

The Dehumanization of Politics: The Case of Carlos Manzo

Sebastián Duran

Mexico is one of the most important countries in Latin America, as it is the second-largest economy and the second-most populous nation in the region, only behind Brazil. However, despite its enormous economic potential, the country has faced a profound security crisis over the past twenty years, resulting in thousands of people being killed and disappeared, triggering forced displacement, and generating millions of victims.

As a result, the Mexican state currently disputes territorial control with more than forty organized crime groups. According to *El Universal* (2025), more than 40 % of the national territory is under the control of one or more groups linked to drug trafficking cartels, with the most dominant being the Jalisco New Generation Cartel (CJNG) and the Sinaloa Cartel.

In 2018, the Morena party, under the leadership of Andrés Manuel López Obrador, reached the Presidency of the Republic after two previous attempts. With a new left-wing landscape at the head of the federal government and several state administrations, the approach to addressing the security crisis shifted, moving away from the so-called “war on drugs” and adopting a discourse centered on prioritizing life and the human rights of the most vulnerable populations.

However, this new public policy failed to achieve the expected results; on the contrary, it produced adverse effects that intensified violence in the country. In general terms, there was a sustained increase in intentional homicides, femicides, and other crimes compared to previous administrations (Osorio et al., 2025). According to Astorga (2016), cited in Osorio (2025), the main shortcomings stemmed from the absence of an articulated public security policy and the limited presence of the state across large territories. This situation facilitated disputes among drug cartels over control of strategic areas for trafficking drugs, weapons, contraband, and other illicit activities. Furthermore, these organizations consolidated mechanisms of social control through practices such as extortion, known as *cobro de piso*, and kidnapping, allowing them to diversify their sources of income and strengthen their dominance over communities.

Despite other crises, such as the corruption cases involving Segalmex and the so-called fiscal fuel theft (*huachicol fiscal*), Morena managed to secure re-election for the 2024–2030 term under the presidency of Claudia Sheinbaum Pardo. Following the line of her predecessor, Sheinbaum maintained the same

security policies, although with even more unfavorable results. During her first year in office, not only did the previously mentioned crimes increase, but political violence targeting the opposition also intensified.

Just six days after taking office, the mayor of Chilpancingo, in the state of Guerrero, was assassinated and subsequently beheaded as a message directed at the local political class that threatened criminal interests in the region. Alejandro Arcos Catalán had been nominated by the Institutional Revolutionary Party (PRI) under a platform aimed at “restoring” security to the local community.

The situation escalated in 2025, affecting ten mayors belonging to the Morena party, the Labor Party (PT), and the Party of the Democratic Revolution (PRD). However, the assassination of Carlos Manzo, independent mayor of Uruapan, Michoacán, triggered a profound crisis of legitimacy for the federal government. This event provoked not only an immediate response from the opposition, but also a civic reaction that translated into social mobilizations demanding security guarantees, as well as improvements in areas such as healthcare, justice, and freedom of expression.

Us vs. Them Theory

This theory describes a psychological and political communication strategy through which a political or social group constructs and reinforces its identity in opposition to another group, which is perceived as a threat or as the source of all problems that must be eliminated.

In general, two concepts underpin this theory, *Realistic Conflict Theory* and *discursive framing*, both of which are used in the development of this article. Realistic Conflict Theory, developed by Muzafer Sherif (1966), argues that conflicts between groups do not primarily arise from individual prejudice, but from real or perceived competition over limited resources, such as power, money, legitimacy, or security. When two groups believe they are competing for something valuable, they tend to develop hostility, negative stereotypes, and an antagonistic narrative of the “us versus them” type, even if no prior rivalry existed.

Sherif demonstrated this process in his famous Robbers Cave experiment, where two groups of children, when competing for

a discourse aimed at minimizing the tone and relevance of the assassination of the mayor of Uruapan.

To identify the patterns and thematic groupings present in this digital conversation, the **Hierarchical Clustering** technique was employed, a method that organizes data into a hierarchical tree-like structure (dendrogram). This approach allows messages to be grouped according to their similarity, either through an agglomerative process (bottom-up), in which clusters are progressively merged, or a divisive process (top-down), in which a general set is subdivided into more specific groups.

Thus, as shown in Figure 2, the main theme centered on redirecting the discussion toward a “corrupt right wing in Mexico”. From this, three major subgroups emerge: those who defend the government based on the work carried out by the current president and former president López Obrador; a second group focused on attacking the opposition through personal labels such as “corrupt”, “indoctrination”, “blackmail”, and “coward”; and finally, a smaller group that defends the mayor’s work, portraying him as a “hero”.

The discussion on X, as illustrated in the following network graph, reveals a highly dominant community when analyzed through retweets. In general, the accounts in purple were grouped together due to representing a similar and more influential narrative; the account “mysteryprecious” functions as a key articulator, whose tweet was the most widely disseminated, effectively setting the agenda. The accounts within this community range from defending the government to attacking the opposition.

In general, agenda-setting is driven not directly by government accounts, but by militant or partisan actors. For example, the main account has no formal connection to the government; however, official accounts such as “LuisaAlcalde” have actively contributed to sharing and amplifying the discussion.

The construction of this narrative responds to the main issues on Mexico’s public agenda during the period analyzed and can be understood through the lens of Realistic Conflict Theory, developed by Muzafer Sherif. In a context of a legitimacy crisis, the assassination of the mayor of Uruapan becomes a contested symbolic resource, in which different actors compete to impose interpretations that favor their political position.

From this perspective, the ruling party and its support networks on X articulated a discursive framing aimed at minimizing the relevance of the crime and shifting attention toward the notion of a “corrupt right wing”, thereby constructing an intergroup antagonism characteristic of realistic conflict. This suggests that the struggle over agenda-setting does not occur exclusively through institutional channels, but also through

peripheral actors who reinforce cohesion within the ruling group while deepening polarization against the political opponent.

Figure 3. Network graph



Source: Author’s own elaboration.

Methodological Precision

The identification and classification of messages within the discourse categorization were supported by the use of a Large Language Model (LLM) as a tool for automated language analysis. The LLM was employed to detect semantic, lexical, and narrative patterns associated with the construction of the “us vs. them” antagonism, as well as to identify indicators of affective polarization, moral delegitimization, dehumanization, and incitement to harm. Its primary function was to assist in content coding by recognizing complex discursive structures beyond isolated words, such as metaphors, generalizations, and implicit dichotomous frames.

Methodologically, the LLM operated as a classifier assisted by theoretical rules, in which the categories (levels 0–5) were defined based on theoretical contributions derived from the “us vs. them” framework. Each message was analyzed according to its contextual content, moral load, and affective orientation, allowing for the assignment of an ordinal level of hate based on the established criteria. The use of the LLM did not replace the theoretical framework or analytical judgment; rather, it enhanced the consistency, replicability, and scalability of the analysis, reducing subjectivity in classification and facilitating systematic comparison across messages.

In order to systematically analyze the intensity of hostile discourse in political messages related to the Uruapan case, this study proposes a categorization on a scale from 0 to 5. This enables the operationalization of the escalation of antagonism through progressive categories ranging from communicative neutrality (level 0) to the explicit incitement or legitimation of harm against an out-group (level 5). The construction of this scale is grounded in theoretical contributions on affective

polarization, framing, moral delegitimization, dehumanization, and hate speech, making it possible to identify not only the presence of negative language but also its degree of symbolic radicalization. In this way, a replicable analytical tool is provided to measure how discourses construct boundaries between “us” and “them”, and how these narratives may intensify toward extreme forms of exclusion and symbolic violence.

The development of this categorization draws on well-established literature on polarization, discursive construction of antagonism, and the escalation of hate in political language. In particular, it incorporates the contributions of Iyengar, Sood, and Lelkes (2012) on affective polarization to identify initial levels of emotional hostility; the theoretical developments of Entman (1993) and Lakoff (2004) on framing and dichotomous narrative structures to understand the discursive production of the “us versus them” divide; as well as the approaches of Mudde (2004) and Müller (2016) regarding the moral delegitimization of adversaries in populist discourse.

For the more extreme levels, the framework integrates Haslam’s (2006) analysis of dehumanization as a symbolic mechanism of exclusion and Bandura’s (1999) concept of moral disengagement, which explains how harm toward the out-group becomes normalized. Finally, the methodological criteria for distinguishing between offensive language and hate speech are complemented by empirical studies on hate speech detection in social media, such as Davidson et al. (2017) and Waseem and Hovy (2016).

Taken together, these contributions enabled the construction of a gradual, theoretically grounded, and methodologically replicable scale to measure the intensity of discursive hate in political messages.

Level 0. Neutral The message does not contain hostility, antagonism, or negative evaluations toward a group. It is descriptive or informational. Example: Factual language, absence of negative adjectives, no construction of enemies.

Level 1. Moderate rational criticism: There is disagreement or political criticism, but within democratic and argumentative parameters. There is no insult or strong emotional rejection. Example: Questioning of decisions or actions, moderate tone, institutional critique.

Level 2. Affective hostility: Explicit emotional rejection toward the adversary or out-group appears. Emotional polarization predominates. Example: Expressions of anger, frustration, or contempt (“they disgust me”, “they always ruin everything”).

Level 3. Moral delegitimization: The adversary is portrayed as corrupt, traitorous, or illegitimate. A moral division of “us

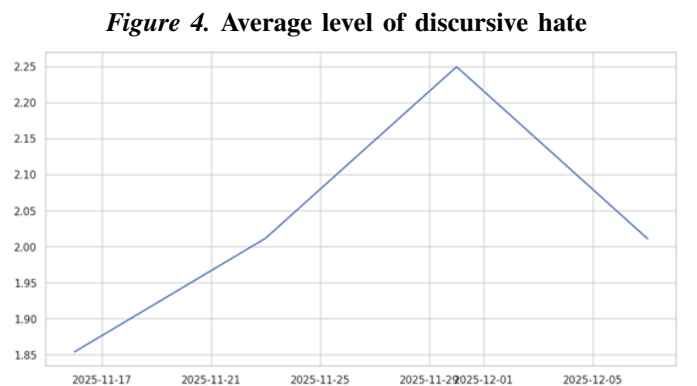
(good) vs. them (bad)” is activated. Example: Generalizations (“they are all corrupt”), denial of legitimacy, absolute moral labels.

Level 4. Dehumanization: The out-group is represented as inferior, dangerous, or non-human through metaphors of animalization or contamination. Example: Use of terms such as “rats”, “plague”, “scum”, “parasites”, symbolic extermination language.

Level 5. Incitement or legitimization of harm: The message promotes, justifies, or normalizes violence or the elimination of the targeted group. This is the highest level of discursive hate. Example: Explicit calls (“they must be eliminated”), justification of harm (“they deserve it”), direct threats.

Temporal Escalation of Discursive Hate and Its Concentration in High-Amplification Actors

The results from more than 1,000 tweets show a range with a minimum of 1.85 points and a maximum of 2.25 points, indicating a gradual and sustained increase over time. This pattern suggests a process of discursive escalation, that is, a growing presence of messages transitioning from lower levels of moderate criticism toward more emotionally charged forms.



Source: Author’s own elaboration.

Under the logic of Table 1, this shift implies that the discourse increasingly incorporates elements characteristic of affective hostility (level 2) and, in some cases, approaches moral delegitimization of the adversary (level 3).

The highest point (2.25) occurred on November 30. This peak can be interpreted as a moment of heightened emotional polarization, in which affective hostility predominates, in some cases approaching moral delegitimization. On November 25, Morena party member and senator Gerardo Fernández Noroña made negative statements against Grecia Quiroz, widow of Carlos Manzo. Additionally, November 30 marked one month since the assassination; both events triggered a surge

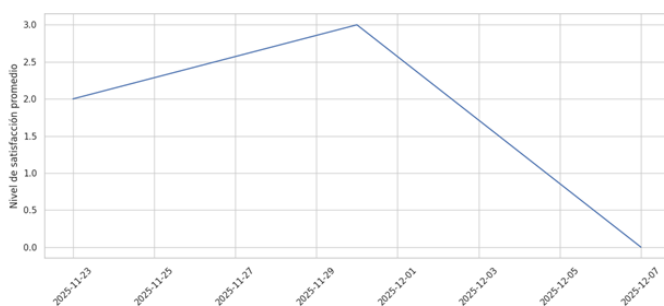
in emotionally charged comments, particularly directed at the victim’s family and, more broadly, at the opposition.

In summary, the figure shows a progressive increase in the average level of discursive hate throughout November, reaching its highest point at the end of the month, followed by a moderate decline in December. The discourse remains primarily at intermediate levels (1–2), characterized by criticism and affective hostility, without predominantly reaching extreme levels of dehumanization or direct incitement to violence. It gradually subsides in December 2025, as the public agenda shifts to other issues.

However, the main accounts functioning as key agenda-setting actors exhibit higher levels within the categorization. In the initial period (November 23), the average stands at around 2.0, corresponding to a level of moderate affective hostility. This indicates that the discourse amplified by these accounts contains explicit emotional rejection, but does not yet predominantly reach extreme registers such as dehumanization or incitement to violence.

Subsequently, toward the end of November (November 30), the average rises to approximately 3.0, representing a significant shift toward levels of moral delegitimization of the adversary. At this level, messages tend to portray the opposition as illegitimate, corrupt, or morally inferior, intensifying the dichotomous “us versus them” logic. This increase is particularly significant because it originates from highly retweeted accounts, suggesting that the most amplified discourses during this period were also the most polarizing.

Figure 5. Average level of discursive hate among the top 5 most retweeted accounts



Source: Author’s own elaboration.

The observed peak (3.0) indicates that, at that point in time, the accounts with the greatest dissemination capacity were producing messages that went beyond conventional political criticism and operated within a moralizing and antagonistic framework.

In the literature on framing and populism (Mudde, 2004; Müller, 2016), this type of discourse typically emerges in contexts where the goal is to consolidate political identities through the construction of the enemy as a moral threat, rather than as a mere opponent.

This result is significant because it demonstrates that discursive hate is not evenly distributed across the network; rather, it can concentrate among actors with high dissemination capacity. Therefore, analyzing highly retweeted accounts provides a better understanding of the mechanisms through which antagonism and polarization are amplified. Moreover, the abrupt pattern observed suggests that hate discourse may be episodic, linked to specific events, rather than a sustained constant.

Final Thoughts

The case of the assassination of Carlos Manzo in Uruapan demonstrates not only the depth of Mexico’s security crisis, but also how this violence becomes a symbolic resource contested within the political and communicative arena. Rather than constituting merely an isolated criminal act, the episode triggered a crisis of legitimacy that was rapidly absorbed into the discursive dynamics of contemporary polarization, particularly through the narrative construction of the “us versus them” antagonism. In this context, the ruling party and its support networks articulated a framing strategy aimed at shifting attention away from the crime and toward the opposition, portraying it as morally and politically responsible for the country’s structural problems.

The results of the analysis on X show that this dispute does not unfold exclusively through formal institutional channels, but rather through highly influential peripheral actors who function as agenda-setters. Through amplification mechanisms such as retweets, a dominant discourse centered on the moral delegitimization of the opposition was consolidated, employing labels such as “fascists”, “corrupt”, or “priianistas”. This reinforces internal cohesion within the ruling group while simultaneously deepening intergroup polarization, as anticipated by Sherif’s Realistic Conflict Theory.

Furthermore, the categorization of discourse made it possible to identify that, although the overall average of the conversation remained at intermediate levels of affective hostility, the most amplified actors reached higher levels associated with moral delegitimization. This confirms that polarizing discourse is not evenly distributed, but rather tends to concentrate in central nodes capable of imposing interpretations and emotional frames.

In sum, the Manzo case illustrates how, in contexts of crisis, digital political communication can function as a space for contesting legitimacy, where the opposition is constructed not as a democratic adversary, but as a moral enemy, thereby generating symbolic conditions that erode public debate and reinforce narratives of exclusion.

References

- Bandura, A. (1999). Moral Disengagement in the Perpetration of Inhumanities, Albert Bandura. Sagepub https://journals.sagepub.com/doi/10.1207/s15327957pspr0303_3
- Davidson, T., Warmesley, D., Macy, M., Weber, I. (2017). Automated Hate Speech Detection and the Problem of Offensive Language. *Proceedings of the International AAAI Conference on Web and Social Media*, 11(1), 512-515. <https://doi.org/10.1609/icwsm.v11i1.14955>
- El Universal. (2025). Mapa del Crimen Organizado en México. *El Universal*. <https://www.eluniversal.com.mx/interactivos/2023/mapa-crimen-organizado>
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51-58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Haslam, N. (2006). Dehumanization: An Integrative Review - Nick Haslam, 2006. *Personality and Social Psychology Review*. https://journals.sagepub.com/doi/10.1207/s15327957pspr1003_4
- Iyengar, S., Gaurav, S., Yphtach, L. (2026). Affect, Not Ideology, Public Opinion Quarterly, *Oxford Academic*. <https://academic.oup.com/poq/article-abstract/76/3/405/1894274?redirectedFrom=fulltextlogin=false>
- Lakoff, G. (with Internet Archive). (2004). Don't think of an elephant!: Know your values and frame the debate: the essential guide for progressives. White River Junction, Vt. *Chelsea Green Pub. Co.* <http://archive.org/details/dontthinkofeleph00lako>
- Mudde, C. (2004). The Populist Zeitgeist. *Government and Opposition*, 39(4), 541-563. <https://doi.org/10.1111/j.1477-7053.2004.00135.x>
- Müller, J. (2026). What Is Populism? *University of Pennsylvania Press*. <https://www.pennpress.org/9780812248982/what-is-populism/>
- Osorio, M. C. R., Jiménez, R. G., Salmorán, U. P. (2025). La estrategia de abrazos, no balazos: Causas y efectos de la delincuencia en México. *Yachana*, 14(2). <https://doi.org/10.62325/10.62325/yachana.v14.n2.2025.953>
- Sherif, M. (2015). Group Conflict and Co-operation: Their Social Psychology. *Group Conflict and Co-operation: Their Social Psychology*, 29, 1-192. <https://doi.org/10.4324/9781315717005>
- Waseem, Z., Hovy, D. (2016). Hateful Symbols or Hateful People? Predictive Features for Hate Speech Detection on Twitter. En J. Andreas, E. Choi, A. Lazaridou (Eds.), *Proceedings of the NAACL Student Research Workshop* (pp. 88-93). *Association for Computational Linguistics*. <https://doi.org/10.18653/v1/N16-2013>