) and U.S. 2016 (Republicans vs. Democrats).

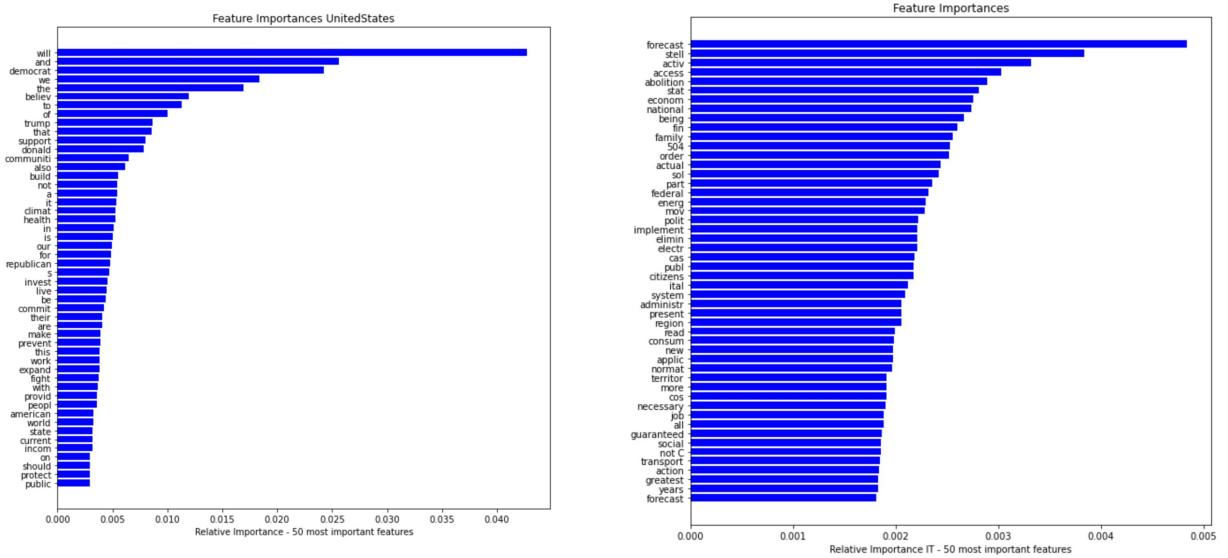


Figure A.2: Top classifier features from U.S. party texts and Italian political speech.

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

The introduction describes three advantages of Populist-PULSAR over existing methods: sentence-level granularity, structural representation via narrative motifs, and cross-context portability. Here we briefly illustrate the second of these with an example that highlights a challenge existing tools cannot address.

As politicians increasingly engage in meta-commentary on populism, measurement systems must distinguish between sincere deployment of populist rhetoric and other uses of populist language, such as quotation, critique, or satire. Consider a statement by Senator Jon Ossoff on March 23, 2025: "they are literally

the elites they pretend to hate.” This sentence simultaneously references and critiques populist framing: the speaker attributes an elite-conspiracy narrative to ”them” while rejecting its sincerity. A keyword-based system would likely flag ”elites” and ”hate” as populist markers. Populist-PULSAR, by contrast, can recover the relational structure: the judgment holder is not the speaker, and the evaluative direction is reversed. Without encoding who judges whom and from what perspective, such statements are liable to be misclassified.

The portability advantage is demonstrated empirically in the cross-national analyses in section 6 and Figure A.8, where the same motif framework is applied to Spanish and British political communication without modification to the underlying annotation scheme.

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.

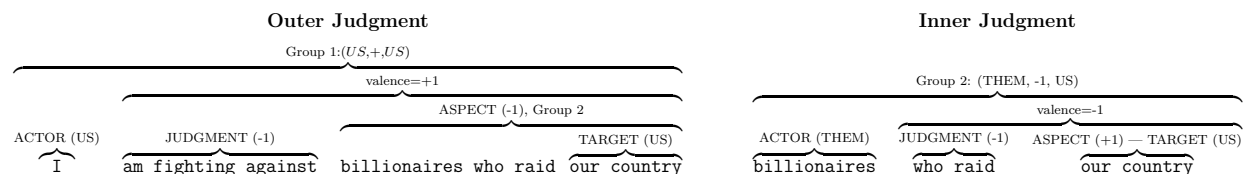
A.4.1 Sampling

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, which we illustrate through our brief cross-national analysis of Spain and the UK.

A.4.2 Handling Complex Cases

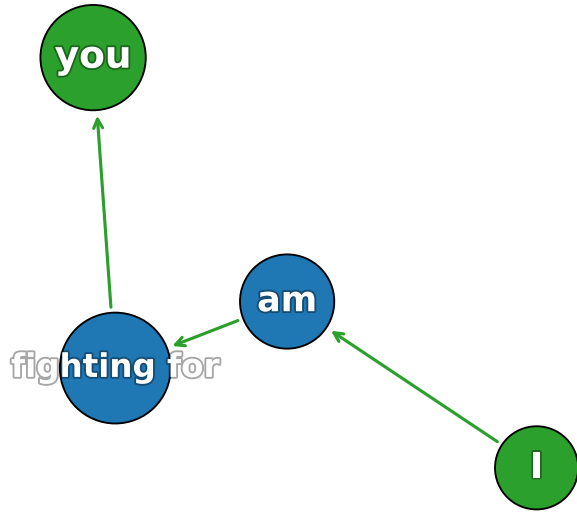
Nested Aspects: Our system also captures nested narratives within judgment-on-aspects, which are distinct from second-order judgment holders. Nested judgments-on-aspects within aspects allow us to identify when a speaker evaluates not only an action, but also the motivations or effects behind that action.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

“I am fighting against billionaires who raid our country.”

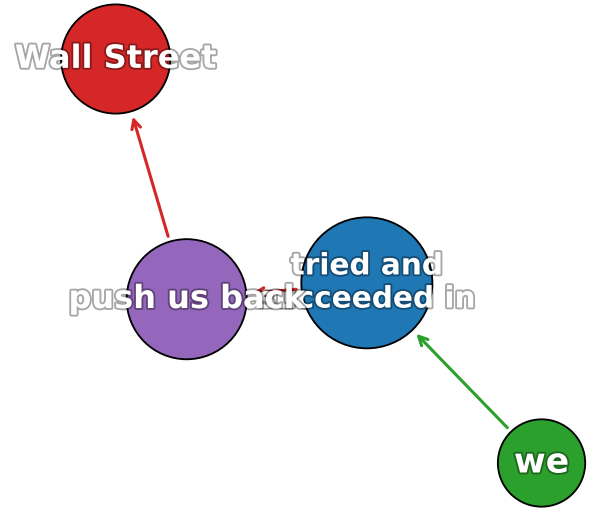


In this example, the outer layer evaluates the speaker’s opposition to harmful actors, while the inner layer attributes harm to those actors. The nesting allows the system to represent how the speaker (*US*) positions themselves against *THEM* (the billionaires) by condemning their actions toward a valued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

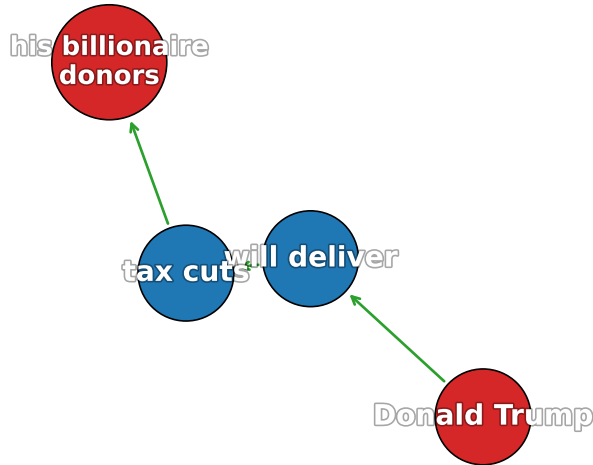
Non-Speaker Judgment-holders: Our approach effectively captures judgments attributed to individuals other than the speaker, a critical feature for identifying indirect populist assertions or rebuttals. Political discourse frequently involves rhetorical structures such as “my opponent says... but I say...,” where speakers reference or critique judgments voiced by others. Ignoring this distinction could misattribute



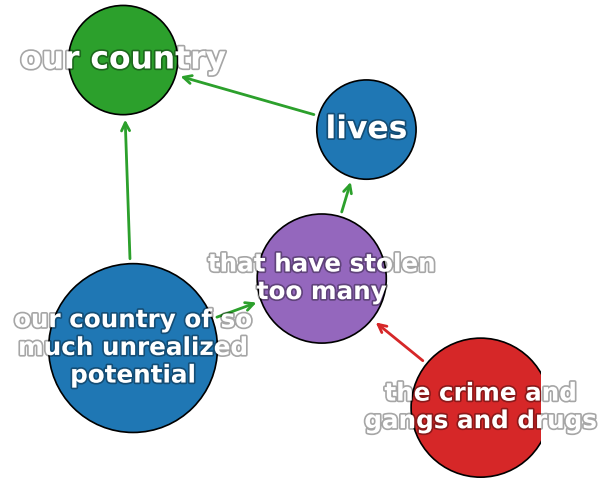
(a) Empowerment (US, +, US,)
I am fighting for you.



(b) Resistance (US, -, THEM,)
We tried and succeeded in reforming Wall Street.



(c) Elite Conspiracy (THEM, +, THEM,)
Donald Trump will deliver tax cuts to his billionaire donors.



(d) Victimization (THEM, -, US,)
The crime and gangs and drugs that have stolen too many lives and robbed our country of so much unrealized potential.

Figure A.3: Examples of Populist Motifs with Labels by Actor-Target Valence

Note: The graph structure represents the semantic path: **Actor** → **Judgment** → **Aspect** → **Target**. The extreme nodes represent the *actor* (where the arrow starts) and the *target* (where the arrow ends). Nodes are shaded to reflect group identity: actors or targets with the role **THEM** are marked in red color, and those with the role **US** in green. Edges reflect the overall valence of the statement: positive edges are rendered with green color and negative ones with red. Following the actor node, the *judgment* appears, with its tone indicated by styling—positive or neutral judgments (blue) to differ from negative ones (purple). The same logic applies to the *aspect* that follows the judgment.

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.¹⁹ 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.

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. The full prompt is reproduced below for transparency and replicability.

Prompt Used

```
prompt_1 = (
    "You are analyzing political speech to detect populist narratives. Your goal is to extract structured information "
    "from each sentence that identifies who is acting, what is being judged, what the judgment is, and who is affected. "
    "This is done by identifying the core elements of the sentence | ACTOR, ASPECT, JUDGMENT, TARGET | and their attributes.\n\n"

    "This framework is designed to identify **populist motifs**, especially those that portray a conflict between 'US' (the people) "
    "and 'THEM' (elites or outsiders), and rely on moral evaluations of actions and topics.\n\n"

    "### Important Concept: Valence Multiplication\n"
```

¹⁹As Reynolds and McDonell (2021) explain, few-shot examples activate pre-trained capabilities rather than teaching new tasks.

"Each judgment has two separate valence components:\n"

"- **JUDGMENT valence**: how the **action or evaluation** is judged (positive/neutral = 1, negative = -1)\n"

"- **ASPECT valence**: whether the **thing being judged** is itself good or bad (positive/neutral = 1, negative = -1)\n"

"**Multiply** JUDGMENT_VALENCE × ASPECT_VALENCE to determine the *moral directionality* of the act:\n"

"- If the result is '1', the actor is doing something good to something good (or bad to something bad) → *positive effect for target*.\n"

"- If the result is '-1', the actor is doing something bad to something good (or good to something bad) → *harmful effect for target*.\n"

Required Fields:\n"

"- 'ACTOR': Who is performing the action. Use 'UNKNOWN' if implicit. Add 'role': 'US' or 'THEM' or 'NA'.\n"

"- 'ASPECT': The topic or object being judged. Always required. Must include a valence.\n"

"- 'JUDGMENT': What action or opinion is expressed. Must include a valence.\n"

"- 'TARGET': Who is affected by the judgment. Use 'UNKNOWN' if implicit. Add 'role': 'US' or 'THEM' or 'NA'.\n"

"- 'JUDGMENT_HOLDER': (optional) who is expressing the judgment (e.g., the speaker or a quoted source). Also has optional 'negation'.\n"

"- In some cases, the ASPECT and TARGET will be the same span (e.g., 'us') | especially when the judgment is a harmful action affecting the in-group.\n"

"- This is valid and should not be treated as a conflict.\n"

Valence Tip:\n"

"- Some verbs like 'spent', 'signed', 'used', or 'implemented' are not inherently positive or negative.\n"

"- In these cases, assign a **positive valence (1)** to the judgment, and let the **ASPECT's valence** carry the meaning.\n"

"- Example: 'They spent millions trying to stop us' → judgment = 1 (spent), aspect = -1 (stop us) → result = -1\n"

"- For identity or possession statements (e.g., 'is', 'is not', 'belongs to'):\n"

"- The **judgment valence** reflects affirmation or denial.\n"

"- 'is', 'belongs to' → judgment = 1\n"

"- 'is not', 'does not belong to', 'cannot be' → judgment = -1\n"

"- The **aspect valence** reflects the object being discussed (e.g., 'money', 'surplus') | often positive (1).\n"

"- Example: 'The surplus is not the government's money' → judgment = -1, aspect = 1 → result = -1\n"

"- Example: 'The surplus is the people's money' → judgment = 1, aspect = 1 → result = +1\n"

"- When a speaker criticizes actors (e.g., 'politicians', 'elites') for harming good things like 'security', 'prosperity', or 'freedom':\n"

"- These **aspects should retain positive valence (1)**.\n"

"- The **judgment reflects the harm**: e.g., 'sacrificed', 'sold out', 'betrayed' → valence = -1\n"

"- Example: 'They sacrificed our security' → judgment = -1, aspect = 1 → result = -1\n"

"- Example: 'They sold out the country' → judgment = -1, aspect = 1 → result = -1\n"

"- If someone corrupts or damages something that is normally good (e.g., 'the system', 'our democracy'):\n"

"- The **aspect remains positive (1)** even if modified by negative adjectives like 'rigged' or 'broken'.\n"

"- Example: 'They rigged the system' → judgment = -1, aspect = 1 → result = -1\n"

"- Do **not** assign aspect valence = -1 just because it's described as 'rigged'.\n"

"- Do not assign negative valence to an ASPECT just because it includes abstract or elite-sounding phrases (e.g., 'status quo', 'bureaucracy', 'Washington') unless they are explicitly the focus of the judgment.\n"

"- Often, the **true aspect is the in-group** or values being harmed.\n"

"- Example: 'They have fought us' → judgment = -1, aspect = 'us' → valence = 1 → result = -1\n"

Output Format:\n"

"Write only the part inside 'judgment_relations' in valid JSON format:\n"

```
[
  {
    "ACTOR": {
      "text": "Actor text",
      "start": start_idx,
      "end": end_idx,
      "role": "US or THEM or NA"
    },
    "ASPECT": {
      "text": "Aspect text",
      "start": start_idx,
      "end": end_idx,
      "valence": valence
    },
    "JUDGMENT": {
      "text": "Judgment text",
      "start": start_idx,
      "end": end_idx,
      "valence": valence
    },
    "TARGET": {
      "text": "Target text",
      "start": start_idx,
      "end": end_idx,
      "role": "US or THEM or NA"
    },
    "JUDGMENT_HOLDER": {
      "text": "Speaker",
      "negation": 0
    }
  }
]
```

Important:\n"

"- Do **not** include explanations or comments in the final output. Use the reasoning to label, but do not print it out.\n"

"- Only output valid JSON as shown.\n"

Examples:\n"

Sentence: 'We will build the greatest infrastructure on the planet earth - the roads and railways and airports of tomorrow.'\n"

```
[
  {
    "ACTOR": {
      "text": "We",
      "start": 0,
      "end": 2,
      "role": "US"
    },
    "ASPECT": {
      "text": "airports of tomorrow",
      "start": 91,
      "end": 111,
      "valence": 1
    },
    "JUDGMENT": {
      "text": "will build",
      "start": 3,
      "end": 13,
      "valence": 1
    },
    "TARGET": {
      "text": "UNKNOWN",
      "start": 0,
      "end": 0,
      "role": "NA"
    }
  }
]
```

Explanation: 1 (build) × 1 (airports) = +1 → actor is doing something good to something good → net positive.\n"

Sentence: 'Donald Trump intends to use the United States military against American citizens who simply disagree with him.'\n"

```
[
  {
    "ACTOR": {
      "text": "Donald Trump",
      "start": 0,
      "end": 12,
      "role": "Them"
    },
    "ASPECT": {
      "text": "the United States military against American citizens",
      "start": 28,
      "end": 109,
      "valence": -1
    },
    "JUDGMENT": {
      "text": "intends to use",
      "start": 13,
      "end": 27,
      "valence": 1
    },
    "TARGET": {
      "text": "American citizens",
      "start": 63,
      "end": 109,
      "role": "US"
    }
  }
]
```

```

" ]\n"
"]\n"
"# Explanation: 1 (intends to use) * -1 (military against citizens) = -1 → actor supports something harmful → net negative.\n\n"

"Sentence: 'I am fighting against billionaires who raid our country.'\n"
"\n"
"  {\n"
"    {\n"
"      'ACTOR': {'text': 'I', 'start': 0, 'end': 1, 'role': 'US'},\n"
"      'ASPECT': {'text': 'billionaires who raid our country', 'start': 19, 'end': 49, 'valence': -1},\n"
"      'JUDGMENT': {'text': 'fighting against', 'start': 2, 'end': 18, 'valence': -1},\n"
"      'TARGET': {'text': 'our', 'start': 45, 'end': 49, 'role': 'US'}\n"
"    },\n"
"    'NA'\n"
"  ],\n"
"  {\n"
"    {\n"
"      'ACTOR': {'text': 'billionaires', 'start': 19, 'end': 30, 'role': 'THEM'},\n"
"      'ASPECT': {'text': 'our country', 'start': 45, 'end': 49, 'valence': 1},\n"
"      'JUDGMENT': {'text': 'raid', 'start': 31, 'end': 35, 'valence': -1},\n"
"      'TARGET': {'text': 'our', 'start': 45, 'end': 49, 'role': 'US'}\n"
"    },\n"
"    'NA'\n"
"  ],\n"
"  ]\n"
"]\n"
"# Explanation:\n"
"# Outer valence: -1 (fighting) * -1 (billionaires who raid) = +1 → speaker is opposing a bad thing → net positive.\n"
"# Inner valence: -1 (raid) * 1 (our country) = -1 → THEM is harming something good → net negative.\n"
"# This is a textbook populist motif: THEM hurt US; speaker protects US.\n\n"

"Sentence: 'The politicians that have sacrificed their security, betrayed their prosperity, and sold out their country.'\n"
"\n"
"  {\n"
"    {\n"
"      'ACTOR': {'text': 'the politicians', 'start': 45, 'end': 59, 'role': 'THEM'},\n"
"      'ASPECT': {'text': 'their security', 'start': 66, 'end': 81, 'valence': 1},\n"
"      'JUDGMENT': {'text': 'sacrificed', 'start': 60, 'end': 71, 'valence': -1},\n"
"      'TARGET': {'text': 'the American people', 'start': 5, 'end': 23, 'role': 'US'}\n"
"    },\n"
"    'NA'\n"
"  ],\n"
"  {\n"
"    {\n"
"      'ACTOR': {'text': 'the politicians', 'start': 45, 'end': 59, 'role': 'THEM'},\n"
"      'ASPECT': {'text': 'their prosperity', 'start': 83, 'end': 99, 'valence': 1},\n"
"      'JUDGMENT': {'text': 'betrayed', 'start': 72, 'end': 80, 'valence': -1},\n"
"      'TARGET': {'text': 'the American people', 'start': 5, 'end': 23, 'role': 'US'}\n"
"    },\n"
"    'NA'\n"
"  ],\n"
"  {\n"
"    {\n"
"      'ACTOR': {'text': 'the politicians', 'start': 45, 'end': 59, 'role': 'THEM'},\n"
"      'ASPECT': {'text': 'their country', 'start': 101, 'end': 113, 'valence': 1},\n"
"      'JUDGMENT': {'text': 'sold out', 'start': 100, 'end': 108, 'valence': -1},\n"
"      'TARGET': {'text': 'the American people', 'start': 5, 'end': 23, 'role': 'US'}\n"
"    },\n"
"    'NA'\n"
"  ],\n"
"  ]\n"
"]\n"
"Explanation: In the example each clause represents THEM harming something good that belongs to US → all should result in -1."

```

```

### Your Task:\n

```

```

"Analyze the sentence and return only the JSON object inside 'judgment_relations'. Do not include explanations, reasoning, or additional comments in your output.\n"
)

```

Model Setting

```

# Function to classify a sentence using OpenAI's correct API format
def classify_sentence(sentence):
    prompt = f"{prompt_1}\n\nSentence: {sentence}"

    response = openai.chat.completions.create(
        model="gpt-4o-mini-2024-07-18",
        messages=[
            {"role": "system", "content": "You are an assistant that helps categorize sentences."},
            {"role": "user", "content": prompt}
        ],
        max_tokens=500,
        temperature=0
    )

    classification = response.choices[0].message.content
    return classification

```

A.5.2 Synthetic Data Generation

To expand the training set while preserving the substantive meaning of the original sentences, we generated synthetic variants of annotated political speech using a separate prompting step. The goal of this augmentation procedure was to create alternative formulations that retained the original sentence’s meaning and sentiment while introducing lexical and syntactic variation. This allowed us to increase the diversity of the training data without altering the underlying rhetorical content being classified.

The augmentation prompt explicitly instructed the model to generate sentences that were logically equivalent to the original, while varying wording and grammatical structure. Importantly, the prompt emphasized that the roles of the key elements in the annotation framework should remain clear, even if the sentence structure changed. Each original sentence was used to generate five alternative versions.

Prompt Used

```
prompt = (
    "You are an AI trained to provide high-quality synonyms specifically tailored for political speech. "
    "Given a JSON structure containing a full sentence under 'text', and labeled components (Actor, Judgment, Aspect, Target), "
    "provide exactly 5 contextually accurate synonyms for each component, strictly preserving sentiment, logic, and clarity of roles. "
    "If a component (Actor, Judgment, Aspect, Target) is labeled as 'UNKNOWN', do not create synonyms and return 'UNKNOWN'. "
    "Ensure each synonym is suitable for political discourse, natural-sounding, and avoids redundancy. "
    "Return a structured JSON precisely formatted as follows:\n\n"
    "{\n"
    "  'ACTOR': ['synonym1', 'synonym2', ..., 'synonym5'],\n"
    "  'JUDGMENT': ['synonym1', 'synonym2', ..., 'synonym5'],\n"
    "  'ASPECT': ['synonym1', 'synonym2', ..., 'synonym5'],\n"
    "  'TARGET': ['synonym1', 'synonym2', ..., 'synonym5']\n"
    "}\n\n"
    f"Input JSON:\n{json.dumps(json_input, indent=2)}"
)
```

Model Setting

```
# Function to classify a sentence using OpenAI's correct API format
def classify_sentence(sentence):
    prompt = f"{prompt_augment_step1}\n\nSentence: {sentence}"

    response = openai.chat.completions.create(
        model="gpt-4o-mini-2024-07-18",
        messages=[
            {"role": "system", "content": "You are an assistant that helps categorize sentences."},
            {"role": "user", "content": prompt}
        ],
        max_tokens=400,
        temperature=0
    )

    classification = response.choices[0].message.content
    return classification
```

The synthetic data generation step was designed to preserve label-relevant semantic structure while increasing surface-level diversity. By generating multiple paraphrased variants of each sentence, this procedure helped reduce the dependence of downstream models on narrow lexical patterns and improved robustness to variation in phrasing. Because the prompt constrained the model to preserve logical equivalence and the clarity of the core roles in the sentence, the augmented examples remained aligned with the original annotation framework.

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

Section A.5.3 describes the GED metric in detail. In this appendix, we provide additional implementation notes on our GED process and evaluation. First, 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.

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: 100 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.

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 100 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.

Label	Attribute	Precision	Recall	F1
Actor	Them	0.83	0.88	0.85
	Residual	0.87	0.92	0.89
	Us	0.91	0.82	0.86
Target	Them	0.70	0.76	0.73
	Residual	0.88	0.84	0.86
	Us	0.84	0.85	0.84
Judgment	-1	0.86	0.82	0.84
	+1	0.91	0.93	0.92
Aspect	-1	0.67	0.61	0.64
	+1	0.76	0.80	0.78

Table A.2: Sentence-level test-set evaluation metrics.

Note: Results are based on a hybrid pool of annotations, combining labels from multiple annotation sources into a single training set.

A.7 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 Figures A.5–A.6.

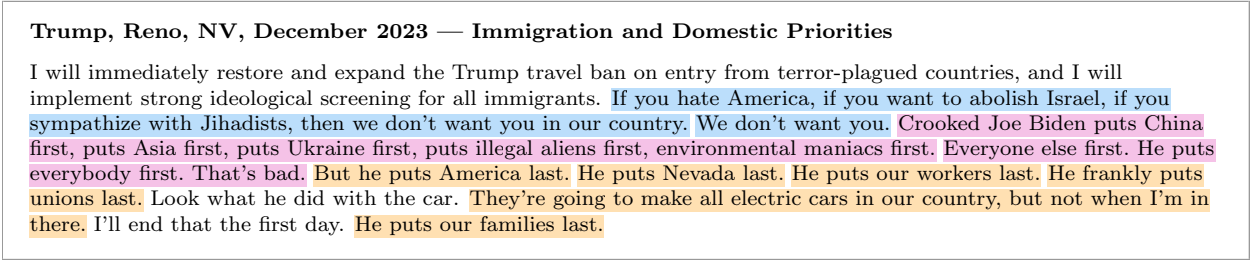
The first paragraph, which focuses on immigration, foreign policy, and domestic priorities, features a series of declarative judgments that consistently portray the out-group (e.g., “Crooked Joe Biden,” “environmental maniacs”) as prioritizing others over “our workers,” “our families,” or the country itself. Out of 13 sentences, 10 are classified as populist according to our annotation criteria. This segment exemplifies 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. In this case, 4 out of 9 sentences are classified as populist, all mapping onto the victimization motif.

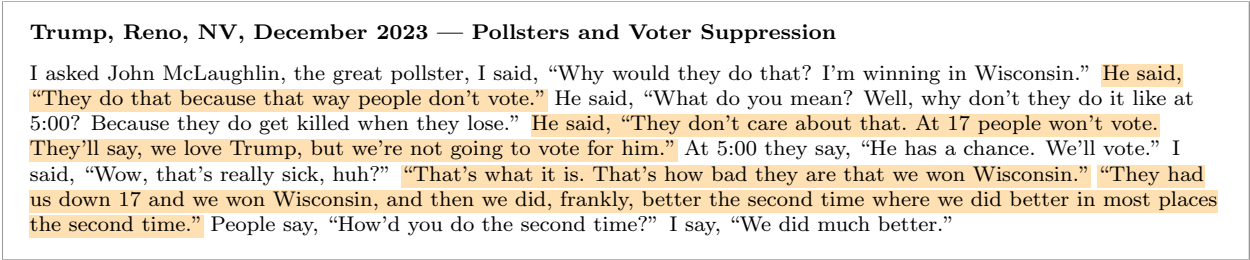
Turning to a leader typically labeled as non-populist by previous classification methods, we find that Barack Obama’s August 2024 speech in Chicago nonetheless contains a high proportion of populist motifs: 7 out of 10 sentences follow one of the defined populist structures (see Figure A.7). Most prominently, Obama relies on empowerment motifs emphasizing in-group moral strength and unity—demonstrating how non-populist leaders can strategically draw on populist elements to mobilize support and establish moral contrast.

These excerpts are presented in Figures A.5–A.7 with motif-specific color coding that matches the scheme used in Figure 6: elite conspiracy (THEM helps THEM), victimization (THEM harms US), empowerment (US helps US), and resistance (US opposes THEM). Uncolored sentences are not classified as populist.



Color key: Elite Conspiracy Victimization Empowerment Resistance Uncolored = not populist.
10 of 13 sentences classified as populist.

Figure A.5: Motif-annotated excerpt from Trump’s December 2023 Reno rally. Victimization motifs dominate, with elite-conspiracy framing when attributing motive to Biden.



Color key: Elite Conspiracy Victimization Empowerment Resistance Uncolored = not populist. 4
of 9 sentences classified as populist.

Figure A.6: Motif-annotated excerpt from Trump’s December 2023 Reno rally: pollster anecdote. All populist sentences are coded as victimization, portraying hostile actors as discouraging in-group political participation.

Obama, Chicago, IL, August 2024 — DNC Address

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.

Color key: Elite Conspiracy Victimization Empowerment Resistance Uncolored = not populist. 7 of 10 sentences classified as populist.

Figure A.7: Motif-annotated excerpt from Obama's August 2024 DNC address. Empowerment motifs dominate, with one victimization sentence directed at the opposing party and two resistance sentences framing the election as a fight. The contrast with Trump's Reno speech illustrates that populist motifs are not exclusive to conventionally populist politicians and that their composition varies across speakers and contexts.

A.8 Tables and Figures

Table A.3: Relationship Between Populist Motifs and Dictionary-Based Measures (refers to Figure 9)

	Occurrence (OLS)	Dummy (Logit)
Intercept	0.195*** (0.007)	-1.672*** (0.043)
Dictionary	0.023*** (0.007)	
Dictionary Dummy		0.414*** (0.080)
R-squared	0.002	
AIC	7162.97	4891.59
Pseudo R-squared		0.005
N	5331	5331

Notes: The table reports regressions examining the relationship between populist motifs identified through the motif extraction procedure and dictionary-based measures of populist language. Column (1) reports OLS estimates where the dependent variable is the standardized count of populist motifs. Column (2) reports a logistic regression where the dependent variable is a binary indicator for the presence of a populist motif.

Dictionary variables indicate whether populist language is detected using a predefined populism dictionary. Positive coefficients imply that texts flagged by the dictionary are more likely to contain motifs identified by the motif extraction approach.

Standard errors are reported in parentheses. Statistical significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4 presents correlation between our measure and Dai and Kustov (2022). The Dai-Kustov corpus used in the analysis contains 587 annotated speech chunks, which expand to 16,618 sentence-level observations. A chunk is a contiguous speech excerpt from a single speaker that may contain multiple paragraphs and is treated as the aggregation unit for the chunk-level comparisons; in this sample, chunks contain between 4 and 20 paragraphs and between 8 and 111 sentences. The corpus covers 6,185 paragraphs, 34 speakers, and 25 election years ranging from 1912 to 2018. The speakers included are Adlai Stevenson, Barack Obama, Bernie

Sanders, Bill Richardson, Donald J. Trump, Dwight Eisenhower, Franklin D. Roosevelt, Fred Thompson, George Bush, George McGovern, George W. Bush, Hillary Clinton, Hubert H. Humphrey, Jeb Bush, Jimmy Carter, John Edwards, John F. Kennedy, John F. Kerry, John McCain, Martin O'Malley, Michael S. Dukakis, Michele Bachmann, Mitt Romney, Newt Gingrich, Richard Nixon, Rick Perry, Rick Santorum, Robert Dole, Ron Paul, Ronald Reagan, Ted Cruz, Walter F. Mondale, William Howard Taft, and William J. Clinton. The years represented are 1912, 1932, 1952, 1956, 1960, 1968, 1972, 1976, 1980, 1984, 1988, 1992, 1996, 2000, 2003, 2004, 2007, 2008, 2009, 2011, 2012, 2014, 2015, 2016, and 2018.

Table A.4: Relationship Between Populist Motifs and the Dai-Kustov Measure (refers to Figure 9b)

	Occurrence (OLS)	Dummy (Logit)
Intercept	2.535*** (0.114)	1.455*** (0.118)
Dai_Kustov	0.440* (0.250)	0.671** (0.316)
R-squared	0.005	
AIC	2722.36	537.95
Pseudo R-squared		0.009
N	587	587

Notes: The table reports regressions examining the relationship between populist motifs identified through the motif extraction procedure and the Dai-Kustov measure of populist rhetoric. Column (1) reports OLS estimates where the dependent variable is the count of populist motifs. Column (2) reports a logistic regression where the dependent variable is a binary indicator for the presence of a populist motif. The corresponding coefficient plots are shown in Figure 9b.

The Dai-Kustov variable indicates whether a text is classified as populist using the Dai-Kustov measure. The data comes from replication data from Dai and Kustov (2022), the unit of analysis is text-chunk (a chunk is 1-3 paragraphs). Positive coefficients imply that texts classified as populist by that measure are more likely to contain motifs identified by the motif extraction approach.

Standard errors are reported in parentheses. Statistical significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Regression Estimates Corresponding to Figure 10

	Occurrence	Dummy
Intercept	0.438*** (0.117)	0.377*** (0.084)
Workers Exp. to Auto	0.000*** (0.000)	0.000** (0.000)
Close	0.055** (0.023)	0.046** (0.018)
Share of Foreign workers	-0.029*** (0.009)	-0.025*** (0.006)
Hate inc x 100K Pop	-0.000 (0.000)	-0.001 (0.000)
State FE	Yes	Yes
R-squared	0.495	0.504
AIC	-188.49	-220.86
N	70	70

Notes: The table reports the regression estimates corresponding to the coefficient plots shown in Figure 10. Column (1) uses DV as the counting of populist occurrences, and column (2) uses a dummy. Coefficients correspond to the point estimates plotted in the figure, and standard errors are reported in parentheses. The unit of analysis is speech.

Table A.6: Regression Estimates Corresponding to Figure 11

	Victimization	Empowerment	Elitist_Conspiracy	Resistance
Intercept	0.025*** (0.007)	0.032*** (0.009)	0.012*** (0.004)	0.017*** (0.002)
Workers Exp. to Auto	0.000 (0.000)	0.000* (0.000)	0.000 (0.000)	0.000 (0.000)
Close	0.016** (0.007)	-0.012 (0.008)	0.015*** (0.004)	0.003** (0.001)
State FE	Yes	Yes	Yes	Yes
R-squared	0.307	0.374	0.351	0.417
AIC	-298.61	-241.34	-341.95	-427.71
N	70	70	70	70

Notes: The table reports the regression estimates underlying the coefficient plots shown in Figure 11. The dependent variable is a dummy equal to one when a text contains a populist motif. Column (1) estimates the relationship between swing elections (defined as states where the 2012 presidential election was decided by less than 5 percentage points) and populist motifs. Column (2) estimates the relationship between exposure to at-risk workers and populist motifs. Coefficients correspond to the point estimates shown in the figure, and standard errors are reported in parentheses. Confidence intervals displayed in Figure 11 correspond to 90% intervals constructed from these estimates.

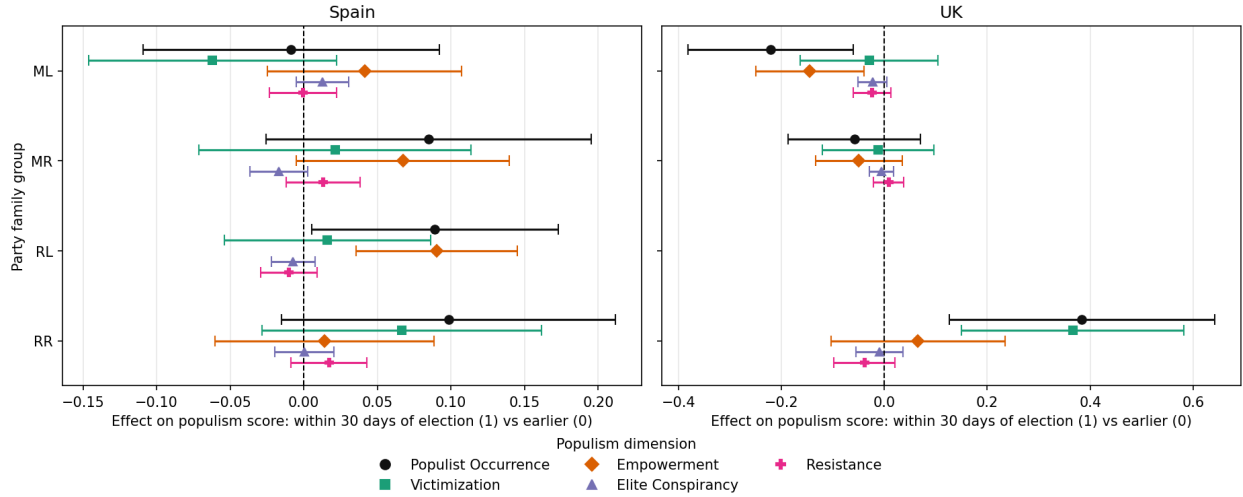


Figure A.8: Populist Motifs in Campaign Videos Close to Elections: Spain and the United Kingdom

Notes: The figure reports coefficient estimates from regressions of populist rhetoric in party YouTube videos on an indicator for videos posted within 30 days of an election. The horizontal axis shows the estimated effect of posting close to an election (within 30 days = 1) relative to earlier campaign communication (0).

The analysis uses party videos posted during the year preceding each national election. The Spain sample includes elections from 2014–2023 (year prior to the 2015, 2016, 2019a, 2019b, and 2023 elections), while the United Kingdom sample covers 2014–2024 (year prior to the 2015, 2017, 2019, and 2024 elections). The dataset contains 8,753 videos collected from official party YouTube channels. Markers correspond to different dimensions of the Populist-PULSAR measure: overall populist occurrence, victimization, empowerment, elite conspiracy, and resistance motifs. Rows group parties by ideological family (ML, MR, RL, RR). All models include country and period fixed effects and are estimated on pooled data across elections.

A.9 Identifying Distinctive Words Across Motifs

To identify words that are most characteristic of each populist motif, we employ the “fighting words” method. This approach compares word frequencies between two sets of documents while adjusting for sampling variability and rare terms.

For each word w and motif i , we compute a weighted log-odds ratio comparing the frequency of the word in texts assigned to motif i with its frequency in the remainder of the corpus. Let y_{wi} denote the number of times word w appears in motif i , and n_i the total number of tokens in that motif. Let $y_{w\neg i}$ and $n_{\neg i}$ denote the corresponding counts in all other motifs.

To stabilize estimates for low-frequency words, the method incorporates an informative Dirichlet prior with parameter α_w for each word and $\alpha_0 = \sum_w \alpha_w$. The weighted log-odds ratio is:

$$\delta_w^{(i)} = \log \left(\frac{y_{wi} + \alpha_w}{n_i + \alpha_0 - (y_{wi} + \alpha_w)} \right) - \log \left(\frac{y_{w\neg i} + \alpha_w}{n_{\neg i} + \alpha_0 - (y_{w\neg i} + \alpha_w)} \right)$$

The variance of this estimate is approximated by

$$\sigma^2(\delta_w^{(i)}) = \frac{1}{y_{wi} + \alpha_w} + \frac{1}{y_{w\neg i} + \alpha_w}.$$

We report standardized scores

$$z_w^{(i)} = \frac{\delta_w^{(i)}}{\sqrt{\sigma^2(\delta_w^{(i)})}},$$

which measure how strongly each word is associated with motif i relative to the rest of the corpus. Larger positive values indicate that a word is more distinctive of the motif, while negative values indicate that the word is relatively more common in other motifs.

In the heatmap figures in the main text, we display the words with the largest absolute z -scores for each motif.

A.10 LLM-Based Topic Classification

To classify the substantive content of the texts, we use a large language model (LLM) to assign each observation to a set of predefined policy topics (e.g., corruption, economy, governance). The classification is implemented through prompt-based annotation using the OpenAI API.

Specifically, we use the model `gpt-4.1-mini`. The model is queried through the Chat Completions API with temperature set to zero in order to produce consistent classifications. Observations are processed in batches of 20 texts per request, and the model is instructed to return its predictions in structured JSON format.

The prompt defines the set of allowed topic labels and instructs the model to assign the most appropriate category based on the semantic content of the text. The model is explicitly instructed not to generate new categories and to select only from the predefined label set. This constraint ensures consistency across observations and facilitates reproducibility.

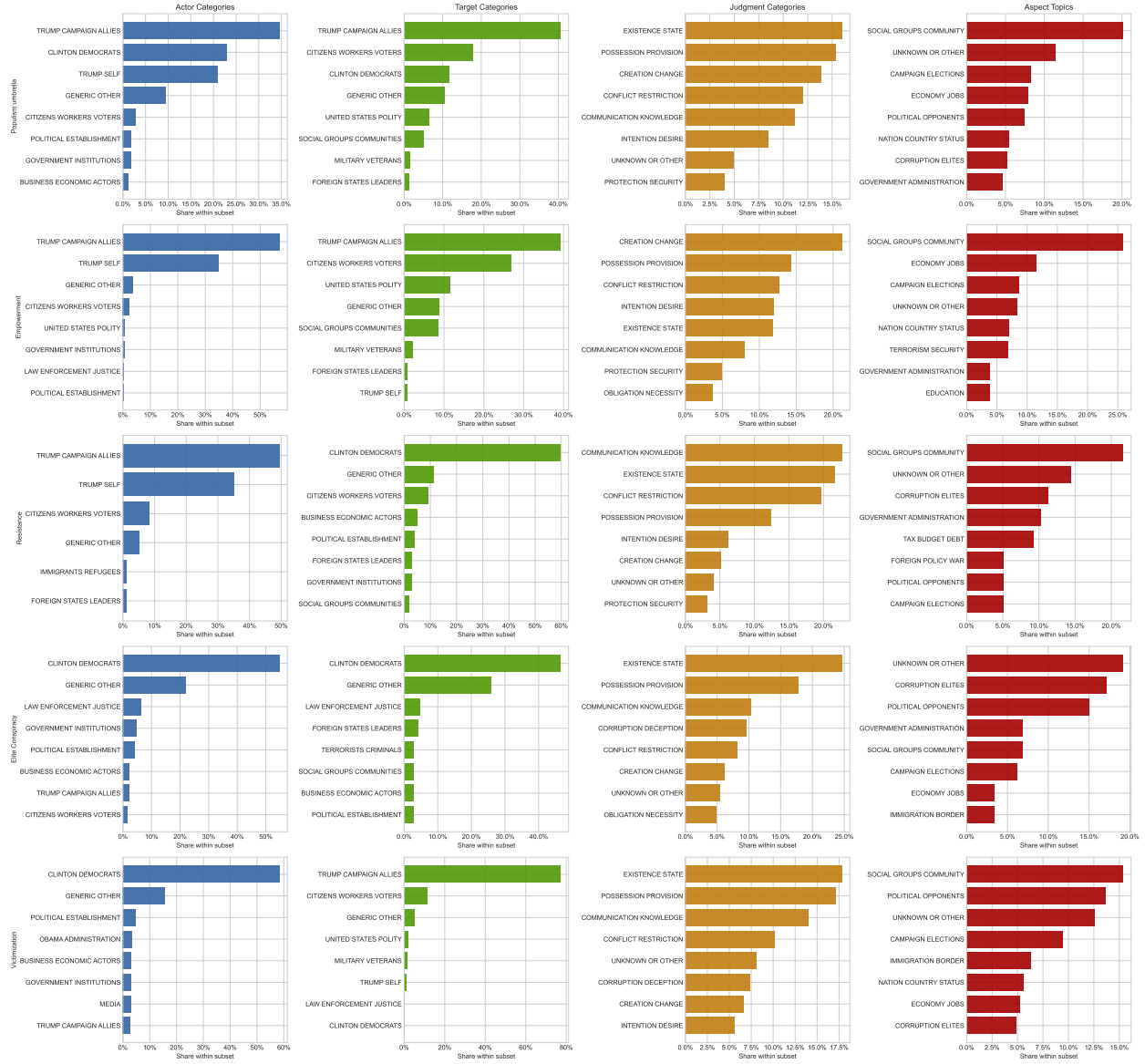


Figure A.9: Populist umbrella and exact motifs across actor, target, judgment, and issue categories

Notes: This figure shows the distribution of annotations across four dimensions: actor categories, target categories, judgment categories, and aspect topics. The top row reports the overall populism umbrella, while the remaining panels report the exact motifs: Empowerment, Resistance, Elite Conspiracy, and Victimization. Bars indicate the share within each subset. Category construction was based on LLM-assisted annotations of the relational structure of each text, including who is speaking about whom, in what evaluative or action-oriented terms, and on which issue dimension; details of the annotation framework are provided in Appendix A.10. For the corresponding within-motif heatmaps, see Figure 13.