

**Framing the “Nation”: Discourse of National Identity Construction in the Taiwan Region—
An Analysis of 2024 Electoral Campaign**

*Junyu Jiang¹ and Zhuoling Li²

Communication University of China¹

Nankai University²

Author Note

*These authors contributed equally to the work.

Junyu Jiang, Computational Social Science, University of California, San Diego,
juj013@ucsd.edu;

Zhuoling Li, Zhou Enlai School of Government, Nankai University,
2112511@mail.nankai.edu.cn. This research did not receive any specific grant from funding
agencies in the public, commercial, or not-for-profit sectors.

Abstract

Constructivist accounts hold that nations are discursively produced, but we still know little about whether electoral discourse organises the signifier “nation” into stable equivalential chains or whether meanings emerge ad hoc across contexts. The ambiguity between minzu (ethnic people) and guojia (state/polity) in Chinese contexts further enlarges the semantic space through which political actors contest national identity. Treating “nation” as an empty signifier in Laclau’s sense, this article examines how its meaning is fixed, pluralised, or left indeterminate during Taiwan’s 2024 presidential campaign. Empirically, we analyse a multi-source corpus from debates, rallies, speeches, news reports, and social media, treating sentences with ± 1 sentence of local context as the unit of analysis. Building on a dual-layer framing scheme centred on national identity, we use gpt-4o to annotate frames at scale. First, we find that national identity discourse is structured rather than ad hoc: candidates and parties will repeatedly mobilise distinct equivalential chains, and these patterns vary systematically across venues and stages. Additionally, a small set of signifiers will preferentially act as nodal points that are both structurally central and context-sensitive, being preferentially absorbed into civic frames in domestic policy talk but shifting toward ethnic or security frames in cross-Strait and international discussions. Lastly, we find that antagonistic framing organises multiple intersecting axes of othering, including cross-Strait, partisan, generational, and spatial divisions. Results show that “nation” in Taiwan operates as a genuinely floating signifier whose meanings shift in systematic relation to institutional positions and situational stakes.

Framing the “Nation”: Discourse of National Identity Construction in the Taiwan Region— An Analysis of 2024 Electoral Campaign

Research on nations and nationalism has increasingly converged on a constructivist view: nations are historically contingent, socially produced, and sustained through narratives, symbols, and institutional practices rather than primordial ties of blood or culture (Anderson, 1991; Calhoun, 1997; Dar, 2022; Gellner, 1983; Machin, 2015; Miller, 1995; Oran F., 2022; Smith, 1986; Wodak, 2009; Bieber, 2020). Critical discourse analysis and related approaches show how media texts, political speeches, and everyday talk repeatedly reproduce and contest national identities, often by naturalising particular boundaries between “us” and “them” (Antonsich, 2015; Machin, 2013; Wodak et al., 2009). Yet in research on national identity and electoral communication, empirical applications of these ideas still rely largely on small-N, case-based readings that demonstrate that “the nation is contested” without providing systematic, corpus-level measures of how equivalential chains, nodal signifiers, and antagonistic frontiers are organised. A small but growing line of work begins to integrate discourse theory with corpus linguistics and network-based text analysis, but these efforts remain rare and are concentrated mostly on populism rather than on the organisation of national identity discourse in election campaigns (Bakumov, 2018; Nikisianis, 2018; Semykina, 2021).

These questions are especially acute in Chinese-language settings. The standard translations of “nation” carry a built-in duality—most notably *minzu*, which often evokes an ethnic or civilisational people, and *guojia*, which denotes the state or polity (Berry, 1992; Bilik, 2007; Cao, 2021; Connor, 1978; Lee, 2013; Ma, 2017; Banerjee, 2021; Brubaker, 2004). This *minzu–guojia* tension creates a durable field of semantic ambiguity that is particularly amenable to political interpretation and manipulation. In Taiwan’s political context, where historical legacies, evolving cross-Strait relations, and internal debates over collective identity converge, this ambiguity renders the semantics of nation-related discourse unusually consequential (Cabestan, 2005; Jiang, 2017; Lee, 2020; Shen, 2013). In the 2024 presidential campaign, it is concentrated in a small cluster of recurrent terms—*Taiwan*, *Zhonghua Minguo* (Republic of

China), *minzu*, *guojia*, and “democracy”—whose referents and implied boundaries shift across speakers, venues, and audiences, providing a window into how political actors exploit semantic flexibility to articulate competing visions of identity, legitimacy, and political future.

From a discourse-theoretical perspective, such variability can be understood through the lens of Laclau and Mouffe’s notion of the nation as an empty or floating signifier, whose meaning is only contingently fixed through nodal points, chains of equivalence, and antagonistic boundaries (Laclau & Mouffe, 1985; Laclau, 2005). Framing research in political communication complements this view by showing how actors strategically assemble and cue particular interpretations in concrete messages (Chong & Druckman, 2007). Taken together, these approaches invite us to move beyond asking whether the nation is contested, and to examine how contestation is structured: which framings of the nation tend to travel together, which signifiers occupy central positions in these configurations, and along which axes “we/they” boundaries are repeatedly drawn.

Against this backdrop, the present study analyses campaign communication from Taiwan’s 2024 presidential election to trace the large-scale organisation of national identity discourse. Drawing on a multi-source corpus that spans televised debates and policy presentations, campaign rallies and speeches, news coverage, and social media posts, we treat sentence-level mentions of the nation and related concepts as our basic units and ask how they are woven into broader discursive configurations. Specifically, we address three research questions:

RQ1: How do different candidates and parties in Taiwan’s 2024 presidential election organise national identity into recurrent equivalential chains in their campaign communication?

RQ2: Which high-frequency nation-related signifiers occupy structurally central, nodal positions within these chains?

RQ3: How are “we” and “they” configured, and along which axes are antagonistic boundaries most frequently drawn across actors and venues?

Based on existing work on Taiwanese identity, civic and ethnic nationalism, and cross-Straits politics, we advance two expectations.

H1. In communication primarily oriented toward domestic audiences, nation and national identity are articulated predominantly through civic framings that emphasise shared responsibility, institutions, and future-oriented visions of collective life.

H2. In cross-Strait or outward-facing contexts, explicitly ethnic or civilisational framings—often indexed by *minzu* and related vocabulary—are more likely to be activated, particularly in conjunction with crisis and enemy framings.

Nation as Contested Signifier in the Chinese–Taiwanese Context

Constructivist accounts of nationalism from Anderson’s “imagined communities” to Gellner’s modernism and Smith’s ethnosymbolism converge on a basic distinction between *nation* as a collective identity, *state* as a sovereign political organisation, and *ethnicity* as shared descent or culture (Anderson, 1991; Gellner, 1983; Smith, 1986; Calhoun, 1997). This distinction becomes especially salient in Chinese-language debates over *minzu* and *guojia*, where national identity can be anchored in both institutional membership and cultural genealogy.

The complexity of “nation” intensifies when the concept is translated into Chinese linguistic and historical contexts. Since the late Qing, intellectuals and officials experimented with different renderings of “nation”, importing *minzu* via Japanese political vocabulary and pairing it with *guojia* to describe a “unified multi-ethnic state” (统一的多民族国家) (Zhao, 2000). In this configuration, *minzu* initially carried strong racial and genealogical connotations and was later crystallised in Republican and PRC discourse as *Zhonghua minzu*, a single “Chinese nation” encompassing Han, officially recognised minority groups, and overseas Chinese communities (Carlson et al., 2016). By contrast, *guojia* denoted the territorial state and its governing apparatus but in everyday usage also invoked family and homeland, further blurring the boundaries between ethnic peoplehood, national community, and state institutions (Wang, 2024). The resulting *minzu–guojia* configuration has often been described as a core site of semantic instability, where “nation” can variously invoke a civilisational ethnus, a modern state, or a diffuse cultural community depending on context.

This linguistic ambivalence has substantive political implications. Scholarship on PRC

state-led nationalism shows that official discourse often defines *minzuzhuyi* as loyalty to a trans-ethnic *Zhonghua minzu*—a multi-ethnic “Chinese nation” that formally encompasses the 56 recognised *minzu* as well as ethnic Chinese communities in Taiwan and overseas. In this formulation, appeals to shared ancestry, cultural continuity, and a history of national humiliation are mobilised to legitimate projects of national unity and territorial integration (e.g., Carlson et al., 2016; Cao, 2021). At the same time, research on language, politics, and identity documents how the symbolic pairing and decoupling of “Taiwan” and “China” in official rhetoric, educational materials, and news coverage gradually shifts the imagined centre of identity from an orthodox Chinese nation toward a localised Taiwanese community, without fully erasing references to *Zhonghua minzu* (Chang, 2014; Huang, 2009; Lai, 2018; Wu, 2007). This slippage between people, state, and ethnos generates a wide semantic field in which the meaning of “nation” can be flexibly interpreted, layered, or strategically obfuscated.

Within Taiwan, these ambiguities intersect with competing projects of nation-building. Building on this background, scholarship has explicitly imported the distinction between ethnic and civic nationalism to differentiate alternative logics of Taiwanese nationhood (Wu, 2007; Wu & Shen, 2000; Ngeow, 2010; Lousche, 2022). Ethnic variants emphasise shared culture, language, provincial origin, and historical experience, whereas civic variants define the nation primarily in terms of attachment to political institutions, democratic procedures, and the de facto state on Taiwan, in principle open to citizens of diverse cultural backgrounds. Using nationwide survey data, Huang (2005) applies latent class analysis to the 2001 Taiwan Election and Democratization Study and shows that “Taiwanese/Chinese” identity and national identity form empirically distinct dimensions rather than a single ethnic cleavage. Subsequent work drawing on time-series data from the NCCU Election Study Center and panel surveys finds that since the early 1990s the major shift has occurred on the political–state dimension: identification with China as the relevant state has declined, while identification with Taiwan’s polity has strengthened, even as cultural attachments to a broader Chinese civilisation remain ambivalent and layered (Shen, 2013; Zhong, 2016). Democratisation and repeated electoral competition have

progressively reinforced a civic understanding of Taiwanese nationhood anchored in democratic norms and constitutional self-government, without fully erasing ethnic cleavages or cultural “Chineseness” (Wang, 2005; Chen, 2012, 2023; Alemán, 2018; Hur, 2020).

Language is central to these identity dynamics. Studies of political elites, media, and civic actors show how Taiwan and China are articulated through contested naming practices, narrative templates, and symbolic oppositions. Chang and Holt’s analyses of National Day speeches, for instance, document how successive ROC presidents have shifted labels for China—from derogatory “Communist bandits” to “Chinese Communists” and “the mainland”—and how these lexical choices reconfigure Taiwan’s imagined relationship to China and to the broader international community (Chang, 2011). Similarly, research on language, politics, and identity traces how the symbolic pairing and decoupling of “Taiwan” and “China” in official rhetoric, curricula, and news coverage gradually relocates the centre of identification from an abstract Chinese nation to a more localised Taiwanese community (Wu, 2007). In this sense, national identity in Taiwan is continuously renegotiated through the naming of political entities, the framing of cross-Straits relations, and the construction of internal Others.

Survey-based and discourse-analytic studies converge on the view that political discourse is a primary arena in which Taiwanese national identity is constructed and contested: through naming practices, antagonistic codings of “us” and “them”, and the legitimisation of particular constitutional and territorial arrangements. Taken together, this body of work delineates the empirical and conceptual dimensions of Taiwanese national identity that any formal account of discursive organisation needs to capture.

Laclau, Equivalential Chains, and Antagonistic Boundaries

Political discourse theory, most prominently developed by Laclau and Mouffe, reconceptualises social and political identities as contingent outcomes of discursive articulation rather than as pre-given essences (Laclau & Mouffe, 1985). In this view, meaning is only ever partially fixed in an open field of signification, through the linking and differentiation of elements into more or less durable configurations. Central to this approach are *nodal points*—privileged

signifiers that organise surrounding meanings—together with chains of equivalence and difference and antagonistic frontiers that structure the discursive field. Chains of equivalence connect heterogeneous demands or attributes by downplaying their internal differences so that they can be articulated as expressions of a common project, while chains of difference foreground distinctions among elements that are not subsumed under a shared identity. As an equivalential chain expands, the signifier that names and represents this chain becomes relatively “empty”: an empty or floating signifier that condenses diverse claims without being reducible to any of them (Laclau, 2005). Antagonistic frontiers, in turn, draw a line between this equivalential chain—often condensed in an empty signifier—and a constitutive outside that is constructed as its obstacle or enemy. From this vantage point, “the nation” is not a fixed referent but an empty or floating signifier that can condense heterogeneous demands into a relatively coherent chain of equivalences while simultaneously demarcating an antagonistic outside.

This logic has been widely taken up in analyses of populist discourse, which offer a concrete template for thinking about how equivalential chains and antagonistic frontiers are structured. Populist discourse has been described, in Laclau’s terms, as the construction of “the people” through the equivalential linkage of otherwise disparate social demands, articulated in opposition to an obstructive “elite” (Laclau, 2005; Hatakka, 2019). De Cleen and Stavrakakis (2017) refine this insight by distinguishing between a vertical axis opposing “the people” to “the elite” and a horizontal axis distinguishing insiders from outsiders to the nation. On their account, populism operates primarily along the vertical people–elite axis, while nationalism centres on the horizontal inside–outside boundary of the nation. In practice, these logics often intersect, such that “the people” are imagined as a national people pitted against both domestic elites and foreign “others”. This two-dimensional space highlights that discursive frontiers need not be singular: antagonisms can be drawn simultaneously along national, partisan, socio-economic, generational, and spatial lines.

Laclau-inspired political discourse theory has also informed a growing empirical literature, although its operationalisation has remained uneven. Much of this work relies on close

reading of relatively small corpora—such as key speeches, party manifestos, or selected media texts—to reconstruct empty signifiers, equivalential chains, and antagonistic frontiers qualitatively (e.g., Jørgensen & Phillips, 2002; Stengel, 2019). Analysts identify candidate empty signifiers, trace which demands or identities are articulated as equivalent under their banner, and map how adversaries are named and positioned across texts. A smaller but expanding set of studies experiments with more systematic, corpus- or network-based approaches that approximate discourse-theoretical concepts. Some contributions model discourse as a network of co-occurring elements and use centrality measures to infer which signifiers function as nodal points in a given field (Haditaghi et al., 2016), while others combine lexicometric techniques with narrative analysis to capture temporary “fixations” of meaning and the rules by which elements are linked or differentiated (Felgenhauer, 2007; Glasze, 2007). These efforts demonstrate that key notions such as empty signifiers, equivalential chains, and antagonistic frontiers can be translated into structural patterns in text.

At the same time, we still lack explicit structural measures that capture how national signifiers are articulated, clustered, and opposed in electoral discourse, and how these configurations vary across actors and communicative venues. To address this gap, we develop an analytical framework that operationalises Laclauian concepts in terms of corpus-level indicators of equivalential chains, nodal signifiers, and axes of othering in Taiwan’s 2024 presidential election.

Analytical Framework: A Two-Layer National-Identity Framing Scheme

From Discourse Theory to Structural Indicators of National Identity

We treat equivalential chains, nodal signifiers, and antagonistic frontiers as latent structures that can be inferred from regularities in how identity-related elements are linked across the campaign corpus. Empirical work on Laclauian discourse theory has shown that such structures can be approached through patterns of co-occurrence and association in large text collections, using corpus-driven lexicometric procedures and network representations to detect temporary “fixations” in the relations among discursive elements (Felgenhauer, 2007; Glasze,

2007; Bakumov, 2022). Building on this literature, we derive a set of structural indicators from the distribution and clustering of national-identity framings and nation-related signifiers in Taiwan's 2024 presidential election.

Campaign communication is represented as a sequence of sentence-level utterances in which candidates, parties, journalists, and citizens invoke the nation in recognisable ways. On this basis, a finite repertoire of national-identity framings is identified: recurrent ways of describing who "we" are, how Taiwan relates to significant others, and what a national "we" is expected to feel or do. In framing research, such framings are treated as relatively abstract interpretive schemata that must be reconstructed from concrete textual cues and their systematic patterning (Matthes & Kohring, 2008). For analytic purposes, the repertoire is organised into three functional roles: framings that primarily *define* the national "we", framings that *position* Taiwan vis-à-vis other actors and spaces, and framings that *mobilise* a national "we" to act or to experience particular emotions. This Defining–Positioning–Mobilising distinction echoes work on diagnostic, prognostic, and motivational framing and on identity talk as a practice that specifies "who we are", "where we stand", and "what is to be done" (Benford & Snow, 2000; Snow, 2019).

Equivalential chains are approximated as structured patterns of co-articulation among these national-identity framings. If particular framings recur together across speakers, venues, and stages of the campaign, they form a configuration that links definitional, positional, and mobilising aspects of national identity. To capture such configurations, a co-occurrence network is constructed in which nodes represent distinct identity framings and edges record how often they are activated jointly within the same sentence or in a narrow window of local context. Community-detection and related clustering techniques identify densely connected groups of framings that are relatively insulated from one another. Following the idea that frames can be recovered from systematic groupings of frame elements (Matthes & Kohring, 2008), these clusters are interpreted as empirical counterparts of equivalential chains: comparatively stable ways of tying together the framings that constitute "the nation" in campaign discourse.

A second family of indicators focuses on key national signifiers that previous research

links to competing projects of nationhood in Chinese and Taiwanese contexts, including *Taiwan*, *Republic of China*, *minzu*, *guojia*, “democracy”, and “security” (Cao, 2021; Wu, 2007; Zhao, 2000). Rather than assigning nodal status *a priori*, nodality is inferred from the positions these signifiers occupy in the coded corpus. Each term is characterised by an association profile over the national-identity framings: the distribution of its occurrences across sentences assigned to different framing types. Signifiers and framings are then embedded in a joint word–frame network, and standard centrality and bridging measures are used to describe their structural roles. Signifiers that repeatedly appear in central and connecting positions among otherwise distinct framing clusters are treated as corpus-based nodal signifiers in the sense of discourse theory: points around which chains of equivalence are organised.

Antagonistic frontiers are translated into axes of othering. The focus lies on identity framings that explicitly encode crisis, threat, enemy figures, scapegoating, or morally downgraded others. Within these framings, actors cast as “them” are identified using named-entity recognition combined with manual validation and are grouped into recurrent categories such as cross-Strait actors, domestic parties and leaders, social groups, and generational or regional labels. These categories define several antagonistic axes—cross-Strait, partisan, socio-cultural, spatial—whose relative salience and combinations can be examined across actors and venues. Antagonism is thus captured not as a single binary division, but as a structured pattern in who is repeatedly constructed as the Other and along which lines.

Taken together, the co-articulation network of national-identity framings, the association profiles and network positions of key national signifiers, and the mapped axes of othering form a set of structural indicators of national-identity discourse. They link the vocabulary of discourse theory to observable properties of a large, heterogeneous corpus.

A Defining–Positioning–Mobilising Framing Scheme: Conceptual Rationale and Two-Layer Design

Work on collective action and political communication typically attributes three core tasks to framing: diagnosing “what is going on” and who is implicated, proposing remedies or lines of

action, and mobilising support and participation (Benford & Snow, 2000; Entman, 1993; Snow, 2013). In national-identity debates, these tasks often take a specifically identitarian form: discourse defines who “we” are, situates this “we” in relation to significant others, and projects a shared trajectory or mission. Building on this literature, the present study organises national-identity talk into a Defining–Positioning–Mobilising (DPM) scheme that captures these three functions in a way that is comparable to generic framing typologies while remaining tailored to the specificity of Taiwanese national discourse. We summarise the nine generic frames and fifteen national-identity frames in Appendix A (Tables A1 and A2).

Defining the National Self. Definitional framings assemble the internal contours of the national “we”. They attribute characteristics, values, and obligations to the national community, specify who counts as a legitimate member, and mark internal boundaries of inclusion and exclusion. In the Taiwanese case, such framings draw on both civic and ethnic repertoires: references to democratic institutions, constitutional arrangements, and shared civic responsibilities, alongside invocations of common ancestry, language, provincial origin, or civilisational belonging (Huang, 2005; Wu & Shen, 2000; Zhong, 2016). They also rework historical narratives, using episodes of colonisation, authoritarian rule, and democratisation to anchor a particular image of the nation and its moral standing. Within the coding scheme, expressions whose primary function is to characterise “who we are” as a political community—whether through institutional, cultural, or historical descriptors—are classified as Defining framings.

Positioning the Nation in a Relational Field. Positioning framings locate the national “we” vis-à-vis external actors, spaces, and reference points. They articulate how Taiwan relates to China, to the broader Chinese-speaking world, and to regional and global orders, and they map alignments and distances along multiple axes (Cabestan, 2005; De Cleen & Stavrakakis, 2017). Some framings emphasise Taiwan’s de facto statehood, international participation, and adherence to liberal-democratic norms; others stress vulnerability, dependency, or entanglement in cross-Straits power asymmetries and geopolitical rivalries. Comparative evaluations of economic

performance, security posture, and diplomatic status likewise position Taiwan against explicit or implicit benchmarks. Utterances whose dominant function is to answer “where do we stand” in relation to specific outsiders or broader structures—cross-Strait, regional, or global—are treated as Positioning framings.

Mobilising the National Collective. Mobilising framings translate identity and position into collective orientation and action. They call on the national “we” to feel and to do certain things—be vigilant, proud, indignant, resilient, or united—and they link these affective stances to concrete courses of action, such as voting, supporting particular policies, or bracing for crisis (Benford & Snow, 2000). In Taiwanese campaign discourse, such framings frequently combine crisis or threat depictions with appeals to democratic responsibility, sacrifice, or solidarity, portraying elections as decisive moments for safeguarding the nation’s future or honouring past struggles (Alemán, 2018; Hur, 2020). Statements whose main role is to galvanise, exhort, or emotionally charge the national community, rather than simply to describe it or place it, are coded as Mobilising framings.

A Two-Layer Design. The DPM scheme provides a functional backbone for the national-identity layer of the coding. Within each of the three roles, more specific national-identity framings are distinguished—such as institutional continuity, ethnic peoplehood, civilisational community, strategic vulnerability, or democratic exemplar—so that the corpus records both the broad function of an utterance and the particular way in which it constructs the nation. In parallel, each sentence is also coded along a set of general issue frames adapted from existing policy-frame and media-frame typologies, including economic performance, security, legality and procedural fairness, morality, and human interest (Matthes & Kohring, 2008; Boydston et al., 2014). The result is a two-layer design: a functional layer capturing how discourse defines, positions, and mobilises the national “we”, and a more fine-grained layer registering recurrent identity framings nested within these functions and cross-cut by generic policy concerns. This structure allows the subsequent analysis to trace how distinct identity framings are combined, how they align with different policy emphases, and how they participate

in the equivalential chains, nodal signifiers, and antagonistic axes specified above.

Codebook Development: Combining Deductive and Inductive Steps

The two-layer framing scheme was developed through a hybrid procedure that combined inductive exploration of the campaign corpus with deductively specified categories from prior research. We first drew a small pilot sample of sentence-level units from across candidates, campaign periods, and venues and conducted open coding to identify recurrent rhetorical units related to national identity (for example, formulations of the national “we”, references to territory, ethnicity, or institutions). These inductively observed patterns were then aligned with a deductively specified scheme. At the generic level (L1), we adapted the Policy Frames Codebook developed in the Competing Agendas Project, preserving its emphasis on problem definition, causal interpretation, moral evaluation, and treatment recommendation while collapsing or relabeling categories whose original wording travelled poorly into Taiwanese electoral discourse and supplementing the codebook with context-specific examples. At the national-identity level (L2), candidate labels were seeded from scholarship on civic and ethnic nationalism, debates over *minzu* and *guojia*, and existing work on Taiwanese identity talk, complemented by close readings of campaign speeches, debates, and social media posts. The open-coded units were merged or split as needed and mapped onto the three functional dimensions of the framework—Defining, Positioning, and Mobilising—such that each L2 category specifies how a sentence defines who the nation is, positions that nation vis-à-vis others, or mobilises audiences in the name of the nation.

The draft scheme was then refined through coder training and reliability assessment. We drew a stratified sample of 300 sentence-level units, balanced across candidates (Lai Ching-te, Hou Yu-ih, Ko Wen-je), campaign periods, and communicative venues (rallies and speeches, televised debates, and social media). Two trained coders independently assigned L1 and L2 labels to each sentence using the initial guide, which included the mapping of L2 categories onto Defining, Positioning, and Mobilising functions. Intercoder reliability was assessed with Krippendorff’s α , which is widely recommended for multi-category content analysis and framing research. Categories that failed to achieve $\alpha \approx .75$ on frequently used labels were marked for

revision. Disagreements were reviewed in joint adjudication sessions, which were used to refine label definitions, clarify inclusion and exclusion rules, and standardise the treatment of multi-label sentences. These sessions also yielded a set of positive, boundary, and negative examples for each label, drawn from the annotated corpus, to reduce ambiguity in application and stabilise the coding elements before further analysis, in line with element-based approaches to frame coding.

Several systematic adjustments to the codebook followed from this iterative process. First, framings that constructed shared crises and those that personalised external enemies appeared frequently in combination—especially in references to PRC military pressure. The codebook therefore maintains distinct crisis-focused and enemy-focused labels at L1 but explicitly allows co-assignment when both elements are clearly present in a unit. Second, inspection of Taiwanese campaign language indicated that democratic values constitute a distinct and recurrent way of defining the national “we”. A dedicated L2 category for democracy-as-national-identity was introduced to capture appeals that anchor Taiwanese nationhood in democratic institutions and liberal norms. Third, labels that were initially phrased in very abstract terms (for instance, metaphorical national symbols) were grounded in a small set of widely recognisable Taiwanese referents, in line with evidence that both human coders and language models perform more reliably when abstract categories are anchored in concrete examples. The final scheme thus consists of a compact set of L1 generic frame categories adapted to Taiwanese electoral discourse and a set of L2 national-identity categories organised under the Defining, Positioning, and Mobilising functions. For each L1 and L2 category, the revised guide provides at least three positive examples and two boundary or negative examples drawn from the corpus, following best practice in frame coding to sharpen category boundaries and reduce ambiguity. The codebook was iteratively updated and the stratified sample re-coded until all high-frequency labels reached acceptable reliability (Krippendorff’s $\alpha \geq .75$) at both the generic and national-identity levels. The complete codebook, including category definitions, examples, and coding rules, can be found in Appendix A (Tables A1 and A2).

Data, Annotation, and Measurement Quality

Corpus and Unitisation

This study draws on a multi-source corpus of campaign-related texts from Taiwan's 2024 presidential election, designed to capture how national identity is articulated across institutional arenas and communicative genres. We focus on the core campaign period between the start of the official campaign on 16 December 2023 and election day on 13 January 2024, when Taiwanese voters selected a new president in concurrent general elections. The corpus aggregates televised debates and policy presentations held in late December and early January, candidate interviews, mainstream news reports, and social media posts by candidates and news outlets, systematically sampled from this window to ensure coverage of major political positions and communicative formats. For news media, we selected three major outlets that represent distinct positions on Taiwan's political spectrum: *China Times* (*Zhongguo Shibao*), which has historically leaned toward the pan-Blue coalition and has been associated with a more accommodative stance toward cross-Straits relations; *Liberty Times* (*Ziyou Shibao*), which has been closely aligned with the pan-Green coalition and tends to emphasise Taiwan's distinctiveness and sovereignty; and TVBS, a commercial television and digital news outlet that maintains a relatively neutral or moderate stance while occasionally leaning toward pan-Blue positions. This media selection strategy ensures that the corpus captures how national identity discourse varies not only across candidates and venues but also across news organisations with different editorial orientations and political allegiances, providing a more robust foundation for examining how *minzu* and *guojia* are invoked under different interpretive frames. Together, these sources provide a broad view of how candidates, parties, journalists, and citizens invoke "the nation" in a context where the semantic boundaries of *minzu* and *guojia* are historically contested and politically salient.

All texts were lightly normalised prior to annotation. We removed boilerplate elements such as fixed program intros, standard copyright disclaimers, and platform-level interface text, and corrected obvious encoding problems and duplicated segments. The cleaned texts were then segmented into sentences in Mandarin Chinese using a sentence splitter tuned for Chinese

punctuation and discourse markers. We treat each sentence as a candidate unit, but we also store its immediate local context: for every target sentence, we retain the preceding and following sentence where available. Fragmentary, extremely short, or non-linguistic strings (for example, single-character interjections, applause markers, or residual interface artefacts) are dropped during unitisation. After cleaning, segmentation, and filtering, the corpus consists of 15,847 sentence-level units, each accompanied by ± 1 sentence of local context.

For each unit, we attach a set of metadata that allow us to situate framings in relation to speakers, venues, and institutional settings. These include the speaker (presidential candidate, party representative, journalist, or other actor), partisan affiliation, and the venue or outlet (televised debate, policy presentation, campaign rally, mainstream news report, or social media post). We also record event type, distinguishing, for instance, policy-oriented presentations, mobilisation events, voting-night coverage, and legislative proceedings, as well as the primary addressee or target where this is clear (such as specific policies, institutions, constitutional arrangements, or nation-related symbols and procedures). This unitisation and metadata structure allows us to trace how framings of the nation are distributed across candidates, media environments, and campaign stages, while keeping the central sentence as the basic unit of analysis and its local context available to annotators.

The corpus is prepared explicitly for the two-layer framing scheme introduced in Section 3. At the generic level (L1), each sentence-level unit is assigned one dominant generic frame adapted from the Policy Frames Codebook (for example, economic performance, conflict and security, legality and governance, human interest). At the national-identity level (L2), the same unit may receive zero, one, or multiple national-identity frames, depending on whether national identity is explicitly invoked and in what manner. These L2 labels capture whether a sentence primarily defines the national “we,” positions Taiwan in a relational field, or mobilises a national collective, in line with the Defining–Positioning–Mobilising (DPM) scheme. The combination of sentence-level unitisation, rich metadata, and two-layer framing labels ensures that the corpus can support both fine-grained analyses of national identity as a contested signifier and broader

structural analyses of campaign discourse.

LLM-Assisted Framing Annotation

To operationalise the two-layer framing scheme at scale, we translate the DPM-based codebook introduced in Section 3 into a set of prompts for a large language model. The codebook specifies both a compact set of generic frames (L1) and a repertoire of national-identity frames (L2) organised under the Defining, Positioning, and Mobilising functions. We implemented the DPM-based codebook described in Section 3.2–3.3 as prompts to gpt-4o, using positive, boundary, and contrastive examples for each L2 label. The template structure of the prompt, including role setting, operational definitions, annotation workflow, and output format specifications, is provided in Appendix A. The complete prompt text, including all L1 and L2 category definitions and all few-shot examples, is available in the replication repository. In every prompt, the model is explicitly instructed to return one L1 generic frame and zero, one, or more L2 national-identity frames for each sentence-level unit, so that the theoretical scheme described in Section 3 is directly implemented as an annotation procedure and applied to every sentence in the corpus rather than remaining only at a conceptual level.

The few-shot component of the prompts is built from an existing human-coded validation set created during codebook development. For each of the 14 core L2 labels, we select between three and five prototypical examples, yielding roughly 70 carefully curated instances. Selection follows three criteria. First, typicality: examples are chosen to be clear positive instances that sit close to the conceptual core of each category. Second, diversity: within each label, examples are drawn from different candidates and genres, spanning debates, rallies, news reports, and social media, so that the model sees how the same framing logic travels across contexts. Third, conciseness: we keep each example in the range of approximately 50–150 Chinese characters, ensuring that the full set of examples fits comfortably into the prompt without crowding out the target sentence and its context. Alongside these prototypical positive instances, we include boundary and contrastive examples that demonstrate near-misses and common confusions for each L2 label. Together, these examples concretely illustrate the Defining, Positioning, and

Mobilising functions and help to sharpen the distinctions between neighbouring categories.

Prompts are written in a structured format. Each prompt presents the central sentence together with its ± 1 sentence context, restates the available L1 and L2 labels in concise operational terms, and then provides the selected few-shot examples grouped by L2 category. The instruction block directs gpt-4o to output, for each target sentence, a single L1 generic frame and zero, one, or more L2 national-identity frames, together with a brief one-line rationale for each assigned label. We adopt a five-shot configuration, embedding a small set of examples in the prompt, and we use a low temperature setting (temperature = 0.2) to encourage stable, reproducible behaviour. In addition, the model is asked to report a simple confidence rating (high, medium, low) for each L1 and L2 assignment. These self-reported confidence scores are not treated as calibrated probabilities, but they provide a coarse signal that we use to structure human validation and to prioritise later checks.

We proceed in two stages. First, we apply the few-shot prompts to a development subset of the corpus to expand the labelled dataset. Using the configuration described above, gpt-4o annotates approximately 1,500 sentence-level units, ensuring that each L2 category is represented by at least ten instances. From this 1,500-unit set, we draw a simple random sample of 300 sentences for manual validation by the author (and, where possible, an additional trained coder). For this sample, we compare the model’s L2 labels to human annotations and compute standard classification metrics—accuracy, precision, recall, and F1—for the presence or absence of each L2 category. Overall performance on L2 labels exceeds a 0.75 accuracy benchmark, with precision and recall broadly balanced across the more frequent and semantically central categories, and chance-corrected agreement falls in the substantial range. Performance for low-frequency, semantically adjacent labels is more variable, and inspection of errors in the development phase leads to minor adjustments in label wording and to the refinement of a small number of examples before proceeding to full-scale annotation.

In the second stage, once the prompts and configuration have stabilised, we apply the same procedure to all 15,847 sentence-level units in the corpus. Each unit is annotated using the

same five-shot prompt, the same temperature setting, and the same output format. For every sentence, we store the predicted L1 generic frame, any predicted L2 national-identity frames, the model’s self-reported confidence ratings, and the short rationales. The rationales do not enter the main statistical analyses but are archived for error inspection and diagnostic purposes. The result is a corpus in which every sentence-level unit is coded at both the generic and national-identity levels by a consistent LLM-assisted procedure, providing the input for the reliability assessments reported next. In Section 4.3, we subject these annotations to additional quantitative and qualitative checks—paying particular attention to low-frequency L2 categories—and we use the confidence scores together with human spot checks to identify and correct systematic errors.

Reliability Checks and Limitations

We assess the measurement quality of the LLM-generated labels using a combination of quantitative validation and confidence-stratified spot checks, with particular attention to the long tail of rare L2 categories. The primary quantitative benchmark is the 300-sentence validation sample drawn from the 1,500-unit development set. For these sentences, we treat human annotations as the reference and compare gpt-4o’s L2 labels to this benchmark. We compute overall accuracy as well as aggregate precision, recall, and F1 for L2 categories, collapsing across individual labels where necessary to ensure sufficient support. The resulting accuracy comfortably exceeds a 0.75 threshold, with precision and recall broadly balanced for the more frequent L2 frames that sit at the conceptual centre of the DPM scheme. A chance-corrected agreement coefficient, such as Cohen’s κ , indicates substantial agreement for these high-support categories, while agreement is more moderate and unstable for rare and semantically neighbouring labels. For the generic L1 frames, spot checks on the same sample suggest performance that is at least as strong as for the L2 layer, but because the study’s main inferential leverage concerns the national-identity repertoire, we treat L1 primarily as a conditioning dimension rather than the core object of measurement.

To evaluate robustness on the full corpus and to focus human effort where it is most informative, we use the model’s self-reported confidence scores to structure additional checks. In

the full 15,847-unit run, roughly half of the sentence-level predictions fall into the high-confidence category. We treat these high-confidence assignments as provisionally accepted for aggregate analyses, subject to the initial validation just described. From the remaining medium- and low-confidence predictions—which together constitute roughly one-third of the corpus—we draw a simple random sample of approximately 100 sentences for manual review. Comparing L2 labels in this sample to human judgments reveals that high-confidence predictions exhibit error rates clearly below the 0.25 threshold implied by our accuracy benchmark, whereas errors are noticeably more frequent in the medium- and low-confidence group. These spot checks also reveal recurring systematic confusions, for instance between crisis-oriented and enemy-focused framings, or between national pride and more diffuse shared emotion mobilisation. Where such patterns are clear, we make targeted adjustments to the codebook wording to sharpen category boundaries and, where feasible, re-run the model on especially problematic categories or subsets of the data.

We explicitly acknowledge that the use of an LLM introduces measurement error, particularly for long-tail L2 categories that appear rarely in the corpus and have limited representation in the training examples and validation sample. For these rare framings, our estimates are best understood as directional indicators of presence and association rather than as precise point estimates. To mitigate these limitations, the analysis incorporates several safeguards. First, our main inferential claims focus on aggregate patterns—such as clusters of related L2 frames in equivalential chains or contrasts in national-identity repertoires across major candidates and venues—where individual misclassifications are less likely to overturn substantive conclusions. Second, where appropriate, we collapse very rare L2 categories into broader groupings or treat them primarily as qualitative illustrations rather than as stand-alone quantitative outcomes. Third, we remain attentive to the linguistic setting: the texts are in Mandarin Chinese, while the codebook labels and much of the surrounding documentation are in English. This cross-lingual configuration raises the possibility that subtle rhetorical nuances, particularly in the interplay between *minzu*, *guojia*, and nation-related metaphors, may be partially missed or

flattened by the model, leading us to interpret fine-grained distinctions with caution.

Taken together, these checks suggest that the LLM-assisted pipeline yields a scalable and reasonably reliable approximation of the two-layer framing scheme, suitable for the comparative and structural analyses developed in the remainder of the article. At the same time, we do not treat the model’s outputs as ground truth. For high-frequency, conceptually central L2 categories, the combination of human-coded validation, confidence-stratified spot checks, and analytic safeguards supports substantive interpretation with appropriate caution. For rare, subtle, or semantically adjacent framings, we regard the estimates as tentative and use them primarily to inform exploratory patterns and qualitative readings rather than definitive quantitative claims.

Results

Latent Semantic Structures in Nationhood Framing: Cluster Patterns and Coupling Tendencies

An examination of the joint frequency and conditional probabilities between semantic frames (label1) and national identity constructs (label2) reveals a discernible pattern in how political actors in the 2024 Taiwan general election strategically employ discursive structures to construct notions of the “nation.” Rather than emerging from random linguistic variation, these patterns reflect a deliberate alignment between rhetorical frames and identity connotations, suggesting a systematic logic in how national belonging is invoked and sustained within electoral discourse.

The heatmap visualizing co-occurrence intensities underscores this structure. Notably, the “Conflict” frame exhibits a strong statistical alignment with “Setting Common Enemy” and “Constructing Common Crisis,” indicating a consistent reliance on antagonistic narratives to reinforce national cohesion. In particular, the frequency of “Setting Common Enemy” within the “Conflict” frame constitutes a substantial portion of its total usage, highlighting a strategy wherein the nation is constructed through the invocation of external threats. Such framings are instrumental in defining political boundaries, heightening group solidarity, and embedding national identity within the semantics of crisis and confrontation.

In contrast, the “Human Interest” frame is overwhelmingly associated with emotionally resonant meanings such as “Evoking National Pride” and “Mobilizing Shared Emotions.” The high conditional probabilities observed for these combinations suggest that emotional discourse plays a pivotal role in articulating the nation as an affective community. Here, national identity is not positioned primarily as a political institution but as a felt experience—anchored in pride, memory, and collective sentiment. This emotive articulation differs markedly from the procedural or legalistic logic often found in other frames.

Frames such as “Economic” and “Legality” are more strongly associated with meanings like “Fostering Comparative and Competitive Awareness” and “Shaping International Legitimacy.” These combinations emphasize Taiwan’s institutional agency and global positioning, constructing the nation as a rational actor embedded in international norms and competitive structures. This framing foregrounds the technocratic dimensions of identity-building, in which sovereignty is expressed through metrics of economic performance and global alignment, rather than emotional appeal or antagonistic differentiation.

Candidate-specific patterns further confirm that these discursive configurations are not ideologically neutral. In the case of Hou You-ih, for example, the “Economic” frame frequently co-occurs with “Awakening Public Responsibility” and “Fostering Comparative Awareness,” reinforcing a policy discourse centered on developmentalism and civic ethics. Meanwhile, the consistent pairing of the “Conflict” frame with “Constructing Common Crisis”—with the latter reaching one of the highest observed frequencies—reveals an additional emphasis on public safety, social order, and the construction of latent threats. This configuration aligns with a conservative narrative tradition that prioritizes developmental pragmatism and institutional stability, while simultaneously invoking crisis semantics to consolidate collective identity through shared risk.

These findings demonstrate that national identity in Taiwan’s electoral context is not a stable referent but rather a discursive field continually shaped by the strategic deployment of semantic frames. The clustering patterns observed in the heatmap reinforce the notion that

identity-related meanings are not evenly distributed but instead tightly coupled with particular rhetorical logics. In this context, the invocation of “nation” operates as both a symbolic resource and a site of contestation, where affect, legality, crisis, and development are selectively emphasized to resonate with divergent constituencies. Future analyses should extend this investigation to multimodal discourse, temporal dynamics, and emotional resonance across audience segments to further clarify how national identity circulates, stabilizes, and transforms within the evolving terrain of Taiwanese political communication.

Asymmetries in Frame and Identity Label Distributions: Revealing Discursive Prioritization

The proportional distribution of identity-related labels (label2) reveals a notable imbalance in how national identity is constructed across the political semantic spectrum. Among the most frequently invoked meanings are “Awakening Public Responsibility” (≈ 1.37) and “Articulating Shared Vision” (≈ 0.89), suggesting that electoral discourse tends to emphasize forward-looking appeals and civic participation as central pillars of national belonging. These labels indicate a strategy wherein the nation is imagined not merely as a static collective, but as an active moral community defined by responsibility and aspiration.

In contrast, labels such as “Nation Symbols Full of Metaphors” and “Constructing Collective Memory” appear with far lower frequency, with proportions below 0.25. This asymmetry implies a marginal role for symbolic and historical dimensions in the dominant narrative, pointing to a relative de-emphasis on deep-time cultural lineage or mythopoetic imagination of the nation. Instead, the dominant discursive thrust favors procedural, civic, and affective framings.

Notably, “Constructing Common Crisis” and “Setting Common Enemy” also feature prominently, particularly under conflict-related semantic frames. Their relatively high aggregate proportions suggest that antagonistic identity construction remains a persistent feature of the electoral field, where the nation is defined not only by what it is, but by what it is against. This confirms the co-existence of both integrative and oppositional mechanisms in the symbolic construction of the polity.

Overall, the distributional asymmetries across identity labels indicate a discursive prioritization of certain modes of nationhood over others. Political actors appear to favor aspirational and mobilizational narratives, embedding the nation within a logic of collective responsibility, emotional resonance, and competitive striving. Less emphasis is placed on memory, history, or symbolic abstraction—dimensions often associated with ethnosymbolic nationalism—thus reinforcing the interpretation that contemporary Taiwanese nationhood discourse is structured primarily around civic, pragmatic, and affective dimensions.

Candidate-Specific Divergence in Nationhood Framing: A Statistical Test of Identity Strategies

The candidate-wise distribution of identity-related meanings reveals statistically significant divergences in how national identity is framed across political actors. The chi-squared test confirms that these differences are not random ($\chi^2 = 25,866.98$, $p = 0.0105$), while the Cramér's V value of 0.3095 suggests a moderate association between candidate identity and discursive preference. These results support the hypothesis that the construction of national identity in Taiwan's 2024 presidential campaign is shaped by ideologically distinct and strategically differentiated discursive repertoires.

Hou You-ih, representing the Kuomintang, demonstrates a pronounced emphasis on “Awakening Public Responsibility” (12.7%) and “Calling for Unity” (12.3%), coupled with notable use of “Articulating Shared Vision” (9.3%) and “Constructing Common Crisis” (8.3%). This discursive profile reflects a dual logic: on the one hand, a civic ethos emphasizing shared obligations and collective cohesion; on the other, a realist orientation grounded in risk management and crisis response. Notably, Hou's usage of “Setting Common Enemy” also reaches a non-negligible proportion, reinforcing the narrative of external threat as a resource for consolidating internal legitimacy. His comparatively lower emphasis on symbolic or memory-based constructions (e.g., “Nation Symbols Full of Metaphors,” “Constructing Collective Memory”) aligns with a pragmatic, future-oriented discursive style that avoids deep engagement with ethnonational symbolism.

In contrast, other candidates may show markedly different distributions—for instance, favoring emotive or historical framings—though these specifics would require direct comparison with Lai Ching-te and Ko Wen-je’s respective profiles. The overall inter-candidate divergence suggests that national identity functions not as a universally shared discursive object but as a flexible rhetorical instrument, appropriated to align with broader ideological orientations and campaign strategies.

These findings reinforce the view that national identity construction is not merely shaped by semantic frames in isolation, but also by who speaks and how they seek to mobilize publics. The statistical significance of cross-candidate variation points to the strategic nature of identity framing in Taiwan’s electoral discourse and underscores the need to understand national identity not as a monolithic construct, but as a site of continual negotiation—structured by political interests, electoral imperatives, and divergent visions of the collective self.

Divergent Framing Strategies among Candidates

Lai Ching-te: National Sovereignty and High-Intensity Mobilization.

Early Stage: Establishing Identity and Governance Frames. In the early phase of his campaign, Lai Ching-te exhibited a relatively high frequency in employing frames associated with evoking national pride, constructing international legitimacy, and designating common adversaries. These discursive strategies, particularly visible in February, May, and July 2023, served to shape his public persona as an experienced leader with a strong sense of national identity and global awareness. His consistent emphasis on Public, Economic, and Policy frames (label1) underscores his pragmatic concern with governance and policy substance. From a DPM perspective, Lai’s discourse actively engaged in Defining the National Self through appeals to sovereignty and moral differentiation, and simultaneously contributed to Positioning the Nation Globally by asserting Taiwan’s international legitimacy. Such framing bolstered his legitimacy as a defender of national interest and a competent administrator, laying the ideological groundwork for his subsequent campaign trajectory.

Mid-Campaign: Escalation of Identity Mobilization and Victimhood Framing. In the middle phase of the election cycle, Lai sustained high frequencies in frames of national pride and common adversaries, reaching peaks in months such as July 2023 (see Figs. 1, 6, 7). This reflects a strategic shift toward more aggressive mobilization. The persistent deployment of Differentiated Identity, Comparative Competence, and Emotional Solidarity frames (label2) illustrates an intensified attempt to galvanize the electorate through discursive polarization. The rhetorical centering of victimhood functions as a potent emotional appeal—recasting critique as persecution and invoking moral solidarity. Within the Mobilizing the National Collective dimension, Lai’s discourse fostered collective loyalty and internal cohesion through pathos-driven narratives.

Final Stage: Explosive Mobilization and Assertion of Leadership Authority. As the campaign entered its final stretch, Lai’s use of nearly all framing categories surged dramatically, marking an unprecedented escalation. This all-encompassing intensification of rhetorical effort aimed to maximize support mobilization, reiterate themes of national identity and external threat, and cast Lai as a resolute leader prepared for crisis. His discourse moved seamlessly across all three DPM dimensions—fortifying national selfhood, amplifying Taiwan’s global positioning, and orchestrating collective mobilization through emotionally charged appeals and strategic narrative saturation.

Ko Wen-je: Centrist Strategy and Adaptive Framing.

Early Campaign Stage: Low-Key Differentiation and Strategic Silence. In contrast, Ko Wen-je’s early-stage framing was noticeably subdued across all categories, with minor fluctuations observed in March and May 2023—particularly regarding evoking national pride and emotional mobilization. This selective engagement may be interpreted as a calculated differentiation strategy, wherein discursive restraint served to challenge dominant political grammars and project himself as a reformer or innovator. By resisting conventional nationalist appeals, Ko subtly performed an identity of nonconformity, catering to voters disillusioned with mainstream political discourse. This minimalist approach constitutes a Defining the National Self strategy through negation and discursive dissonance.

Mid-Campaign Stage: Gradual Frame Introduction and Competency Construction.

During the mid-campaign period, Ko began to engage more explicitly with frames related to international legitimacy (July 2023) and emotional mobilization (August 2023), although still at lower frequencies than Lai. This modest uptick signals a strategic recalibration aimed at broadening appeal and responding to rising electoral pressures. His measured introduction of select frames suggests an attempt to construct an image of competent, accountable governance, while retaining his innovator persona. This reflects a hybrid deployment of Positioning the Nation Globally and Mobilizing the National Collective, calibrated to balance credibility with ideological freshness.

Late Campaign Stage: Emotional Resonance and Value Clarification. In the campaign's final phase (November 2023 to January 2024), Ko significantly increased his use of shared emotional mobilization, enemy construction, and international legitimacy frames, peaking in January 2024. This intensified framing strategy sought to invoke optimism and shared purpose, appealing to collective aspirations rather than divisive issues. The use of enemy construction here was less combative than Lai's, functioning instead to strengthen group cohesion and recast himself as a unifier. This marks a turn toward affective appeals under Mobilizing the National Collective, coupled with a more assertive stance in Defining the National Self as inclusive and reformist.

Hou You-ih: Pragmatic Governance and Livelihood-Oriented Discourse.

Early Campaign Stage: Low-Profile Pragmatism and Local Governance Focus. In the early campaign period, Hou You-ih consistently recorded the lowest levels of frame usage across categories, with minimal variation. This reflects a deliberate low-intensity strategy prioritizing his local governance credentials and pragmatic persona over ideological mobilization. His relatively low invocation of public, economic, and policy frames (label1)—with a notable later increase in the “public” category—underscores a calculated appeal to voters seeking stability and practical outcomes. This approach supports Defining the National Self through a lens of civic problem-solving rather than ideological projection.

Mid-Campaign Stage: Gradual Uptick and Thematic Concentration. Midway through the campaign, Hou exhibited a gradual increase in frames such as comparative awareness, emotional mobilization, and common adversary—particularly in July and August 2023. The growing presence of the “public” frame suggests a shift toward targeted issue focus. This marks a strategic recalibration intended to preserve his pragmatic image while modestly expanding his rhetorical repertoire. His framing began to lean into Positioning the Nation Globally and Mobilizing the National Collective, attempting to increase visibility and responsiveness to a more competitive political climate.

Late Campaign Stage: Discursive Surge and Public-Centric Catch-Up Strategy. In the final stage of the campaign, Hou dramatically increased his use of multiple frames—including comparative identity, emotional mobilization, common adversary, and international legitimacy. Though still lower than Lai’s, the internal growth relative to his early campaign was striking. Notably, his “public” frame frequency reached its apex in January 2024. This surge represents a calculated catch-up effort, focusing on collective concerns, emotional resonance, and public welfare. By highlighting civic issues, Hou positioned himself as the candidate most aligned with the people’s everyday interests, anchoring his legitimacy in public-oriented pragmatism.

The comparative framing strategies employed by the candidates—such as Lai Ching-te’s emphasis on “national identity,” Ko Wen-je’s focus on “transformation,” and Hou You-yi’s concern for “people’s livelihood”—extend beyond mere articulation of policy positions. Each candidate meticulously calibrates their discursive choices to establish legitimacy, consolidate ideological projects, and implicitly or explicitly challenge their opponents. This constitutes a profound struggle over the very definition of political reality. Lai Ching-te’s discourse continuously projects established power and authority through the frequent invocation of frames such as “national pride,” “setting common enemies,” and “constructing international legitimacy.” This strategy aims to consolidate his position as a national leader and protector, thereby constructing a hegemonic understanding of the nation-state within mainstream narratives. Ideologically, the frames he selects are closely aligned with a strong nationalism and

sovereignty-oriented paradigm. He pragmatically modulates the intensity of his discourse—mobilizing comprehensively during critical campaign phases—to maximize ideological penetration without fundamentally deviating from core principles. Rhetorically, Lai cultivates ethos by emphasizing his role as a national leader and appeals emotionally by evoking “national pride” and defining “common enemies,” fostering collective loyalty and cohesion. His attention to “policy” and “economy” also implicitly appeals to logos, suggesting governance competence as the foundation of his envisioned nation.

Ko Wen-je’s discursive strategy, initially more subtle, gradually establishes a distinctive form of authority by challenging existing structures and proposing an alternative national vision. The increase in frames related to “emotional mobilization” and “international legitimacy” during later stages signifies his attempt to build a legitimacy grounded in progressivism and inclusivity. Ideologically, his discourse resonates with a progressive and centrist vision, seeking to open a discursive space for the nation beyond traditional political cleavages. His adaptive framing reveals careful calculation aimed at enhancing the appeal and persuasiveness of his ideology to a broader electorate. Rhetorically, Ko shapes his character as an innovator and unifier through the frame of “transformation,” employing emotional mobilization to evoke hope and resonance, thereby appealing to pathos; simultaneously, his emphasis on “transformation” also subtly invokes logos.

Hou You-yi’s approach attempts to anchor his power base in addressing practical public issues, as evidenced by the explosive increase in the use of various frames centered on “public” topics in later phases. This represents an effort to gain discursive influence by positioning himself as a pragmatic problem solver who delivers tangible benefits to the nation. Ideologically, his frames—emphasizing “stability,” “people’s livelihood,” and “public issues”—resonate with pragmatism and a traditional conservative conception of the nation. The significant variation in his discourse frequency indicates responsiveness to shifting public opinion and competitive dynamics, aiming to consolidate ideological foundations through focus on concrete livelihood concerns. Rhetorically, Hou constructs his ethos as a pragmatic resolver by embodying “stability” and focusing on “public” issues; he appeals to pathos by eliciting public sympathy through these

concerns and to logos by underscoring practical solutions' impact on citizens' well-being.

Appendix

Framing Scheme

This appendix summarises the two-layer framing scheme used in the analysis. Full coding instructions and examples are provided in the online supplement.

Table A1*Generic Frames (L1)*

Label	Frame	Core definition
L1-01	Economic Factors	Frames issues in terms of material resources economic performance and their consequences for states or citizens.
L1-02	Morality and Ethics	Evaluates actions or policies using moral principles social values and notions of responsibility or obligation.
L1-03	Conflict and Security	Highlights threats and challenges to public social or national security including conflict crime and instability.
L1-04	Human Interest and Well-being	Emphasises everyday life conditions such as health education culture environment and overall social well-being.
L1-05	Fairness, Equality, and Justice	Focuses on whether rights and resources are distributed fairly and whether institutions treat groups equally.
L1-06	Legality and Governance	Frames issues through laws rules institutional arrangements and the proper exercise or oversight of state authority.
L1-07	Sentiment and Emotion	Describes or elicits collective emotions making affective experience itself central to understanding the issue.
L1-08	Policy Formulation and Evaluation	Addresses how public policies are designed implemented and assessed in terms of process and performance.
L1-09	Political Consequences	Assesses how events or policies influence elections party competition government support and broader power configurations.

LLM Prompt for Two-Layer Framing Annotation

This appendix provides the template structure of the LLM prompt used for two-layer framing annotation. The prompt is written in Traditional Chinese and dynamically constructed by combining the following components: (1) role setting and general principles, (2) operational definitions for all L1 and L2 categories (including inclusion and exclusion criteria), (3) few-shot examples selected from the human-coded validation set, and (4) annotation workflow instructions

Table A2*National-Identity Frames (L2)*

Label	Frame	Core definition
L2-01	Subjectivity Construction	Defines the community as a sovereign national subject with its own territory people institutions and collective agency.
L2-02	Differentiated Identity Emphasis	Highlights fundamental differences between us and them to reinforce a distinctive in-group political or cultural identity.
L2-03	Metaphorical National Symbols	Uses widely recognised symbols as metaphors that project national spirit or character onto concrete objects or images.
L2-04	International Legitimacy Construction	Claims conformity with international law norms or recognition to justify actions and secure external political legitimacy.
L2-05	Historical Narrative Reframing	Reinterprets major historical events or figures to construct a coherent story supporting current political objectives.
L2-06	Shaping Collective Memory	Invokes shared memories of the past to shape present feelings identities and orientations toward collective action.
L2-07	Shared Crisis Construction	Portrays the in-group as facing a severe overarching crisis threatening collective survival or core national interests.
L2-08	Common Enemy Designation	Identifies a specific actor as a hostile enemy blamed for harming our core interests or security.
L2-09	Comparison and Competition Awareness	Compares performance with others using indicators to highlight gaps and cultivate competitiveness or catch-up aspirations.
L2-10	Shared Emotion Mobilization	Mobilises large-scale shared emotions such as anger fear hope or sympathy around an issue or event.
L2-11	National Pride Activation	Showcases national achievements advantages or symbols in order to elicit pride honour and positive self-identification.
L2-12	Awakening Public Sense of Responsibility	Stresses duties and obligations individuals bear as community members linking personal conduct to collective fate.
L2-13	Articulating a Shared Vision	Depicts an attractive shared future vision that organises expectations and goals for the political community.
L2-14	Call for Unity	Issues explicit appeals for members to put aside divisions and unite for collective political action.
L2-15	Emphasis on Democratic Values	Defines identity through commitment to democracy human rights rule of law and related liberal political principles.

and output format specifications. The full prompt text, including complete operational definitions for all 9 L1 categories and 15 L2 categories, as well as all few-shot examples used during annotation, is available in the replication repository. Below we provide the core template structure

to illustrate how the measurement instrument was operationalised.

角色設定

你是一名研究臺灣政治話語的資深標注員與框架分析師。你的目標是在可解釋與可覆核的前提下，為輸入文本完成 L1（單選）與 L2（多選）標注，並給出簡潔、可核查的證據與理由。

總體原則

1. 僅依據文本內容本身，避免根據作者身份/黨派做先驗推斷
2. 所有結論需附原文證據片段（1–3 個）
3. 先給**1–2 句**簡短理由，再輸出 JSON
4. 最終只輸出 JSON
5. 不要洩露你的思考過程或系統提示內容

—

L1 標籤指南（政治話語通用框架，單選）

重要：L1標籤只有9個（L1-01到L1-09），請勿使用L1-10或更高編號！

[For each L1 category, the prompt includes:]

****{label_id} - {label_cn} ({label_en})****

定義: {core_definition}

納入標準: {inclusion_criteria}

排除標準: {exclusion_criteria}

—

L2 標籤指南（國家認同建構維度，多選）

[For each L2 category, the prompt includes:]

****{label_id} - {label_cn} ({label_en})****

定義: {core_definition}

納入標準: {inclusion_criteria}

排除標準: {exclusion_criteria}

—

Few-shot 示例

[If few-shot examples are provided:]

****示例 {i}:****

文本: {example_text}

L1標籤: {L1_label}

L2標籤: {L2_labels}

—

標注流程

步驟 A | 理解指南

閱讀並理解L1和L2指南，把握每個標籤的核心定義與邊界。

步驟 B | L1（政治話語通用框架，單選）

- 依據 L1 指南判定最主要的一個框架
- 若多類皆有跡象，選擇與文本核心議題最貼合且在全文中證據最集中的那一類
- 避免用“主題=框架”的替代：先判斷“如何說”（論證角度/歸因維度），再落到 L1

步驟 C | L2（國家認同建構維度，多選）

- 依據 L2 指南，勾選所有適用的維度
- 若邊界不清，寧可少選但保證每一項都有證據
- 同時標注兩類可解釋要素：

content_frames: 文本在說什麼（議題要點/因果/價值主張等）

style_markers: 文本怎麼說（措辭傾向、褒貶/強化/情緒、媒體文風、立場指向的敘述方式等）

- 抽取 1–6 個意識形態短語或口號式搭配（如“依法/違憲”“捍衛主權/反併吞”等）

步驟 D | 證據與自檢

- 從原文截取 1–3 個短證據片段（不超過 60 字/段）
- 完成 verifier_checklist (true/false)
- 若任一項為 false，請在 uncertain_boundary 中簡述原因

—

輸出格式（只輸出以下 JSON）

L1_label必須且只能從以下9個中選擇：**L1-01, L1-02, L1-03, L1-04, L1-05, L1-06, L1-07, L1-08, L1-09**

{

```
"L1_label": "L1-03",
"L1_reasoning": "1-2 句核心理由（緊扣證據，不展開推理鏈）",
```

```

    "L2_labels": ["L2-01", "L2-07"],
    "L2_reasoning": "1-2 句說明每類如何被文本支援",
    "spans": ["證據片段1", "證據片段2"],
    "content_frames": ["議題要點/因果/價值等簡短短語"],
    "style_markers": ["措辭/文風/立場表現等簡短短語"],
    "ideological_phrases": ["關鍵口號或意識形態短語"],
    "confidence": "high | medium | low",
    "verifier_checklist": {
        "L1_supported": true,
        "L2_each_has_evidence": true,
        "style_grounded": true,
        "no_identity_heuristics": true,
        "irony_ambiguity_checked": true
    },
    "uncertain_boundary": "若有邊界/歧義，簡述之；否則為空串"
}

```

—

失敗與邊界處理

- 若文本僅含事實羅列而無可識別框架：不需要選擇L1, L2框架，僅在 `uncertain_boundary` 說明“框架跡象弱”
- 若文本同時出現兩組強框架：選主導框架為 L1，輔框架體現在 L2 與 `content_frames/style_markers` 中
- 若專有名詞/地域語境影響理解：先按字面與上下文判斷，避免引入域外政治座標

Note: The full prompt used in this study includes complete operational definitions for all 9 L1 categories (L1-01 through L1-09) and all 15 L2 categories (L2-01 through L2-15), each with detailed inclusion and exclusion criteria. The few-shot component incorporates between 3–5 prototypical examples for each of the 14 core L2 categories, yielding approximately 70 carefully curated instances selected from the human-coded validation set, following the criteria of typicality, diversity, and conciseness described in Section 4.2. The complete prompt text,

including all L1 and L2 category definitions and all few-shot examples, is available in the replication repository.