

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Journal of Pragmatics

journal homepage: www.elsevier.com/locate/pragma

Appraising impoliteness on X: A case study of Isabel Díaz Ayuso

Sergei Sikorskii^{*}, María Luisa Carrió-Pastor

Department of Applied Linguistics, Universitat Politècnica de València, Camí de Vera, s/n, 46022 València, Spain



ARTICLE INFO

Article history:

Received 28 July 2025

Received in revised form 5 November 2025

Accepted 26 December 2025

Keywords:

Appraisal theory
 Digital impoliteness
 Gender and politics
 Political discourse
 Social media
 Evaluation
 Isabel Díaz Ayuso
 Spanish politics
 Twitter/X
 Corpus linguistics

ABSTRACT

This study applies Appraisal Theory to analyze evaluative patterns in digital political impoliteness, examining how systematic deployment of Judgment and Graduation resources functions as a form of pragmatic exclusion in hostile replies to Spanish politician Isabel Díaz Ayuso on X (formerly Twitter). Through manual annotation of 498 Spanish-language hostile replies collected between April and May 2025, we coded for five Judgment categories (Capacity, Tenacity, Normality, Veracity, Propriety) and intensification mechanisms (Force, Focus). Results reveal that Social Sanction judgments, such as Propriety- (63.5 %) and Veracity- (20.4 %), sharply outweigh Social Esteem evaluations, indicating a focus on moral delegitimization over competence critique.

Intensification was present in 70.1 % of replies, realized through typography, repetition, irony, and emoji. Notably, 29.5 % of replies contained gendered language, disproportionately associated with Propriety- and Normality-judgments ($\chi^2 = 52.85$, $p < 0.000001$), suggesting gendered language disproportionately co-occurs with moral and behavioral judgments. Eleven recurrent impoliteness strategies were identified, often combining multiple negative evaluations with multimodal amplification. This case study introduces the concept of *evaluative gatekeeping* to describe how hostile replies may use neutral-seeming assessment to challenge political legitimacy in gendered terms. These findings suggest that digital hostility toward female political figures may center on moral judgment rather than policy critique, raising implications for pragmatic exclusion and gendered participation in online political discourse.

© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Impolite replies to female politicians on social media represent a critical intersection of digital discourse, gender politics, and democratic participation. Research consistently demonstrates that women in political leadership face systematic online harassment that differs qualitatively from attacks on male counterparts, with hostility targeting personal characteristics rather than policy competence (Musi et al., 2025; Southern and Harmer, 2021). These interactions matter because they reflect deeper ideological structures perpetuating gender inequality while potentially undermining democratic engagement, as documented evidence shows online aggression causes women to become more cautious about political expression (Nadim and Fladmoe, 2021; Reich and Bachl, 2025).

^{*} Corresponding author

E-mail addresses: ssikors@upv.es (S. Sikorskii), lcarrio@upv.es (M.L. Carrió-Pastor).

In the Spanish context, political discourse on Twitter is notably confrontational, especially during electoral cycles (Gamir-Ríos et al., 2022; Torregrosa et al., 2023). Female politicians, particularly those in leadership positions, often encounter intensified scrutiny in this ideologically polarized space (Camargo-Fernández and Polo-Artal, 2024). Isabel Díaz Ayuso, President of the Community of Madrid, exemplifies this dynamic. Her prominent conservative stance and visibility make responses to her posts a revealing site for analyzing how female political authority is contested online.

Prior studies have mapped impoliteness strategies in digital political discourse, Saz-Rubio's (2023) analysis of Prime Ministers' replies offering a detailed typology of aggression. However, these frameworks focus primarily on the form of hostility rather than its evaluative content. This study builds on that work by examining how aggression against a specific female political figure is structured evaluatively, using Ayuso as a revelatory case for understanding gendered political assessment.

Appraisal Theory (Martin and White, 2005), particularly the Judgment subsystem, provides a structured framework for analyzing evaluations of human behavior along axes of social esteem (capacity, tenacity, normality) and social sanction (veracity, propriety). Prior applications to Spanish political discourse confirm its utility in identifying recurring evaluative patterns (Carrió-Pastor, 2025). Graduation resources, such as intensifiers, typography, and emoji, further modulate the force and focus of these evaluations (Makki and Zappavigna, 2022; Zappavigna, 2013).

This study applies Appraisal Theory to a curated sample of hostile replies to Ayuso to examine how political aggression against a female leader is structured evaluatively. The aim is to identify which traits are most frequently condemned, how intensely they are judged, and whether these patterns reflect gendered ideologies of leadership.

Research Question: How is Isabel Díaz Ayuso evaluated in impolite replies on X, and what Judgment and Graduation resources are mobilized to delegitimize her leadership in this specific political and cultural context?

2. Theoretical background

2.1. Digital impoliteness and incivility

Digital impoliteness encompasses intentional use of language to cause offense in online interactions, distinguished from incivility (subtler breaches of mutual respect) and intolerance (expressions violating democratic equality principles) (Culpeper, 2021; Rega et al., 2023; Rossini, 2022). These distinctions enable precise analysis: impoliteness typically targets individual characteristics while intolerance challenges democratic legitimacy itself.

Foundational frameworks from Culpeper (2011), Locher and Bousfield (2008), and Terkourafi (2015) conceptualize impoliteness through face-threatening acts, relational work, and frame-based conventionalization. However, these approaches focus on delivery mechanisms rather than systematically analyzing evaluative content, which is a limitation particularly problematic for political discourse where moral logic underlying attacks carries significant ideological implications.

Recent empirical research validates systematic patterns through large-scale analysis. Platform-specific studies show X exhibits higher incivility than Facebook due to character limitations and algorithmic amplification (Oz et al., 2018). Large-scale analysis reveals gendered patterns: women politicians receive attacks targeting appearance and sexual behavior rather than policy positions, while cross-cultural evidence demonstrates female leaders across opposing ideologies encounter similar misogynistic abuse strategies (Musi et al., 2025; Vladimirov and Fernández-Amaya, 2024).

Spanish political discourse requires frameworks sensitive to cultural specificities. Systematic corpus studies establish methodological precedents for analyzing Spanish political X discourse (Sánchez-Junquera et al., 2020; Torregrosa et al., 2023), while research consistently shows particularly high tension during electoral periods (Baviera et al., 2019; Guerrero-Solé, 2022). Recent work by Saz-Rubio (2023) and Carrió Pastor (2025) demonstrates how systematic frameworks can be applied to Spanish political leaders' Twitter responses, though these studies focus primarily on identifying strategic patterns rather than examining underlying evaluative logic.

2.2. Appraisal Theory: Judgment and Graduation

Appraisal Theory provides systematic frameworks for analyzing evaluative language through three subsystems: Attitude, Engagement, and Graduation (Martin and White, 2005). Unlike typological approaches listing impoliteness strategies, Appraisal maps discourse evaluative architecture through codable categories enabling intercoder reliability and cross-context comparisons.

Within Appraisal, the Judgment subsystem is particularly relevant for political discourse. It captures how speakers evaluate human behavior according to moral and social norms. Judgment is divided into two domains: social esteem, which includes Capacity (ability and competence), Tenacity (resolve and dependability), and Normality (conformity to expected behavior); and social sanction, which includes Veracity (truthfulness and honesty) and Propriety (ethics and morality).

This framework has proven useful in studies of political discourse. Ross and Caldwell (2020) show how Trump's tweets use Judgment to attack opponents' character. Cavasso and Taboada (2021) find that online comments often rely on character evaluations rather than emotional reactions.

In Spanish-language contexts, Carrió Pastor (2025) applied Appraisal Theory to aggressive replies to Prime Minister Pedro Sánchez, identifying six affect categories including culturally shaped patterns of dissatisfaction and moral

disapproval. This confirms that Appraisal can capture recurring evaluative strategies in Spanish political Twitter. The present study builds on this by focusing more specifically on Judgment and Graduation in replies to a female conservative leader.

Graduation encompasses resources that intensify or attenuate evaluative meanings. In digital contexts, Force mechanisms include lexical intensifiers (“absolutamente incompetente”), typography as emphasis (“¡MENTIROSA!”), repetition and emoji as amplifiers (e.g. angry-face emoji expressing outrage, thumbs-down emoji signalling disapproval) (Zappavigna, 2013). These amplify the impact of Judgment and signal emotional intensity or group alignment (Hansson et al., 2022; Makki and Zappavigna, 2022).

The extension of Graduation to multimodal elements requires theoretical justification, as traditional Appraisal Theory focused primarily on lexico-grammatical resources. However, recent work by Zappavigna (2013, 2022) and Hansson et al. (2022) demonstrates that emoji and typographic emphasis function as functional equivalents to lexical intensifiers in digital political contexts, operating within Graduation:Force to amplify evaluative stance. Their analysis of judgment in political protest online shows how visual elements perform analogous pragmatic work to traditional intensification mechanisms. We therefore treat emoji, capitalization, and excessive punctuation as multimodal Graduation resources that extend rather than replace the traditional framework, recognizing their role in amplifying and sharpening evaluative meaning in platform-specific ways.

Together, Judgment and Graduation offer a replicable framework for identifying not just how online aggression manifests, but what ideological and moral work it performs. The current study applies this framework to impolite replies directed at Isabel Díaz Ayuso, aiming to uncover the evaluative architecture that underpins gendered hostility in Spanish digital political discourse.

The systematic nature of Appraisal Theory's evaluative categories makes it particularly suited for examining how gendered ideologies may systematically structure political evaluation, revealing whether attacks on female leaders cluster around specific types of moral and social judgments that differ from those targeting male politicians.

2.3. Gendered evaluation and legitimacy attacks

Women in political leadership face systematic discursive challenges shaped by gender-role incongruity (Eagly and Karau, 2002). Political behavior that is agentic and assertive can violate normative expectations of femininity, while displays of warmth may be perceived as weak or incompetent. These tensions result in a “double bind” for women, where nearly any performance of leadership can trigger criticism.

This incongruity is not just discursive but evaluative. Delegitimizing replies to female politicians on X frequently cluster around specific Judgment categories within the Appraisal framework, particularly in the domain of social sanction. Insults such as *corrupta* (“corrupt”), *sinvergüenza* (“shameless”), and *mentirosa* (“liar”) signal the prevalence of Propriety and Veracity evaluations-targeting morality and truthfulness. These categories frame women not merely as poor leaders, but as ethically and personally unfit for public office. Such attacks operate as a form of gendered moral policing, invoking expectations of virtue, modesty, and honesty that are disproportionately applied to women in politics (Rohrbach et al., 2023; Vladimirou and Fernández-Amaya, 2024).

Large-scale studies confirm these tendencies. Southern and Harmer (2021) found that online replies to women MPs overrepresented accusations of dishonesty and disgrace. Similarly, Rheault et al. (2019) demonstrate that female politicians are significantly more likely than male counterparts to be targeted with comments attacking their appearance, personality, and personal life, rather than policy positions, highlighting a systematic focus on character over competence in digital aggression. In the Spanish context, Carrió Pastor (2025) identified moral disapproval and culturally negative affect as recurrent themes in hostile replies. Computational studies likewise reveal a high frequency of lexically identifiable hate speech markers targeting women's morality, credibility, and mental stability (Iranzo-Cabrera et al., 2025).

These patterns of evaluative hostility are not without consequence. In a systematic review of 73 studies, Koch et al. (2025) found that 81 % reported psychological distress or withdrawal from political engagement among women targeted by online hostility, underscoring how gendered evaluation functions not just discursively but politically, as a tool of exclusion.

Online harassment is not merely spontaneous aggression but a systematic ideological practice that encodes gender norms through evaluation. By applying Appraisal Theory to this corpus, the present study asks whether gendered hostility follows regular, codable patterns of social and moral judgment or whether it functions as a diffuse, affective phenomenon. The analysis that follows investigates how these Judgment categories, along with Graduation markers, construct delegitimization within digital political conflict.

3. Data and method

This study adopts a mixed-methods approach grounded in Appraisal Theory to examine how hostile political discourse operates evaluatively in replies to Isabel Díaz Ayuso on the platform X (formerly Twitter).

3.1. Corpus selection and scope

The dataset used in this study consists of 498 Spanish-language replies to Isabel Díaz Ayuso, collected via X's advanced search interface between April 1 and May 9, 2025. This time frame was selected to capture a period of high political salience without distortion by a single dominant news cycle.

This represents a purposive sample focused on evaluative discourse, as the study examines patterns of hostile assessment rather than general mention patterns. Replies were collected using search terms based on Ayuso's name and variants, with systematic inclusion criteria: (1) Spanish language content, (2) direct reply format, (3) clear negative evaluative stance toward Ayuso, and (4) individual authorship (excluding identifiable bot accounts). This hostile-reply focus enables intensive analysis of evaluative weaponization patterns, though it precludes comparison with neutral evaluation baselines. We treat hostile replies as a distinct discourse variety requiring systematic analysis of its evaluative architecture, following established practice in impoliteness research (Culpeper, 2011; Saz-Rubio, 2023).

The present study draws on the final analytical 498-reply sample that exhibited clear evaluative stance toward Ayuso. This selection criterion aligns with the study's focus on hostile assessment patterns, excluding neutral news shares, factual queries, and non-evaluative mentions. The resulting sample size balances intensive manual annotation depth and feasibility for qualitative annotation and aligns with prior studies of impoliteness and appraisal in political discourse (e.g. Musi et al., 2025; Ross and Caldwell, 2020). The resulting 498-post sample includes varying tone, intensity, and topic.

This single-case design enables intensive pragmatic analysis of systematic evaluative patterns while maintaining analytical depth. Following established practice in critical discourse analysis (Fairclough, 2013) and case study methodology (Yin, 2018), we treat Ayuso as a revelatory case that illuminates broader mechanisms of gendered political evaluation in digital contexts. While this approach prioritizes analytical depth over comparative breadth, the systematic annotation framework and statistical analysis provide replicable insights into how hostile evaluation operates as a form of digital political exclusion.

3.2. Annotation procedure

Annotation was performed using the Judgment and Graduation subsystems of Appraisal Theory (Martin and White, 2005). Each reply was manually coded for:

- Judgment type (Capacity, Tenacity, Normality, Veracity, Propriety)
- Polarity (Positive/Negative)
- Graduation (Force and Focus)

The coding framework is informed by recent applications of Appraisal Theory to Spanish political discourse (Carrió-Pastor, 2025; Cavasso and Taboada, 2021; Saz Rubio, 2024), particularly for refining the operationalization of Judgment categories in hostile digital environments. The framework was further adapted to account for typographic and multimodal cues frequently observed in online political impoliteness, including capitalization, excessive punctuation, emoji, and repetition. These were treated as Graduation:Force resources that amplify evaluative meaning.

Graduation: Focus, particularly in its ironic realization, was annotated when evaluative polarity was pragmatically inverted. For example, sarcastic statements such as “Muy inteligente” or “Qué profesional eres” were identified as sharpened or ironic evaluations when the communicative context clearly signaled insult rather than praise. These constructions were especially frequent in replies targeting Capacity and Propriety judgments.

Annotation proceeded in two phases. In the pilot phase, a subset of 200 replies (40 % of the corpus) was independently annotated by two trained coders following the Appraisal Theory framework (Judgment and Graduation subsystems). Interrater agreement reached Fleiss' $k = 0.54$, and all discrepancies were reviewed and adjudicated to refine coding guidelines.

The remaining 298 replies (60 %) were then annotated by the lead author using the revised scheme. We understand that this single-annotator approach is common in pragmatics research but it introduces a potential for consistency drift over time. To mitigate this risk, the annotation protocol included detailed reasoning logs and category decision criteria. Future work should incorporate post-pilot coder triangulation, such as subsample re-annotation or adjudication checks, to further validate interpretive consistency across larger datasets.

Judgment categories and gender labels were assigned through interpretive analysis following established semantic tests (Martin and White, 2005). However, in some cases, a single utterance could plausibly be interpreted through multiple evaluative lenses (e.g., Normality-vs. Veracity-). To solve this issue, we have used Gendered Language Identification Criteria:

1. Biological/reproductive references (e.g., “está en sus días”)
2. Appearance-based sexualization (e.g., “enseña las tetas”)
3. Infantilizing diminutives (e.g., “nena,” “churri”)
4. Stereotypical female behavior attribution (e.g., “histérica”) “histérica”)
5. A hypothetical gender-substitution heuristic: Would this expression be used against a male politician in equivalent context?

Borderline cases were excluded from gendered analysis to maintain conservative coding standards.

The following reply (Fig. 1) illustrates how multiple evaluative and impoliteness resources co-occur in a single reply:

Item:	Tiene que ir durísimo ser así de retrasada (clown-face emoji; eye-roll emoji)	
Translation:	It must be so tough to be that retarded (clown-face emoji; eye-roll emoji)	
Annotation Categories		Interpretation
Judgment:	Normality- (explicit slur: retrasada) + implied Veracity-	Targets cognitive abnormality and possibly willful ignorance, delegitimizing Ayuso as irrational or clueless.
Graduation:	Force (así de), Focus (irony, emoji-based ridicule)	Intensifies the insult lexically and pragmatically via mock-sympathy and nonverbal sharpening.
Emoji Functions:	clown-face emoji: ridicule; eye-roll emoji: sarcasm/contempt	Reinforce judgment through visual cues — the clown emoji distills Ayuso into ridicule; eye-roll signals dismissal.
Impoliteness:	Direct Insult, Mock Empathy, Visual Harassment	Combines explicit offense, sarcastic structure, and emoji aggression to construct a layered hostile message.
Gendered?	Borderline (no direct stereotype)	"Retrasada" has a feminized form, but no explicitly gendered frame or stereotype is activated here.

Fig. 1. Annotated example of multilayered impoliteness in a reply targeting Isabel Díaz Ayuso.

All data used in this study were publicly accessible on X. In line with institutional ethical guidelines for social media research, no usernames, profile images, or personal identifiers are included in any data excerpts.

3.3. Strategy typology construction

The impoliteness strategy types outlined in [Table 1](#) were not pre-assigned or automatically extracted. Rather, they were developed inductively, through qualitative clustering of co-occurring Judgment categories and Graduation mechanisms, following the completion of Appraisal annotation. This approach builds on recent work by [Saz-Rubio \(2023\)](#), who emphasizes the importance of data-driven adaptations of Culpeper's (1996, 2011) model when analyzing impoliteness in digitally mediated discourse.

Eleven recurrent strategy types were identified (e.g., Direct Insults, Sarcastic Praise, Credibility Attacks, Infantilization). While some categories (such as Direct Insults or Mock Politeness) align loosely with established impoliteness super-strategies, others (e.g., Sexualized Moral Judgment, Clustered Delegitimization) emerged uniquely from the evaluative and multimodal patterns observed in this corpus. These functional labels offer a systematized, usage-based framework for interpreting how hostile replies construct public delegitimation through patterned combinations of judgmental stance and intensification.

[Table 1](#) summarizes 11 impoliteness strategies identified through inductive analysis of judgmental stance and graduation mechanisms.¹ While some categories align with established super-strategies (e.g., Culpeper, 1996), the framework is grounded in corpus patterns.

Table 1

Functionally derived typology of impoliteness strategies mapped to appraisal resources.

Impoliteness Strategy	Appraisal Category	Spanish Examples (English Translation)	Evaluative Function	Gendered Implications
Direct insults	Judgment: Capacity-	inútil (useless), incompetente (incompetent), no sirve para nada (good for nothing)	Attacks leadership competence	Questions women's fitness to govern
Character assassination	Judgment: Propriety-	corrupta (corrupt), sinvergüenza (shameless), inmoral (immoral)	Moral delegitimization	Moral policing of female behavior

(continued on next page)

¹ Some examples (e.g., 'Qué lista eres', 'CORRUPTA') appear in multiple tables to illustrate intersecting functions (Judgment, Graduation, Gendered Framing).

Table 1 (continued)

Impoliteness Strategy	Appraisal Category	Spanish Examples (English Translation)	Evaluative Function	Gendered Implications
Credibility attacks	Judgment: Veracity-	mentirosa (liar), falsa (fake), embustera (deceiver)	Undermines credibility	Reinforces stereotypes of female dishonesty
Ridicule/mockery	Judgment: Normality-	loca (crazy), ridícula (ridiculous), chalada (nutcase)	Positions as deviant or irrational	Frames women as emotionally unfit or socially inappropriate
Weakness accusations	Judgment: Tenacity-	débil (weak), blanda (soft), cobarde (coward)	Questions resolve and resilience	Targets female emotional fragility
Intensified hostility	Graduation: Force	¡MENTIROSA! (LIAR!), muyyy tonta (verrry stupid), !!! (excessive punctuation)	Amplifies negative evaluations	Escalates emotional aggression
Sarcastic praise	Graduation: Focus (ironic)	Qué lista eres (“Aren’t you clever”), muy inteligente (“so intelligent”) [sarcastic]	Inverts polarity via irony	Patronizing dismissal
Visual harassment	Graduation: Force (multimodal)	angry-face emoji (anger), red-face emoji (rage), lipstick emoji, eggplant emoji, mocking memes	Intensifies evaluation via visuals	Sexualization/intimidation via semiotics
Sexualized appearance critiques	Judgment: Propriety-/Normality-	puta (whore), zorra (slut), enseña las tetas (show your tits)	Sexualized moral/social judgment	Undermines legitimacy via body objectification

The classification is function-oriented, not face-based: strategies are not mapped to positive or negative face orientation but are instead grounded in how hostile evaluation is enacted through language and multimodal emphasis. This analytical stance reflects the shift in recent impoliteness research away from fixed theoretical categories toward context-sensitive, emergent typologies adapted to platform-specific affordances.

Each reply was labeled with one or more strategy types after Appraisal coding was complete. Strategy assignment was performed manually by the lead author through interpretive analysis. Replies frequently involved clusters of strategies, especially when multiple negative judgments (e.g., Propriety- + Veracity-) were stacked and intensified through Graduation: Force or Focus. These co-occurring resources are central to how evaluative impoliteness is performed, and the resulting taxonomy captures the recurrent pragmatic forms through which political figures are delegitimized in hostile replies.

4. Results

This section presents the evaluative and graduation patterns observed in the 498 hostile replies to Isabel Díaz Ayuso. All tweets in the sample contained at least one negative evaluation, confirming the dataset’s focus on impolite discourse. The analysis identifies dominant Judgment categories, Graduation mechanisms, and the intersection of gendered aggression with evaluative stance.

4.1. Judgment patterns: moral condemnation prevails

Out of 498 annotated replies, 415 (83.3 %) contained at least one negative Judgment category. The remaining 84 replies either expressed non-judgmental hostility (e.g., phatic insults, mockery without clear evaluation), were ambiguous, or relied solely on Graduation or gendered framing without explicit judgmental content. Negative assessments were distributed across all five Judgment subtypes, though some appeared far more frequently and forcefully than others.

Propriety-judgments overwhelmingly dominated the corpus, appearing in 316 replies (63.5 %) as either a primary or secondary category. These evaluations framed Ayuso as morally corrupt, dishonest, unethical, or shameless. Lexical triggers included *corrupta*, *sinvergüenza*, *inmoral*, and metaphors like *rata*, often intensified by exclamatives or typographic force (e.g., *¡¡¡CORRUPTA!!!*). These were used to delegitimize her political legitimacy and trustworthiness, positioning her as a morally unfit public servant.

Tenacity-was the second most frequent subtype (173 replies, 34.7 %), typically characterizing Ayuso as weakwilled, unstable, cowardly, or emotionally fragile. Expressions such as *cobarde*, *llora*, *no aguanta nada*, and *histeriquea* implied unreliability or poor leadership resilience. These attacks often intersected with stereotypically gendered framings of emotional excess, though not always explicitly.

Normality-judgments were present in 152 replies (30.5 %), portraying Ayuso as irrational, erratic, or deviant from expected political behavior. Many examples were humorous or sarcastic, invoking absurdity or discreditation (e.g., *vaya personaje*, *estás fatal*, or visual shorthand such as the clown-face emoji functioning as ridicule). These judgments attacked not her morality or strength per se, but her perceived credibility or normalcy as a political actor.

Veracity-judgments, though less common (102 replies, 20.4 %), were pointed. These framed Ayuso as deceptive, manipulative, or out of touch with reality, often using terms like *mentirosa*, *todo falso*, or *vive en su burbuja*. Such evaluations questioned not only her honesty but also her epistemic alignment with the public or truth itself.

Capacity-judgments appeared least often (86 replies, 17.3 %), targeting Ayuso's intelligence, competence, or effectiveness as a leader. Examples included *subnormal*, *cero neuronas*, or sarcastic praise like *qué lideresa*. These were sometimes combined with Normality or Tenacity in clusters that reinforced a portrayal of cognitive or professional unfitness.

Together, these data confirm that Social Sanction judgments (Propriety- + Veracity-) were far more prominent than Social Esteem judgments (Capacity-, Tenacity-, Normality-). The replies focused less on Ayuso's competence or temperament and more on moral blame, dishonesty, and abnormality, aligning with broader patterns of public delegitimation through moral discourse.

This trend is visually summarized in Fig. 2, which displays the distribution of all negative Judgment types, combining primary and secondary occurrences. Propriety-judgments dominate sharply, while Tenacity- and Normality-follow as auxiliary modes of attack. Veracity- and Capacity-are less frequent but often cluster with other judgments.

While these associations are statistically significant, the effect sizes are small to moderate (Cramér's $V = 0.19$), indicating that Judgment categories explain approximately 3.6 % of variance in Graduation patterns.

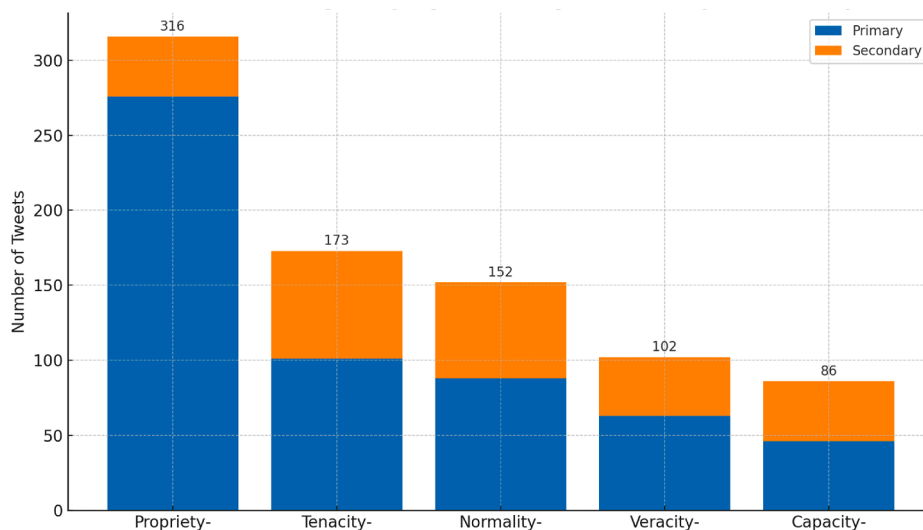


Fig. 2. Distribution of negative judgment categories in 498 replies.

4.2. Graduation: force and focus

Graduation resources were pervasive across the dataset, with 349 of 498 replies (70.1 %) exhibiting some form of intensification or attitudinal sharpening. Both Force and Focus mechanisms were mobilized to amplify judgmental stance, often layering emphasis with sarcasm, repetition, or multimodal cues.

Force was the more common type, appearing in 268 replies, often through typographic emphasis (e.g., all caps, exclamation marks), lexical intensifiers (e.g., *muy*, *tan*, *así de*), or repetition. Examples include *iii CORRUPTA!!!*, *muy muy mal*, or sequences like *patética, patética, patética*. These mechanisms boosted the severity, urgency, or ridicule of the judgment expressed — especially in tweets targeting Propriety- or Tenacity-.

Focus was present in 186 replies, typically realized through irony, quotation distancing, or sarcastic constructions that inverted the surface polarity. Phrases such as *qué crack*, *muy profesional*, or *gran lideresa* carried a sharpened, mocking tone, often identifiable only through context or known evaluative drift. These constructions disproportionately targeted Capacity- and Propriety-judgments. In some cases, users combined both types of Graduation in a single utterance, for instance, *QUÉ LISTA ERES* (*sarcastic praise accompanied by an eye-roll emoji*), where typographic emphasis and sarcasm co-occur.

A subset of 105 replies (21.1 %) featured both Force and Focus, creating dense clusters of evaluative intensity. These combinations were especially common in replies using sarcastic praise amplified by emoji-based ridicule (e.g. *clown-face emoji*, *eye-roll emoji*), mock quotations, or exaggerated exclamations.

A more granular analysis of Graduation revealed five recurring intensification strategies (Fig. 3)²: lexical intensifiers, typographic emphasis, irony/sarcasm, multimodal cues (e.g., emoji), and repetition. These often co-occurred, particularly in tweets combining moral judgment with sarcastic stance. Although 349 replies contained Graduation, the total frequency of

² Fig. 3. Proportional breakdown of Graduation mechanisms observed in the 349 replies (70.1 % of total) that contained at least one intensification or attitudinal sharpening resource. Percentages reflect the proportion of replies containing each type of mechanism. Many replies contain multiple Graduation types (e.g., typographic emphasis + emoji), so total exceeds 100 %.

mechanisms listed reflects only the most salient types. Several replies contain more than one mechanism (e.g., typographic emphasis and emoji), while others employed less frequent forms such as elongation, paralinguistic repetition, or stylized mockery not reflected in the count.

A total of 153 replies employed multimodal Graduation (e.g., emoji-based resources, hyperlinks). Common emoji types included the poop emoji, clown-face emoji, eye-roll emoji, and skull emoji, which typically appeared alongside sarcastic or typographically intensified text. In addition, 127 replies incorporated links or images referencing scandals or circulating memes, thereby contextualizing and reinforcing the evaluative judgment.

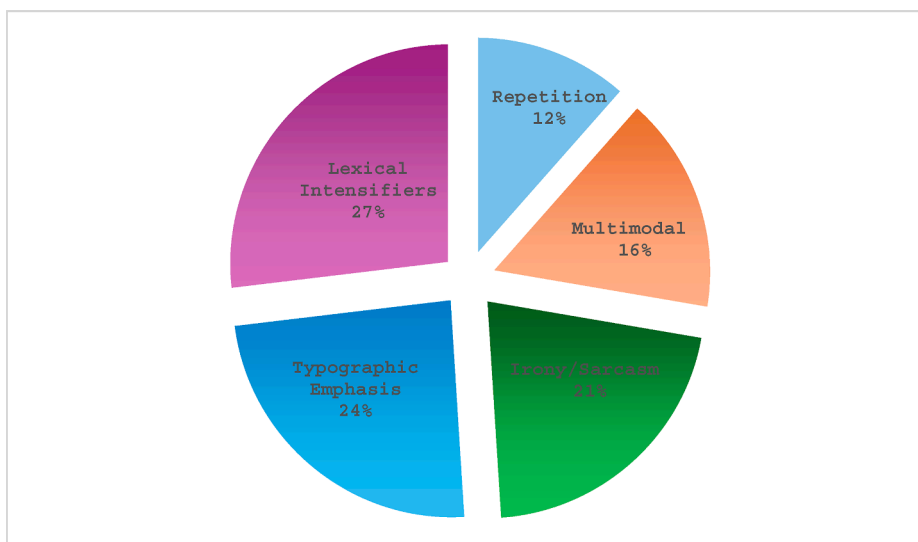


Fig. 3. Distribution of graduation mechanisms in impolite replies.

4.3. Judgment–Graduation clusters

Among the 415 replies containing negative Judgment, a total of 349 (84.1 %) also included one or more Graduation mechanisms. This suggests that intensification resources were present in the majority of evaluative replies in the corpus.

Table 2 presents the distribution of Graduation types (Force, Focus, Both, or None) across the five Judgment categories. Propriety-judgments showed the highest overall frequency of co-occurring Graduation mechanisms, with 121 instances of Force, 63 of Focus, and 65 replies containing both. Normality- and Tenacity-also exhibited frequent combinations, particularly through the use of emoji, ironic constructions, or typographic emphasis. By contrast, Capacity- and Veracity-judgments were less frequently intensified, and more often occurred without additional Graduation.

Table 2
Distribution of graduation mechanisms across judgment categories.

Judgement Type	Force	Focus	Both Force & Focus	None	Total Mentions ^a
Propriety-	121	63	65	123	372
Normality-	41	18	17	36	112
Tenacity-	34	11	9	29	83
Capacity-	14	6	0	26	46
Veracity-	15	10	0	26	51
Total	225	108	91	240	664

^a Rows reflect co-occurrence counts across all tweets where the respective Judgment category appeared (primary or secondary). Because tweets can contain multiple Judgment categories and Graduation types, row totals exceed the number of tweets in the corpus (N = 498).

While Force mechanisms (e.g., lexical amplifiers, CAPS, repetition, punctuation) were the most frequently observed across the dataset, Focus resources such as irony, quotation distancing, and multimodal cues (e.g., emoji) were notably present in tweets involving sarcasm or rhetorical inversion.

Table 3 provides representative examples of these pairings. Each row illustrates a typical combination of Judgment and Graduation types observed in the data, along with its corresponding Spanish example and English translation. For instance, *Muy corrupta, totalmente inmoral* exemplifies Propriety- + Force (Lexical), while *Qué normalita eres* (accompanied by a clown-face emoji and an eye-roll emoji) demonstrates a combination of Normality-judgment and Focus through irony and emoji-based ridicule.

Table 3
Examples of judgment–graduation pairings.

Judgment Type	Graduation Type	Example and its translation
Propriety-	Force (lexical)	Muy corrupta, totalmente inmoral (<i>Very corrupt, totally immoral</i>)
Veracity-	Force (CAPS)	MENTIROSA, TODO FALSO (<i>LIAR, EVERYTHING IS FALSE</i>)
Normality-	Focus (Irony/Emoji)	Qué normalita eres (“Aren't you normal?”, sarcastic praise accompanied by a clown-face emoji and an eye-roll emoji)
Tenacity-	Force + emoji	Cobarde... otra vez, Débil (“coward... Again, weak”; intensified through emoji-based dismissiveness)
Capacity-	Focus (irony)	Toda una líder..., no das una (<i>Such a leader..., you don't get one thing right</i>)

4.4. Gender weaponization

Out of 498 annotated replies, 147 tweets (29.5 %) were marked as containing gendered expressions or evaluations, either explicitly or implicitly. These replies included references to appearance, sexual behavior, emotionality, or stereotypically feminized traits.

Gendered language most frequently co-occurred with Normality- ($n = 41$) and Propriety- ($n = 45$) judgments. Other co-occurrences included Tenacity- ($n = 30$), Capacity- ($n = 18$), and Veracity- ($n = 13$). The remaining 351 judgmental tweets did not contain gendered language. Table 4 summarizes the co-occurrence of gender weaponization by Judgment type.

Table 4
Gendered vs. Non-Gendered Judgments by Category.

Judgment Type	Gendered	Non-Gendered	Total
Propriety-	45	327	372
Normality-	41	71	112
Tenacity-	30	53	83
Capacity-	18	28	46
Veracity-	13	38	51

A Chi-square test of independence confirmed a statistically significant association between Judgment category and the presence of gendered language: $\chi^2(4, N = 498) = 52.85, p < 0.000000001$. This indicates that the use of gendered language was not evenly distributed across Judgment types but was more prevalent in certain evaluative domains, particularly Propriety- and Normality-.

A total of 12 tweets included multiple gendered expressions, often combined with typographic or multimodal Graduation (e.g., , quotation marks). These tweets typically layered insults targeting morality, normality, and competence through gendered frames. To illustrate the range of expressions, Table 5 provides representative examples of gendered language from the corpus, categorized by their primary evaluative function.

Table 5
Representative examples of gendered expressions in judgmental replies.

Spanish Excerpt	English Translation	Judgment Type	Gendered Mechanism
<i>Típica nena con rabieta</i>	Typical baby girl throwing a tantrum	Normality-	<i>Infantilization</i>
<i>Enséñanos las tetas, churri</i>	Show us your tits, babe	Propriety-	<i>Sexualized mockery</i>
<i>Como buena loca, todo mal</i>	Like a proper lunatic, everything wrong	Normality-	<i>Gendered emotionality</i>
<i>Está en sus días, seguro</i>	She must be on her period	Tenacity-	<i>Biological stereotype</i>
<i>Una histérica sin oficio ni beneficio</i>	A hysteric with no purpose or use	Capacity-	<i>Mental instability trope</i>
<i>Calladita estás más mona</i>	You're prettier when you're quiet	Veracity-	<i>Silencing + appearance</i>

Replies were considered gendered if the content met one or more of the following criteria:

- Referenced stereotypical female behavior (e.g., *hysteria, vanity, fragility*)
- Used diminutives or infantilizing forms (e.g., *nena, churri*)
- Contained sexual or appearance-based remarks
- Failed the a hypothetical gender-substitution heuristic, i.e., would not likely be used to critique a male figure in the same context

4.5. Statistical association between Judgment and Graduation

To assess whether the presence and type of Graduation resources varied systematically across different Judgment categories, a Chi-square test of independence was conducted on the full dataset of 498 replies. Graduation was categorized into four types: Force only, Focus only, Both, or None, and cross-tabulated with the five negative Judgment types.

The analysis revealed a statistically significant association between Judgment type and Graduation type: $\chi^2(12, N = 498) = 31.07, p = 0.0019$, with a Cramér's $V = 0.19$, indicating a small to moderate effect size. This effect size (Cramér's $V = 0.19$) represents a small-to-moderate association according to conventional interpretation guidelines (Cohen, 2013), indicating that approximately 3.6 % of the variance in Graduation type can be attributed to Judgment category. While statistically significant, this suggests that Judgment–Graduation co-occurrence patterns show systematic tendencies rather than strong deterministic relationships.

As shown in Table 2, Propriety-judgments were most frequently accompanied by Graduation mechanisms, especially Force (e.g., lexical intensifiers, typographic emphasis) and Both. Of the 372 Propriety-tweets, 249 (66.9 %) contained either Force, Focus, or both. Normality- and Tenacity-judgments also exhibited high Graduation co-occurrence, particularly in the form of Focus resources such as irony, emoji, and quotation distancing.

In contrast, Veracity-judgments were more likely to occur without Graduation, with 25 of 51 tweets (49 %) showing no intensifying or focusing mechanisms. Capacity-judgments were frequently paired with Focus (e.g., sarcastic praise), though overall Graduation presence was lower than in Propriety- or Normality-judgments.

These frequency distributions are detailed in Table 2, and illustrative examples of Judgment–Graduation pairings are provided in Table 3. This pattern of results supports the presence of distinct Judgment–Graduation combinations across the corpus.

4.6. Judgment clusters and strategy types

A total of 192 replies (46.4 %) included multiple negative judgments within the same tweet. These judgment clusters typically combined two or more evaluative stances (e.g., Propriety- + Veracity-, or Normality- + Tenacity-), often layered with intensifying Graduation mechanisms. The most frequent combinations were:

1. Propriety- + Veracity- (e.g., *mentirosa corrupta*)
2. Normality- + tenacity- (e.g., *ridícula y cobarde*)
3. Propriety- + normality- (e.g., *vaya personaje, vergüenza ajena*)

Based on recurrent patterns of co-occurring judgments and Graduation resources, a typology of impoliteness strategies was developed. Eleven strategy types were identified across the corpus, each labeled after annotation and coded by the lead author. Out of 498 replies, 426 tweets (85.5 %) were assigned one or more strategy labels. In Fig. 4, you can see the most frequently observed types.

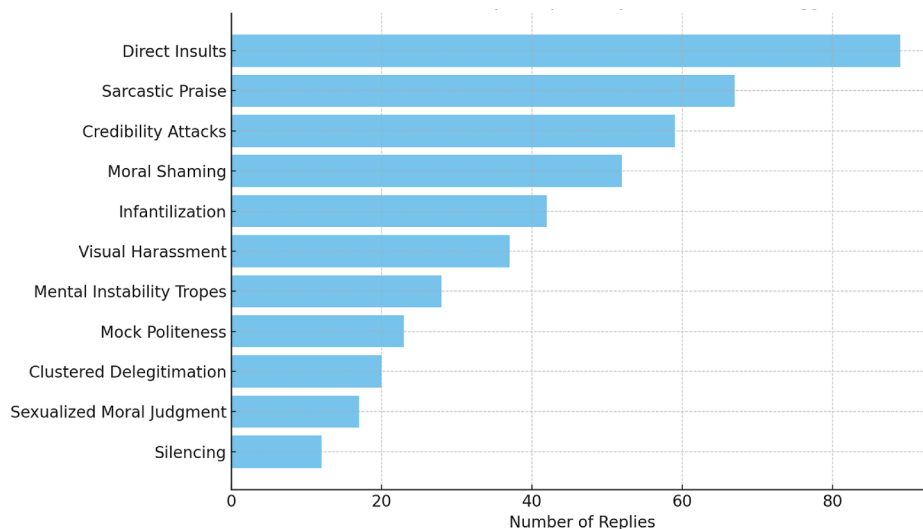


Fig. 4. Number of annotated replies per impoliteness strategy.

5. Discussion

5.1. Research aim and key findings

This study examined how Isabel Díaz Ayuso is evaluated in impolite replies on X, and what Judgment and Graduation resources are mobilized to delegitimize her leadership. The findings revealed a dominance of moral and social sanction judgments (Propriety-, Veracity-), frequently intensified through typographic and multimodal mechanisms. Nearly one-third of replies contained gendered expressions, showing significant association with Normality- and Propriety-judgments. These patterns suggest that hostile evaluation of this prominent conservative female politician operates through systematic deployment of moral assessment resources, warranting investigation of whether similar patterns characterize responses to other female political figures.

5.2. Interpreting judgment patterns: beyond competence-based critique

The predominance of Social Sanction judgments, such as Propriety- (63.5 %) and Veracity- (20.4 %) over Social Esteem categories reveals a distinct evaluative logic in gendered political hostility. While [Martin and White's \(2005\)](#) Appraisal framework provides robust categorical distinctions, the current data exposes how Propriety-judgments in digital political contexts often conflate moral transgression with sexualized evaluation. Terms like *corrupta* and *sinvergüenza* appeared alongside appearance-based attacks and infantilizing diminutives, suggesting that moral evaluation becomes a vehicle for gender-specific delegitimization.

This conflation of moral and sexualized evaluation within Propriety-judgments reveals potential limitations in Appraisal Theory's categorical boundaries when applied to gendered political discourse. The framework's Social Sanction domain may require subcategorization distinguishing moralethical evaluation from gender-specific virtue policing, particularly in contexts where political authority intersects with cultural expectations of feminine behavior.

This pattern aligns with [van Leeuwen's \(2007\)](#) analysis of legitimation strategies, where moral evaluation serves ideological exclusion. However, the systematic clustering of attacks around character rather than competence indicates what might be termed “moral bypassing”, i.e. the strategic avoidance of policy-based critique in favor of character assassination. The relative scarcity of Capacity-judgments (17.3 %) contrasts sharply with typical political discourse, where professional competence often becomes a primary target.

These findings extend [Saz-Rubio's \(2023\)](#) work on hostile political replies by demonstrating how evaluative patterns differ systematically when targeting female versus male political figures. The data suggests that gendered political hostility operates through distinct moral frameworks that frame political assertiveness as virtue violation rather than policy disagreement.

5.3. Graduation as intensification infrastructure

The pervasive presence of Graduation mechanisms (70.1 % of replies) reveals systematic patterns in how hostile evaluation is amplified. While Appraisal Theory conceptualizes Graduation primarily through lexicogrammatical resources, the current findings necessitate expanding this framework to account for multimodal intensification prevalent in digital political discourse.

Force mechanisms—including typographic emphasis (CAPS, excessive punctuation), lexical intensifiers, and repetition—appeared most frequently (53.8 % of graduated replies). However, the 16 % of replies using multimodal resources (emoji, memes, visual links) demonstrates how digital platforms enable new forms of evaluative amplification that exceed traditional linguistic categories. Emoji such as the clown-face emoji and the poop emoji function analogously to lexical intensifiers but add dimensions of visual ridicule absent from purely textual evaluation.

Focus mechanisms, particularly ironic constructions like *qué crack* or *muy profesional*, operated through pragmatic inversion that [Culpeper \(2011\)](#) identifies as characteristic of indirect impoliteness. These constructions enable plausible deniability while sharpening negative evaluation, suggesting that digital political discourse increasingly relies on rhetorical ambivalence rather than direct confrontation.

The systematic co-occurrence of Judgment and Graduation types ($\chi^2(12, N = 498) = 31.07, p = 0.0019$) indicates that intensification is not randomly distributed but follows identifiable patterns that amplify specific evaluative domains—particularly moral and behavioral judgments.

The small-to-moderate effect size (Cramér's $V = 0.19$) in Judgment–Graduation associations suggests that speakers have systematic but flexible pragmatic choices in how they intensify different types of moral and social evaluation. This pattern indicates that while Propriety- and Normality-judgments show stronger co-occurrence with intensification mechanisms, the relationship reflects pragmatic tendencies rather than rigid linguistic rules, consistent with the strategic deployment of evaluation for political exclusion.

5.4. Gendered evaluation: systematic double standards

The significant association between gendered language and specific Judgment categories ($\chi^2(4, N = 498) = 52.85, p < 0.000000001$) reveals how contemporary political misogyny operates through seemingly neutral evaluative

frameworks. The concentration of gendered expressions within Normality- (41 instances) and Propriety- (45 instances) judgments demonstrates systematic activation of gender-role expectations in political evaluation.

The systematic co-occurrence of gendered language with Normality- and Propriety-judgments reflects what [Eagly and Karau \(2002\)](#) term 'role incongruity', i.e. where female political assertiveness violates cultural femininity expectations, triggering evaluative responses that frame leadership behavior as personality deviation rather than policy disagreement. This pattern suggests evaluation functions as a mechanism for gender-role enforcement in digital political spaces.

These patterns extend beyond individual harassment to reflect broader cultural tensions around female political authority. The documented psychological consequences of systematic online aggression ([Koch et al., 2025](#)) suggest that gendered evaluation functions as a gatekeeping mechanism that potentially deters women from political participation through sustained character assassination.

5.5. Methodological reflections and framework limitations

While the annotation framework successfully captured multiple evaluative layers, several challenges emerged that illuminate broader theoretical limitations. Ironic constructions proved particularly difficult to disambiguate, often requiring contextual inference unavailable through isolated tweet analysis. Similarly, emoji interpretation involved subjective judgments about visual meaning that resist systematic codification.

These challenges reflect inherent tensions in applying Appraisal Theory to multimodal, context-collapsed social media discourse where meaning often depends on shared cultural knowledge and pragmatic inference. The 0.76 inter-rater agreement, while substantial, masks areas of interpretive uncertainty that future research must address more systematically.

More fundamentally, the current findings reveal limitations in Appraisal Theory's categorical structure when applied to gendered political discourse. The frequent conflation of moral and sexualized Propriety-judgments suggests potential need for subcategorization within Social Sanction domains. Similarly, the prevalence of multimodal intensification indicates that Graduation's Force dimension requires systematic expansion beyond lexico-grammatical resources.

5.6. Theoretical implications and proposed extensions

The systematic patterns identified in this study suggest several directions for theoretical development. First, Appraisal Theory's Judgment framework might benefit from gender-sensitive modifications that recognize how evaluative categories can encode cultural biases, particularly in political contexts where leadership performance intersects with gender-role expectations. These findings demonstrate how Appraisal Theory can capture pragmatic patterns invisible to traditional impoliteness frameworks.

Second, the prevalence of multimodal Graduation indicates need for systematic inclusion of emoji, typography, and visual elements as evaluative resources within the framework. These mechanisms function analogously to lexical intensifiers but enable additional dimensions of evaluation (visual ridicule, affective amplification) that current theory underrepresents.

Third, the patterns observed in this case suggest a phenomenon we term 'evaluative gatekeeping', i.e. the systematic use of ostensibly neutral assessment to delegitimize political participation. While this concept emerges from analysis of one politician's hostile replies, it offers a framework for investigating how evaluation functions as exclusion across broader political contexts.

This concept extends [van Leeuwen's \(2007\)](#) legitimation analysis by showing how systematic evaluation functions as pragmatic exclusion in digital political contexts.

5.7. Broader implications for digital political discourse

The findings reveal how digital platforms may systematically amplify certain forms of political hostility through affordances that reward affective engagement over substantive critique. When hostile replies receive algorithmic amplification through engagement metrics, systematic character assassination may become more visible and influential than policy-based discussion.

The documented patterns have implications beyond individual harassment for democratic discourse quality. If female politicians consistently face moral delegitimization rather than policy-based opposition, political participation itself becomes conditional on conformity to gendered virtue expectations that may exclude voices challenging traditional power arrangements.

These patterns suggest need for critical digital literacy education that recognizes how seemingly neutral evaluation can encode systematic bias, and platform design considerations that account for how affordances may disadvantage marginalized political actors through amplified hostile assessment.

5.8. Limitations and future directions

Several methodological constraints emerge from this study's specific design choices and analytical approach, affecting the scope and transferability of findings.

While the pilot phase achieved interrater reliability ($K = 0.54$) on 200 replies, the remaining 398 replies (80 % of the corpus) were annotated by the lead author alone. This single-annotator approach, though common in discourse analysis, introduces potential consistency drift over the annotation period. Additionally, several analytical decisions required interpretive judgment that resists systematic codification: distinguishing ironic from literal evaluations, interpreting emoji pragmatic functions, and applying a hypothetical gender-substitution heuristic, assessing whether equivalent expressions would plausibly target male politicians. In many cases, such reversals produced pragmatically incoherent or implausible results, suggesting gender specificity. The annotation log recorded ambiguities between Judgment categories (e.g., Normality-vs. Veracity-), indicating inherent boundary issues that may affect reproducibility across different analysts.

Second, the analysis revealed systematic conflation of moral and sexualized evaluation within Propriety-judgments, suggesting that Appraisal Theory's Social Sanction categories may be insufficient for gendered political discourse. This category strain, while theoretically interesting, introduces analytical uncertainty about whether observed patterns reflect genuine evaluative clustering or framework limitations. The frequent co-occurrence of multiple negative judgments (46.4 % of replies) further complicates attribution of primary evaluative focus.

Another constraint is that the 11 impoliteness strategies were identified inductively and assigned manually by the lead author across all 498 replies without inter-rater validation. This post-hoc typology, while grounded in corpus patterns, lacks independent verification and may reflect interpretive framing more than reproducible discourse features. The functional rather than face-based classification also departs from established impoliteness frameworks, limiting direct comparison with existing research.

Additionally, the April–May 2025 collection period represents routine political discourse rather than crisis or electoral contexts when evaluative patterns might intensify or shift. Ayuso's specific political activities, media coverage, or policy controversies during this period may have shaped the observed evaluation patterns in ways not captured by the analysis. The absence of temporal segmentation also obscures whether evaluative patterns remained stable or evolved over the five-week collection period.

A related limitation is that the study's focus on X/Twitter's character-limited, public reply format shapes both the types of evaluation possible and their multimodal realization. Spanish-language discourse patterns, cultural context around political criticism, and platform-specific norms for emoji use may limit transferability to other linguistic, cultural, or technological contexts. The manual bot exclusion process, while systematic, may have failed to identify sophisticated automated accounts potentially skewing evaluative patterns.

Finally, while the single-case design enables intensive analysis appropriate for discourse research, it prevents distinguishing Ayuso-specific patterns (relating to her conservative ideology, Madrid leadership, media persona, or communication style) from gender-specific evaluation tendencies. The "revelatory case" framing assumes that patterns observed here illuminate broader phenomena, but this assumption remains a hypothesis pending cross-case validation.

These limitations suggest several productive research directions. Comparative analysis across politicians varying in gender, ideology, and political role would isolate gender-specific from other targeting patterns. Multi-platform studies could reveal how different technological affordances shape evaluative discourse. Longitudinal analysis during electoral or crisis periods would capture evaluation pattern dynamics. Crosslinguistic replication would test cultural specificity of observed patterns. Finally, developing semi-automated approaches to Appraisal annotation could address scalability while maintaining interpretive nuance, enabling larger-scale analysis of evaluative bias in digital political discourse.

Despite these constraints, this single-case design offers foundational methodological scaffolding for examining how evaluative hostility operates in digital political discourse, while revealing the challenges of applying Appraisal Theory in multimodal, gendered, and politically volatile environments.

5.9. Conclusion

The systematic deployment of moral evaluation, intensified through multimodal mechanisms and filtered through gendered stereotypes, reveals how digital political discourse can weaponize assessment itself as a tool of exclusion. These replies are not merely impolite—they construct evaluative regimes that systematically challenge female political authority through moral rather than policy-based frameworks.

Understanding these patterns is crucial for protecting democratic inclusivity in digital environments where evaluation becomes a primary mode of political engagement. The findings suggest that contemporary political hostility operates not through overt exclusion but through systematic deployment of seemingly neutral assessment resources that encode cultural biases about political participation and leadership legitimacy.

CRediT authorship contribution statement

Sergei Sikorskii: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **María Luisa Carrió Pastor:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Funding

This study was funded by the Generalitat Valenciana under Project CIAICO/2023/187, “Elaboración de una base de datos multimodal y plurilingüe para la identificación y clasificación de los mecanismos de valoración hacia mujeres y hombres en TikTok, Instagram y X (Base de datos valoración MyH REDES)”, through the Conselleria d’Innovació, Universitats, Ciència i Societat Digital.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Sergei Sikorskii reports financial support was provided by Universitat Politècnica de València (Spain). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Baviera, T., Calvo, D., Llorca-Abad, G., 2019. Mediatisation in Twitter: an exploratory analysis of the 2015 Spanish general election. *J. Int. Commun.* 25 (2), 275–300. <https://doi.org/10.1080/13216597.2019.1634619>.
- Bousfield, D., Locher, M.A. (Eds.), 2008. *Impoliteness in Language: Studies on Its Interplay with Power in Theory and Practice*. Mouton de Gruyter. <https://doi.org/10.1515/9783110208344>.
- Camargo-Fernández, L., Polo-Artal, A., 2024. Representation of women in the digital discourse of Spanish far-right female leaders. *Discourse Commun.* 18 (1), 28–50. <https://doi.org/10.1177/17504813231211985>.
- Carrió-Pastor, M.L., 2025. A functional classification of aggressive digital replies to Pedro Sánchez’ posts on X. *J. Pragmat.* 245, 19–34. <https://doi.org/10.1016/j.pragma.2025.06.003>.
- Cavasso, L., Taboada, M., 2021. A corpus analysis of online news comments using the appraisal framework. *J. Corpora Discourse Stud.* 4 (0). <https://doi.org/10.18573/jcads.61>. Article 0.
- Cohen, J., 2013. *Statistical Power Analysis for the Behavioral Sciences*, 0 ed. Routledge. <https://doi.org/10.4324/9780203771587>.
- Culpeper, J., 2011. *Impoliteness: Using Language to Cause Offence*. Cambridge University Press.
- Culpeper, J., 2021. Impoliteness and hate speech: compare and contrast. *J. Pragmat.* 179, 4–11. <https://doi.org/10.1016/j.pragma.2021.04.019>.
- Culpeper, J., 1996. Towards an anatomy of impoliteness. *Journal of Pragmatics* 25 (3), 349–367. [https://doi.org/10.1016/0378-2166\(95\)00014-3](https://doi.org/10.1016/0378-2166(95)00014-3).
- Eagly, A.H., Karau, S.J., 2002. Role congruity theory of prejudice toward female leaders. *Psychol. Rev.* 109 (3), 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>.
- Fairclough, N., 2013. *Critical Discourse Analysis: the Critical Study of Language*, second ed. Routledge.
- Gamir-Ríos, J., Pop, A.-I., López-García, G., Llorca-Abad, G., Fenoll, V., 2022. Unidireccionalidad, promoción partidista y polarización. Actividad en Twitter de los candidatos a la presidencia del Gobierno de España en las elecciones generales de 2019. *El Prof. Inf.* <https://doi.org/10.3145/epi.2022.mar.03>.
- Guerrero-Solé, F., 2022. The ideology of media. Measuring the political leaning of Spanish news media through Twitter users’ interactions. *Commun. Soc.* 35 (1), 29–43. <https://doi.org/10.15581/003.35.1.29-43>.
- Hansson, S., Page, R., Fuoli, M., 2022. Discursive strategies of blaming: the language of judgment and political protest online. *Soc. Media Soc.* 8 (4). <https://doi.org/10.1177/20563051221138753>.
- Iranzo-Cabrera, M., Castro-Bleda, M.J., Simón-Astudillo, I., Hurtado, L.-F., 2025. Journalists’ ethical responsibility: tackling hate speech against women politicians in social media through natural language processing techniques. *Soc. Sci. Comput. Rev.* 43 (3), 475–502. <https://doi.org/10.1177/08944393241269417>.
- Koch, L., Russo Riva, M.P., Steinert, J.I., 2025. Technology-facilitated gender-based violence against politically active women: a systematic review of psychological and political consequences and women’s coping behaviors. *Trauma Violence Abuse.* <https://doi.org/10.1177/15248380251343185>.
- Makki, M., Zappavigna, M., 2022. Out-grouping and ambient affiliation in Donald Trump’s tweets about Iran: exploring the role of negative evaluation in enacting solidarity. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPra)* 32 (1), 104–130. <https://doi.org/10.1075/prag.20048.mak>.
- Martin, J.R., White, P.R.R., 2005. *The Language of Evaluation: Appraisal in English* (1. Publ). Palgrave Macmillan.
- Musi, E., Federico, L., Mounim, A., 2025. Unveiling Toxic Discourse Towards Female Politicians in Italy: a Corpus-based Approach. *European Journal of Women’s Studies.* <https://doi.org/10.1177/13505068251352636>.
- Nadim, M., Fladmoe, A., 2021. Silencing women? Gender and online harassment. *Soc. Sci. Comput. Rev.* 39 (2), 245–258. <https://doi.org/10.1177/0894439319865518>.
- Oz, M., Zheng, P., Chen, G.M., 2018. Twitter versus Facebook: comparing incivility, impoliteness, and deliberative attributes. *New Media Soc.* 20 (9), 3400–3419. <https://doi.org/10.1177/1461444817749516>.
- Rega, R., Marchetti, R., Stanziano, A., 2023. Incivility in online discussion: an examination of impolite and intolerant comments. *Soc. Media Soc.* 9 (2). <https://doi.org/10.1177/20563051231180638>.
- Reich, S., Bachl, M., 2025. When sexism becomes the norm: the effect of sexism on women’s participation in political online discussions. *Commun. Res.* <https://doi.org/10.1177/00936502251343988>.
- Rheault, L., Rayment, E., Musulan, A., 2019. Politicians in the line of fire: incivility and the treatment of women on social media. *Res. Politics* 6 (1). <https://doi.org/10.1177/2053168018816228>.
- Rohrbach, T., Aaldering, L., Van Der Pas, D.J., 2023. Gender differences and similarities in news media effects on political candidate evaluations: a meta-analysis. *J. Commun.* 73 (2), 101–112. <https://doi.org/10.1093/joc/jqac042>.
- Ross, A.S., Caldwell, D., 2020. ‘Going negative’: an APPRAISAL analysis of the rhetoric of Donald Trump on Twitter. *Lang. Commun.* 70, 13–27. <https://doi.org/10.1016/j.langcom.2019.09.003>.
- Rossini, P., 2022. Beyond incivility: understanding patterns of uncivil and intolerant discourse in online political talk. *Commun. Res.* 49 (3), 399–425. <https://doi.org/10.1177/0093650220921314>.
- Sánchez-Junquera, J., Ponzetto, S.P., Rosso, P., 2020. A Twitter political corpus of the 2019 10N Spanish election. In: *Lecture Notes in Computer Science*, vol 12284. Springer International Publishing, pp. 41–49. https://doi.org/10.1007/978-3-030-58323-1_4.

- Saz Rubio, M.M., 2024. An approach to the rhetoric of boebert and díaz-ayuso's political tweets on X: a corpus-assisted analysis. *Rev. Alicant. Estud. Ingleses* 41, 75–103. <https://doi.org/10.14198/raei.2024.41.04>.
- Saz-Rubio, M.M.D., 2023. Assessing impoliteness-related language in response to a season's greeting posted by the Spanish and English Prime Ministers on Twitter. *J. Pragmat.* 206, 31–55. <https://doi.org/10.1016/j.pragma.2023.01.010>.
- Southern, R., Harmer, E., 2021. Twitter, incivility and “Everyday” gendered othering: an analysis of tweets sent to UK members of parliament. *Soc. Sci. Comput. Rev.* 39 (2), 259–275. <https://doi.org/10.1177/0894439319865519>.
- Terkourafi, M., 2015. Conventionalization: a new agenda for im/politeness research. *J. Pragmat.* 86, 11–18. <https://doi.org/10.1016/j.pragma.2015.06.004>.
- Torregrosa, J., D'Antonio-Maceiras, S., Villar-Rodríguez, G., Hussain, A., Cambria, E., Camacho, D., 2023. A mixed approach for aggressive political discourse analysis on Twitter. *Cognit. Comput.* 15 (2), 440–465. <https://doi.org/10.1007/s12559-022-10048-w>.
- Van Leeuwen, T., 2007. Legitimation in discourse and communication. *Discourse Commun.* 1 (1), 91–112. <https://doi.org/10.1177/1750481307071986>.
- Vladimirou, D., Fernández-Amaya, L., 2024. ‘Shut up and resign’: a contrastive investigation of gender-based online aggression on Twitter in Spanish vs. Greek. *Contrastive Pragmat.* 6 (1), 51–89. <https://doi.org/10.1163/26660393-bja10132>.
- Yin, R.K., 2018. *Case Study Research and Applications: Design and Methods*, sixth ed. SAGE.
- Zappavigna, M., 2013. *Discourse of Twitter and Social Media: How we Use Language to Create Affiliation on the Web*. Bloomsbury Academic.
- Zappavigna, M., 2022. Social media quotation practices and ambient affiliation: weaponising ironic quotation for humorous ridicule in political discourse. *J. Pragmat.* 191, 98–112. <https://doi.org/10.1016/j.pragma.2021.12.003>.

Sergei Sikorskii is a senior research specialist at the Department of Applied Linguistics, Universitat Politècnica de València (Spain), where he is completing his doctoral research in the PhD programme *Languages, Literature, Culture and Their Applications*. He is a member of the GALE research group (Grupo de Análisis de las Lenguas de Especialidad) and focuses on political discourse, impoliteness, online aggressive language, language of hatred, and evaluative language on social media platforms. His research integrates Appraisal Theory, multimodal discourse analysis, and corpus linguistics. He has presented at national and international conferences and recently completed a research stay at the University of Bergamo. His professional profile is available at <https://www.upv.es/ficha-personal/ssikors> and ORCID: 0009-0009-1874-4394

María Luisa Carrió-Pastor is a professor of English language at Universitat Politècnica de Valencia, Spain. Currently, she is the coordinator of the Doctoral degree “Languages, Literature, Culture and, their Applications”, the President of the Spanish Association for Applied Linguistics (AESLA) and a member-at-large of AILA Executive Board. Her research areas are contrastive linguistics, pragmatics and the study of academic and professional discourse both for second language acquisition and discourse analysis. She has been in charge of the projects “Identification and Analysis of Metadiscourse Strategies in Research Articles in English and Spanish”, funded by the Spanish Ministry of Economy (Reference FFI2016-77941-P) and “Creation of a database, identification and classification of aggressive language on Twitter in English, French and Spanish” (AGRETUITEFI). Some of her publications are available at <http://www.upv.es/ficha-personal/lcarrio>, Research Gate, Academia and ORCID <https://orcid.org/0000-0002-3040-5362>.