

“Nosotros” vs “Ellos”: las estrategias discursivas de othering en la comunicación política de la extrema derecha

Nombre y apellidos:	Francisco José Segado Boj
Institución de afiliación:	Universidad Complutense de Madrid
Email:	fsegado@ucm.es
Doctor en Periodismo por la UCM (2008) y Profesor Titular en el Departamento de Periodismo y Comunicación Global de la Facultad de Ciencias de la Información de la UCM. Sus líneas de investigación se centran en comunicación digital, redes sociales, meta-investigación y discursos de odio. Ha publicado más de 70 artículos científicos y cuenta con experiencia en encuestas, análisis de contenido y métodos bibliométricos.	

Nombre y apellidos:	Marco Antonio Ruiz-Pérez
Institución de afiliación:	Universidad Complutense de Madrid
Email:	marcoantoniorp@ucm.es
Personal investigador en formación y doctorando en Periodismo en la UCM. Máster en Comunicación, Cultura, Sociedad y Política por la UNED y graduado en Derecho y en Ciencias Políticas y de la Administración por la UGR. Fue becario de formación para posgraduados del Centro de Investigaciones Sociológicas en 2024. Sus líneas de investigación se centran en comunicación política, campañas y debates electorales, medios de comunicación, discurso de odio y extrema derecha.	

Nombre y apellidos:	Javier García-Marín
Institución de afiliación:	Universidad de Granada
Email:	jgmarin@ugr.es
Profesor Titular en el Departamento de Ciencia Política y de la Administración de la UGR. Doctor en Ciencia Política, licenciado en Historia Contemporánea y máster en Economía y Relaciones Internacionales. Ha sido profesor visitante en la London School of Economics, la Universidad de Glasgow, la Universidad de Ámsterdam, Sciences Po Bordeaux y la Universidad de Buenos Aires. Su investigación se centra en comunicación política y en la aplicación de técnicas de aprendizaje automático al análisis de textos y contenidos audiovisuales.	

Resumen

El auge de la extrema derecha coincide con procesos de erosión democrática y con infraestructuras digitales que amplifican la circulación de discursos hostiles. Sin embargo, la investigación automatizada suele priorizar el odio explícito y es menos eficaz para detectar estrategias implícitas como el *othering* —la construcción de un “nosotros” frente a un “ellos”—. Este estudio analiza cómo opera el *othering* en el ecosistema español de Telegram vinculado a la extrema derecha, cómo se intersecta con narrativas anti-élite y anti-minoría, y cómo se relaciona con la toxicidad de los mensajes. A partir de 4.113.119 mensajes textuales procedentes de 66 canales, desarrollamos y validamos un enfoque automatizado reproducible basado en cinco clasificadores supervisados: uno para *othering* y cuatro para familias de narrativas hostiles. Los resultados indican que los mensajes que activan cualquier narrativa “anti-” tienen una probabilidad sustancialmente mayor de contener *othering*, con la asociación más intensa en los marcos anti-establishment. Asimismo, la presencia de *othering* incrementa la probabilidad de un tono agresivo, aunque este efecto varía según la narrativa y es más pronunciado en mensajes dirigidos contra la inmigración y contra grupos definidos por religión, etnia o cultura.

Palabras clave: Discurso de odio - Extrema derecha - Othering - Telegram - Toxicidad - Discurso tóxico

Introducción

El reciente auge de la extrema derecha coincide con un proceso de erosión democrática. Episodios como el asalto al Capitolio de Estados Unidos en 2021 o las campañas violentas dirigidas contra migrantes, colectivos LGBTIQ+ y mujeres en distintos países europeos son algunas de las manifestaciones de esta tendencia. La extrema derecha ejerce una presión sostenida sobre los principios de la democracia liberal, especialmente sobre el pluralismo, los derechos de las minorías y la inclusión (Dowling, 2023). Bennett y Kneuer (2024) vinculan esta trayectoria con los entornos digitales de comunicación, que la extrema derecha ha sabido aprovechar para organizarse y movilizarse con mayor eficacia que otros actores políticos (Bennett y Livingston, 2025).

La aparición de las denominadas plataformas *alt-tech* (Donovan et al., 2018), entre ellas Gab, Minds, Telegram, Parler o BitChute, ejemplifica esta dinámica. Estas plataformas se han convertido en posibles espacios de refugio para la *alt-right* tras sucesivas oleadas de *deplatforming*, entendido como la suspensión o eliminación de cuentas de las principales redes sociales por incumplir sus términos de servicio (Rogers, 2020). Rogers (2020) llega a caracterizar Telegram como el «rincón oscuro de Internet». En esta plataforma, la extrema derecha ha desarrollado una infraestructura comunicativa propia (Urman y Katz, 2020), que funciona como un importante nodo para la producción y difusión de contenidos de odio (Al-Rawi, 2021; Stewart et al., 2023).

La literatura académica ha identificado de manera consistente una estrecha relación entre la extrema derecha y el discurso de odio. El odio se ha conceptualizado tanto como parte de un proceso gradual de radicalización como uno de los posibles resultados sociales entre los simpatizantes de la extrema derecha, hasta el punto de considerarse una característica definitoria de este actor político (Marwick et al., 2022). Sus estrategias comunicativas se orientan principalmente hacia la construcción de enemigos, entre ellos quienes son representados como enemigos de la nación, de los hombres o de la raza blanca (Cervi et al., 2023). Este proceso de construcción del enemigo o *enemization* (Salvatore et al., 2018) guarda una estrecha relación, en el plano discursivo, con las estrategias de *othering*, articuladas en torno a una oposición binaria entre «nosotros» y «ellos», la evaluación moral y la estereotipación de grupos subordinados (García-Marín et al., 2025). Alorainy et al. (2018) muestran, además, que el *othering* constituye un potente marcador lingüístico del odio, especialmente de sus formas implícitas.

La investigación reciente sobre las estrategias comunicativas de la extrema derecha se ha centrado principalmente en la agenda temática y en los contenidos de su discurso, prestando especial atención al uso estratégico de emociones como el miedo o el resentimiento (Wodak, 2021). Aunque existe una producción considerable sobre discurso tóxico y detección automatizada del discurso de odio, estos trabajos suelen concentrarse en manifestaciones explícitas y abiertas de hostilidad (Vidgen et al., 2021). En comparación, las investigaciones dedicadas específicamente a las estrategias de *othering* dentro del discurso de extrema derecha siguen siendo más limitadas (Assimakopoulos et al., 2017). Se trata de una carencia relevante, ya que el *othering* puede actuar como indicador de formas implícitas de odio y, por tanto, permite aproximarse a una «zona gris» que los métodos actuales de detección automatizada tienen dificultades para identificar cuando el discurso recurre a formas lingüísticas matizadas, indirectas o codificadas (Young, 2025).

Este estudio analiza cómo se articulan las estrategias de *othering* en el discurso político de extrema derecha y de qué manera se insertan en distintas familias narrativas, definidas tanto por su contenido temático como por sus expresiones de hostilidad. La investigación persigue tres objetivos. En primer lugar, identificar qué familias narrativas se asocian con fronteras simbólicas más intensas. En segundo lugar, analizar la relación entre el *othering* y las expresiones de hostilidad, así como las variaciones de esta relación entre las distintas familias narrativas. Por último, desarrollar y validar un procedimiento automatizado y reproducible para detectar tanto el *othering* como dichas familias narrativas. El análisis empírico se basa en más de cuatro millones de mensajes de texto procedentes de 66 canales de Telegram pertenecientes al ecosistema español de extrema derecha.

Extrema derecha y discurso de extrema derecha

Existe un debate doctrinal sostenido en torno a las categorías utilizadas para denominar el espacio político situado más allá de la derecha clásica (Pirro, 2023). Partiendo de la distinción propuesta por Mudde (2019) entre derecha radical —iliberal, pero democrática— y extrema derecha —antidemocrática—, en este trabajo se adopta el concepto paraguas de *far right* o extrema derecha, dado que la muestra de Telegram incluye actores y repertorios heterogéneos. Esta decisión coincide asimismo con el planteamiento de Pirro (2023), quien sostiene que los crecientes vínculos entre actores iliberales-democráticos y antidemocráticos dificultan cada vez más el establecimiento de categorías estrictamente delimitadas.

Los estudios sobre el universo temático de la extrema derecha apuntan a la existencia de un repertorio narrativo amplio, pero ampliamente compartido. Este incluye narrativas conspirativas relacionadas con el coronavirus, la teoría de la conspiración QAnon, el racismo y la victimización blanca, los discursos *anti-woke* y las narrativas antiestablishment (Willaert et al., 2022). Desde una perspectiva operacional, Curley et al. (2022) identifican dos grandes familias narrativas: por un lado, las relacionadas con las protestas contra las medidas adoptadas durante la COVID-19 —aislamiento, antivacunación y narrativas conspirativas— y, por otro, otras narrativas características de la extrema derecha, como las vinculadas al reemplazo poblacional, el racismo, los marcos legales y los discursos antiélite. La evidencia más reciente indica que cuestiones que durante mucho tiempo habían permanecido en un segundo plano han pasado a convertirse en ejes centrales del discurso de extrema derecha.

Las narrativas antiinmigración aparecen de forma recurrente en casos como el de Donald Trump en Estados Unidos (Edilashvili, 2025), Giorgia Meloni en Italia (Indelicato y Magalhães Lopes, 2024) o Viktor Orbán en Hungría (Bastos, 2020). Con menor frecuencia, aunque también presentes en determinados casos —como Rassemblement National en Francia o la Lega en Italia—, aparecen las críticas al movimiento feminista y a los colectivos LGBTQI+ (Scrinzi, 2025). Sin embargo, uno de los temas más recurrentes es la crítica a las élites políticas nacionales e internacionales, a los partidos tradicionales en general y a la izquierda en particular (Winberg, 2017; Vaughan y Heft, 2023; Pérez-Díaz y Arroyas Langa, 2025). Las teorías de la conspiración suelen aparecer asociadas a otros temas, especialmente a marcos antiestablishment (DiMaggio, 2022). También se vinculan con frecuencia a la inmigración, como ocurre con la teoría conspirativa del Gran Reemplazo (Svatoňová y Doerr, 2024). Una lógica similar se observa en el revisionismo histórico y el

ultranacionalismo, que suelen funcionar como marcos interpretativos o vehículos ideológicos de las narrativas anteriormente descritas. En consecuencia, a efectos de esta investigación, se definen como principales ejes discursivos de la extrema derecha global las narrativas antiélite —dirigidas contra el establishment, los gobiernos y los grupos liberales, progresistas o de izquierda— y las narrativas contra las minorías —dirigidas contra la migración, la religión, la cultura, la etnicidad, el feminismo, el género y los colectivos LGBTQI+—.

***Othering*: «nosotros» frente a «ellos»**

Como se ha señalado, las estrategias de *othering* permiten captar los procesos de construcción del otro como enemigo (*enemization*) (Salvatore et al., 2018), en la medida en que constituyen uno de los principales mecanismos discursivos de delimitación de fronteras identitarias. El concepto procede de los estudios poscoloniales y de la sociología del racismo y hace referencia a los procesos de diferenciación y demarcación mediante los cuales se establece una frontera entre «nosotros» y «ellos» (Van Dijk, 1993). La literatura muestra que los objetivos más frecuentes del *othering*, especialmente en el discurso de extrema derecha, son las mujeres (Southern y Harmer, 2019), los migrantes y las minorías étnicas o religiosas (Molla, 2024) y los colectivos LGBTQI+ (Nartey, 2021). Las élites políticas constituyen también un objetivo recurrente (Resende y Freire, 2025), especialmente cuando se trata de élites vinculadas a la izquierda (Machuy, 2025). Por ello, cabe esperar que la presencia de narrativas dirigidas contra minorías y élites esté asociada con patrones de *othering* en los siguientes términos:

H1a. Los mensajes que incluyen al menos una narrativa contra las minorías o las élites presentan una mayor probabilidad de recurrir a estrategias de *othering* que aquellos que no contienen ninguna de estas narrativas.

$$\text{Any(ANTI)} \Rightarrow +P(\text{OTHERING})$$

H1b. A medida que aumenta el número de narrativas contra las minorías o las élites presentes en un mensaje, también aumenta la probabilidad de que este contenga *othering*.

$$\sum \text{ANTI} \Rightarrow \uparrow P(\text{OTHERING})$$

En consonancia con este planteamiento, la Teoría de la Identidad Social (Tajfel y Turner, 1979) proporciona un marco explicativo sólido al mostrar cómo los individuos construyen su autoconcepto mediante categorías grupales que distinguen entre un endogrupo valorado positivamente y un exogrupo devaluado. Este proceso cognitivo se materializa discursivamente a través del *othering*, que fija y refuerza la frontera simbólica entre «nosotros» y «ellos». La Teoría de la Amenaza Intergrupala (Stephan y Stephan, 2000) permite precisar este argumento al plantear que dichas fronteras se intensifican cuando el exogrupo es percibido como fuente de amenazas realistas o simbólicas. En estas circunstancias, el *othering* actúa como un mecanismo discursivo que materializa y amplifica esa frontera, al representar al exogrupo no solo como diferente o inferior, sino también como peligroso, amenazante o invasor. Estas percepciones pueden alimentar dinámicas de victimización colectiva (Vollhardt, 2012), mediante las cuales el endogrupo se percibe a sí

mismo como perjudicado, atacado o asediado, reforzando así su cohesión interna y contribuyendo a legitimar actitudes hostiles hacia el exogrupo.

En este punto resulta útil distinguir entre *othering* descendente (*othering downward*), cuando el grupo objeto del discurso ocupa una posición subordinada o vulnerable —como migrantes, minorías o personas pobres—, y *othering* ascendente (*othering upward*), cuando el objetivo es un grupo situado en una posición de poder, como las élites sociales, políticas y económicas o las propias instituciones (Novais y Christofolletti, 2025). Segado-Boj et al. (2026) sugieren que, en TikTok durante la campaña de las elecciones europeas de 2024, el odio dirigido contra grupos minoritarios se articula principalmente mediante estrategias de *othering*, mientras que, en el odio de carácter político, estas tienen un peso comparativamente menor frente a otras estrategias discursivas. Puesto que los diferentes objetivos del *othering* pueden construirse discursivamente de formas distintas, se propone la siguiente hipótesis:

H2. La asociación entre las distintas narrativas contra las minorías y las élites y la presencia de *othering* es heterogénea entre familias narrativas, aunque en todos los casos se espera una asociación positiva entre la presencia de estas narrativas y el *othering*.

$$\text{ANTI}_{\text{type}} \xrightarrow{\beta > 0} \text{OTHERING}$$

Toxicidad discursiva y *othering*

En investigaciones anteriores, la toxicidad discursiva se ha utilizado con frecuencia como una aproximación al discurso de odio. De hecho, Anjum y Katarya (2024) señalan que conceptos como discurso de odio, discurso tóxico, lenguaje abusivo, *flaming*, contenido agresivo y ciberacoso se emplean con frecuencia de manera solapada en la literatura. La toxicidad puede manifestarse mediante amenazas explícitas de daño físico, insultos directos o formas de comunicación condescendientes. Se caracteriza por patrones de incivilidad e intolerancia y por una clara orientación hacia la degradación simbólica o la agresión contra el interlocutor (Pradel et al., 2024). Zamri et al. (2023) sostienen que, cuando el *othering* intensifica la frontera entre «nosotros» y «ellos», la hostilidad y el odio pueden adquirir formas más virulentas.

Morales et al. (2025) observan, además, que la función social de la toxicidad varía en función del tipo de narrativa. Cuando una narrativa favorece la consolidación de fronteras grupales cohesionadas, el *othering* tiende a activar guiones discursivos que incrementan la probabilidad de niveles elevados de toxicidad. Cuando, por el contrario, la narrativa es objeto de disputa, los insultos pueden desempeñar funciones adicionales, de modo que el incremento de la toxicidad asociado al *othering* puede ser menor o presentar un patrón más irregular. Sobre esta base, se formula la siguiente expectativa acerca de la interacción entre las narrativas, el *othering* y la toxicidad discursiva:

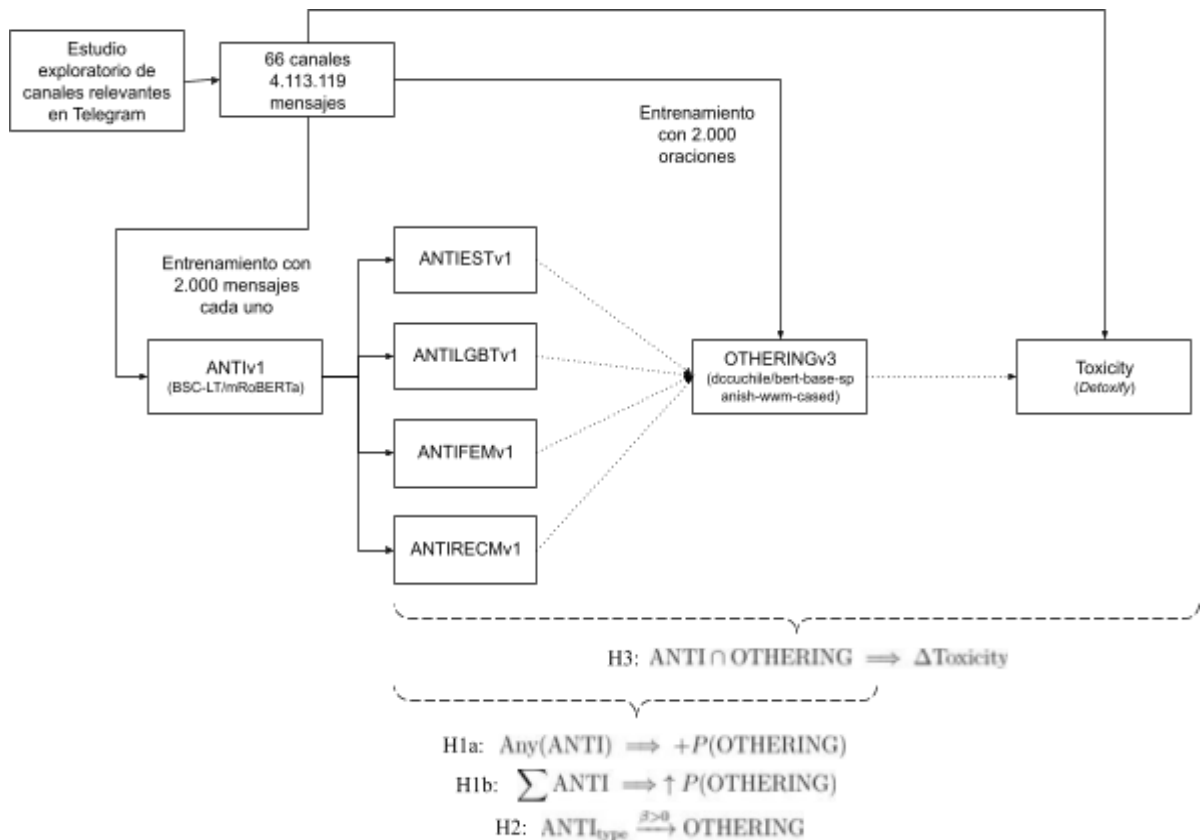
H3. Para cada tipo de narrativa contra las minorías o las élites, los mensajes que contienen *othering* presentan niveles de toxicidad significativamente superiores a los de mensajes equivalentes sin *othering*, aunque la magnitud de este incremento varía entre las distintas narrativas.

$$\text{ANTI} \cap \text{OTHERING} \Rightarrow \Delta \text{Toxicity}$$

Metodología

La Figura 1 sintetiza el diseño completo de la investigación. En una primera fase, el corpus analítico se construyó mediante un procedimiento de muestreo de canales de Telegram desarrollado en dos etapas. Inicialmente, se realizó un estudio exploratorio para identificar los canales más relevantes y visibles asociados con la extrema derecha dentro del ecosistema hispanohablante de Telegram. A partir de esta primera selección se aplicó un muestreo de bola de nieve, incorporando progresivamente nuevos canales a medida que aparecían mencionados o como origen de mensajes reenviados. La selección de los canales se basó en cuatro criterios: (a) contenido, exigiendo una vinculación observable con la extrema derecha o con entornos de discurso de odio; (b) impacto, definido mediante un mínimo aproximado de 1.000 suscriptores; (c) lengua y localización, restringiendo el corpus a canales en español con referencias sustantivas al contexto político, social y cultural español; y (d) exclusión temática, descartando aquellos canales centrados fundamentalmente en realidades latinoamericanas o con audiencias claramente situadas fuera del contexto español. Esta última valoración se realizó a partir de indicios presentes en el contenido de los mensajes, así como en la imagen de perfil, la descripción y el título de cada canal.

Figura 1. Diseño de la investigación



Fuente: Elaboración propia

A partir de los 66 canales identificados se extrajeron 4.960.330 mensajes mediante acceso programático a la API de Telegram utilizando la librería Telethon (Lonami, 2024) para Python. Los mensajes abarcan desde la creación de cada canal hasta el 15 de agosto de 2025. El conjunto de datos conserva únicamente metadatos a nivel de canal y características asociadas a los mensajes; no se almacenaron identificadores de usuarios. Los metadatos de los emisores se utilizaron exclusivamente para obtener recuentos descriptivos agregados y no se conservaron en el conjunto de datos final.

Se eliminaron los mensajes compuestos exclusivamente por enlaces, menciones a usuarios, *hashtags*, emojis o imágenes sin texto asociado. El corpus analítico final quedó compuesto por 4.113.119 mensajes de texto, sobre los que se realizó además una limpieza de enlaces, menciones, *hashtags* y emojis.

Para detectar manifestaciones lingüísticas de *othering* en el discurso digital en español se desarrolló un modelo supervisado de clasificación binaria entrenado mediante ajuste fino de BETO (dccuchile/bert-base-spanish-wwm-cased) (Cañete et al., 2020), un modelo BERT específicamente preentrenado para lengua española.

El modelo se entrenó sobre 2.000 oraciones anotadas manualmente, extraídas de mensajes pertenecientes al corpus, con una distribución equilibrada de 1.000 ejemplos positivos —presencia de *othering*— y 1.000 negativos. La utilización de la oración como unidad de entrenamiento permite captar la estructura semántica fundamental del *othering* —la construcción discursiva de una oposición entre «nosotros» y «ellos» (Van Dijk, 1993)— reduciendo el ruido contextual asociado a mensajes extensos.

Tabla 1. Ejemplos de oraciones anotadas para el entrenamiento del clasificador de *othering*

Etiqueta	Mensaje
<i>Othering</i>	«Los culpables son los Gobiernos, que nos mienten, nos roban, nos traicionan, nos venden y que además se ríen en nuestra cara...»
No <i>othering</i>	«Salvamento Marítimo “recoge” en El Hierro 4 cayucos con más de 500 “inmigrantes”.»

Para trasladar la clasificación al nivel de mensaje, los textos se segmentaron mediante spaCy (Montani et al., 2023) en oraciones o unidades lingüísticamente coherentes. Cuando un segmento superaba el límite de entrada del modelo, se subdividía mediante tokenización. Cada segmento se puntuó independientemente y la probabilidad final del mensaje se obtuvo mediante una regla disyuntiva, tomando el valor máximo entre sus segmentos. De este modo, un mensaje se consideraba *othering* cuando al menos una de sus partes presentaba una señal suficientemente intensa.

El umbral óptimo se estableció en 0,734 a partir de su contraste con dos codificaciones humanas independientes. En este punto de corte, el clasificador de *othering* alcanzó un acuerdo observado del 93,94 %, κ de Cohen = 0,754, F1 = 78,87 %, precisión = 89,36 % y *recall* = 70,59 %.

Para identificar las principales narrativas hostiles se desarrollaron cuatro clasificadores binarios: AntiEstablishment, AntiFeminism, AntiLGBTIQ+ y AntiRelEthCulMig. El clasificador AntiEstablishment identifica discursos hostiles dirigidos contra el *establishment*, las instituciones y las élites políticas, particularmente aquellas asociadas con la izquierda o el globalismo. AntiLGBTIQ+ y AntiFeminism recogen dos grupos objetivo diferenciados dentro de las narrativas dirigidas contra minorías o colectivos sociales. Por su parte, AntiRelEthCulMig integra expresiones hostiles dirigidas contra grupos definidos por su religión, origen étnico o racial, cultura o condición migratoria, dimensiones que aparecen frecuentemente entrelazadas en el corpus.

A diferencia del clasificador de othering, estos clasificadores se entrenaron directamente utilizando el mensaje como unidad de análisis, dado el carácter más contextual y temático de las familias narrativas. Los cuatro modelos fueron desarrollados mediante ajuste fino de BSC-LT/mRoBERTa (BSC-LT, 2025), empleando subconjuntos equilibrados de 2.000 mensajes anotados manualmente para cada tarea —1.000 positivos y 1.000 negativos—.

La utilización de arquitecturas diferentes para el *othering* y las familias narrativas constituye una precaución metodológica adicional. Mientras que el clasificador de othering se basa en BETO, los clasificadores narrativos se estiman mediante mRoBERTa, reduciendo así el riesgo de que las asociaciones observadas entre ambos constructos respondan exclusivamente a una representación lingüística compartida por los modelos.

Tabla 2. Ejemplos de mensajes anotados para los clasificadores de familias narrativas

Etiqueta	Mensaje
AntiEstablishment	«El régimen del 78 está podrido. Se basa en la corrupción, en el clientelismo y en la manipulación del voto.»
AntiLGBTIQ+	«El enemigo es el LGTB. Han hecho de una aberración la normalidad. Mira qué asco de dibujos animados para los niños. Para freírles el tarro con el mariconeo.»
AntiFeminism	«Feminazis aprovechan el 8M para vandalizar varias iglesias en Sevilla.»
AntiRelEthCulMig	«La inmigración ilegal no para de crecer como consecuencia de las políticas de fronteras abiertas y las jugosas ayudas sociales que ofrecen el Estado y las autonomías. Hay que frenar el efecto llamada terminando con esas ayudas y con deportaciones masivas YA.»

En su aplicación al corpus, los clasificadores se ejecutaron directamente a nivel de mensaje. Los textos que superaban el límite de entrada se dividieron respetando, siempre que fuera posible, los límites oracionales y las predicciones de los distintos segmentos se agregaron mediante una regla OR.

Para las cuatro familias narrativas se estableció un umbral operativo común de 0,881, calibrado frente a la codificación humana donde los cuatro clasificadores maximizaban su

F1. En este punto de corte, los niveles de acuerdo observado oscilaron entre el 89,46 % y el 95,90 %, con valores de κ de Cohen comprendidos entre 0,574 y 0,650. Los F1 fueron del 69,11 % para AntiEstablishment, 69,15 % para AntiRelEthCulMig, 62,56 % para AntiFeminism y 63,64 % para AntiLGBTIQ+.

Por su parte, la toxicidad se utiliza como una medida operacional de las formas más explícitamente agresivas del discurso. No se equipara, sin embargo, con el discurso de odio, ya que este puede adoptar también manifestaciones implícitas que no necesariamente presentan niveles elevados de toxicidad.

La variable se estimó mediante Detoxify (Hanu y Unitary team, 2020), utilizando su modelo multilingüe basado en *transformers*. Aunque la herramienta proporciona puntuaciones para distintas dimensiones —toxicidad, toxicidad severa, obscenidad, ataques identitarios, insultos, amenazas y contenido sexual explícito—, en este estudio se empleó exclusivamente la puntuación general de toxicidad, comprendida entre 0 y 1.

La puntuación se transformó en una variable dicotómica utilizando un umbral de 0,510, establecido mediante comparación con la codificación humana. En dicho punto de corte, el modelo alcanzó un acuerdo observado del 81,74 %, κ de Cohen = 0,490 y F1 = 59,38 %. La precisión fue elevada (86,18 %), aunque el *recall* resultó más limitado (45,30 %), lo que indica una clasificación relativamente conservadora de la toxicidad, es decir, los casos identificados positivamente presentan una elevada correspondencia con la referencia humana, aunque una parte de los mensajes considerados tóxicos por la codificación humana, no es recuperada por el modelo.

Resultados

La Tabla 3 ofrece una primera aproximación descriptiva a la relación entre las familias narrativas, el *othering* y la toxicidad. La diferencia más clara aparece al comparar los mensajes que contienen alguna de las narrativas analizadas con aquellos en los que no se identifica ninguna. El *othering* está presente en el 33,26 % de los mensajes que contienen al menos una narrativa, frente al 8,00 % de aquellos sin ninguna narrativa. Esta diferencia anticipa una asociación sustantiva entre la activación de marcos antagonistas y la construcción discursiva de fronteras entre un «nosotros» y un «ellos».

Tabla 3. Presencia de *othering* según familia narrativa y su intersección con toxicidad

Categoría	Othering	No othering	N
Any narrative	237.453 (33,26 %)	476.582 (66,74 %)	714.035
No narrative	272.091 (8,00 %)	3.126.993 (92,00 %)	3.399.084
AntiEstablishment	168.851 (47,11 %)	189.577 (52,89 %)	358.428
AntiFeminism	66.333 (23,16 %)	220.059 (76,84 %)	286.392
AntiRelEthCulMig	58.510 (32,01 %)	124.259 (67,99 %)	182.769

AntiLGBTIQ+	26.320 (31,66 %)	56.808 (68,34 %)	83.128
Toxic × AntiEstablishment	44.530 (59,16 %)	30.739 (40,84 %)	75.269
Toxic × AntiFeminism	18.629 (34,30 %)	35.684 (65,70 %)	54.313
Toxic × AntiRelEthCulMig	9.526 (63,18 %)	5.552 (36,82 %)	15.078
Toxic × AntiLGBTIQ+	8.793 (36,12 %)	15.548 (63,88 %)	24.341

Nota. Las categorías no son mutuamente excluyentes, salvo la comparación entre presencia y ausencia de narrativas.

También se observan diferencias relevantes entre familias. AntiEstablishment presenta descriptivamente la mayor prevalencia de *othering* (47,11 %), seguida de AntiRelEthCulMig (32,01 %), AntiLGBTIQ+ (31,66 %) y AntiFeminism (23,16 %). Las diferencias se acentúan al considerar conjuntamente narrativa y toxicidad. Entre los mensajes tóxicos de AntiRelEthCulMig, el 63,18 % contiene *othering*, porcentaje que alcanza el 59,16 % en AntiEstablishment. En cambio, las proporciones son menores en AntiLGBTIQ+ (36,12 %) y AntiFeminism (34,30 %). Aunque estos porcentajes son únicamente descriptivos y no permiten aislar los efectos de cada variable, anticipan que la relación entre *othering*, narrativa y toxicidad presenta una marcada heterogeneidad.

Tabla 4. Modelos de regresión logística para la predicción de *othering*

Modelo	Predictor	β	SE	Odds Ratio (OR)
M1 (H1a)	Any anti-narrative present (binary)	1,745***	0,003	5,726
M2 (H1b)	Number of anti-narratives (0–4)	1,101***	0,002	3,006
M3 (H2)	AntiEstablishment	2,096***	0,004	8,133
	AntiRelEthCulMig	0,988***	0,006	2,686
	AntiLGBTIQ+	0,552***	0,009	1,736
	AntiFeminism	0,297***	0,006	1,346

*** $p < 0,001$.

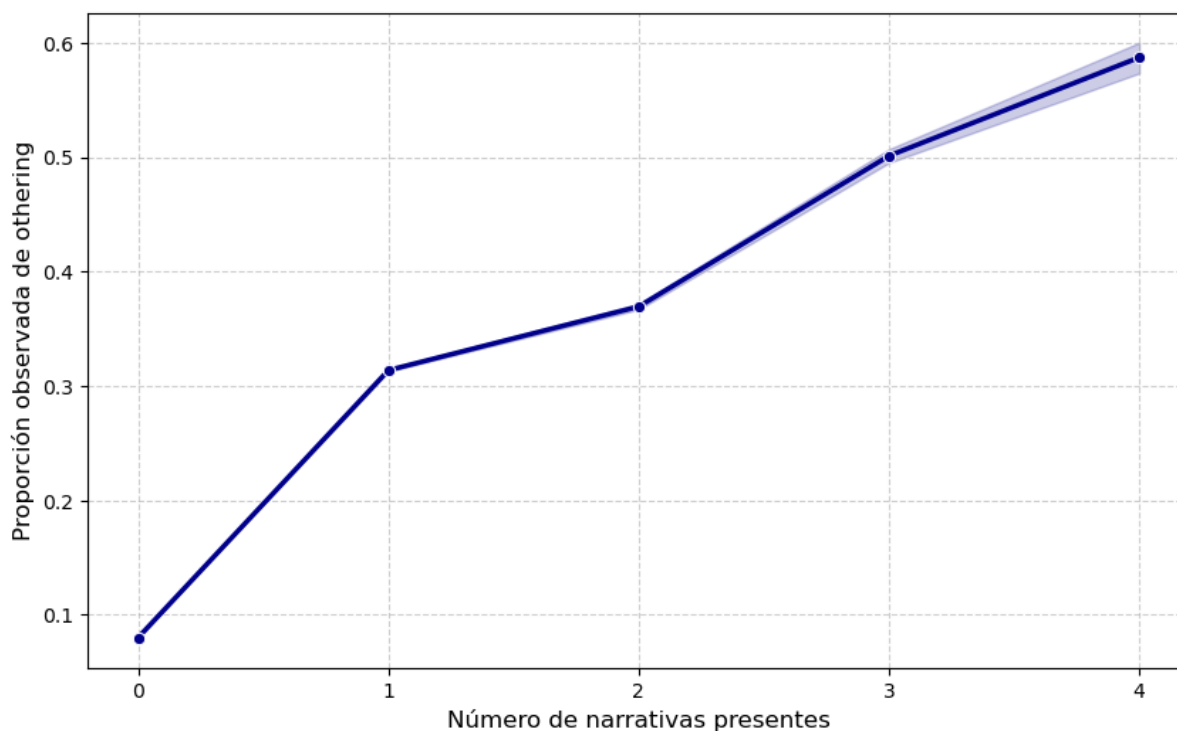
Nota. Debido al elevado tamaño muestral ($N = 4.113.119$), los errores estándar son muy reducidos y los intervalos de confianza del 95 % resultan correspondientemente estrechos; por concisión, no se muestran en la tabla.

Para determinar si la presencia de narrativas contra las minorías o las élites incrementa la probabilidad de que un mensaje contenga *othering*, se estimó un modelo de regresión logística binaria (M1, Tabla 4), en el que la variable independiente indica la presencia de al menos una de las cuatro familias narrativas. El coeficiente obtenido es positivo y estadísticamente significativo ($\beta = 1,745$; $p < 0,001$). En términos sustantivos, los mensajes

que contienen al menos una narrativa presentan unas *odds* de *othering* 5,73 veces superiores a las de aquellos en los que no se identifica ninguna (OR = 5,726).

Este resultado respalda H1a y muestra una asociación intensa entre la activación de narrativas antagonistas y la construcción de fronteras simbólicas mediante *othering*.

Figura 2. Prevalencia de *othering* según el número de narrativas



Fuente: Elaboración propia

La Figura 2 muestra un incremento de la prevalencia de *othering* conforme aumenta el número de narrativas concurrentes. El patrón es compatible con un mecanismo acumulativo: la coexistencia de varios marcos antagonistas en un mismo mensaje se relaciona con una demarcación más intensa entre un endogrupo legítimo y uno o varios exogrupos problematizados.

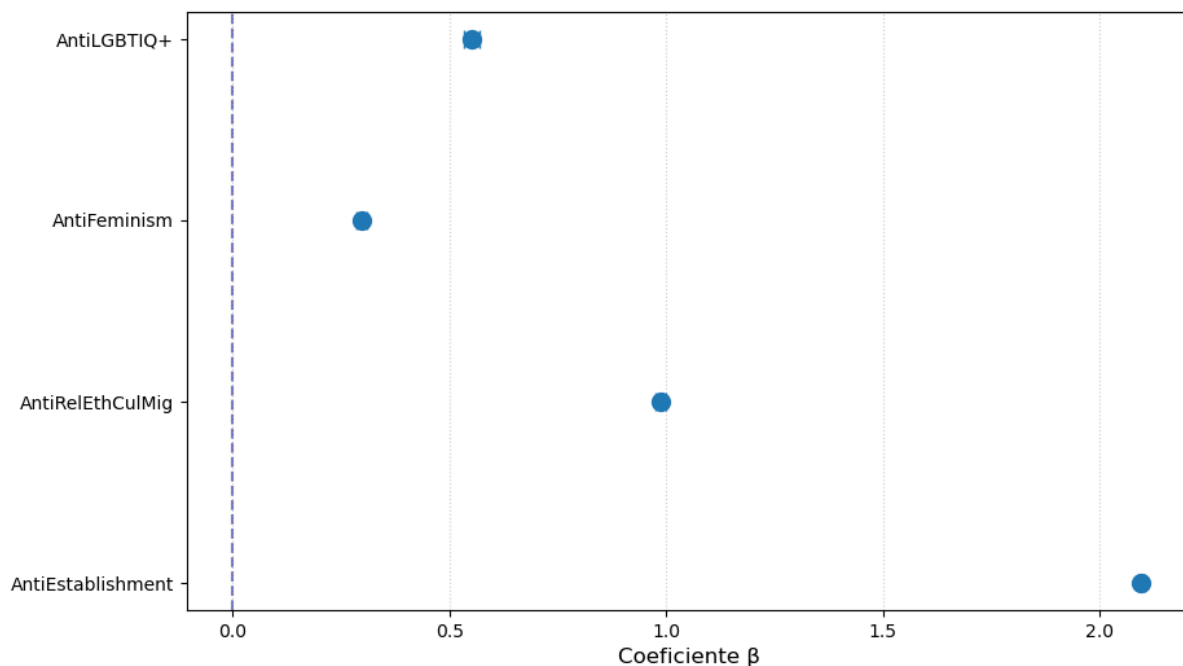
Para contrastar estadísticamente este patrón se estimó un segundo modelo logístico (M2, Tabla 4), sustituyendo el indicador dicotómico anterior por el número de narrativas presentes en cada mensaje, con valores comprendidos entre 0 y 4. El coeficiente es nuevamente positivo y significativo ($\beta = 1,101$; $p < 0,001$). El correspondiente OR = 3,006 indica que cada narrativa adicional se asocia con una multiplicación por aproximadamente tres de las *odds* de que el mensaje contenga *othering*. Los resultados respaldan, por tanto, H1b y muestran una relación incremental entre acumulación narrativa y construcción discursiva de alteridad.

Para evaluar H2, se estimó un modelo multivariante (M3, Tabla 4) que incorpora simultáneamente las cuatro familias narrativas. Esta especificación permite estimar la asociación específica de cada una con el *othering* manteniendo constantes las restantes

narrativas y, por tanto, controlar parcialmente su frecuente coocurrencia dentro de un mismo mensaje.

Los resultados muestran una clara heterogeneidad. Las cuatro familias presentan coeficientes positivos y estadísticamente significativos ($p < 0,001$), pero sus magnitudes difieren considerablemente. AntiEstablishment destaca como el predictor más intenso ($\beta = 2,096$; OR = 8,133): manteniendo constantes las demás narrativas, su presencia se asocia con unas *odds* de *othering* más de ocho veces superiores.

Figura 3. Efectos de cada familia narrativa sobre el *othering*



Fuente: Elaboración propia

Entre las narrativas dirigidas contra grupos sociales, AntiRelEthCulMig presenta la asociación más elevada ($\beta = 0,988$; OR = 2,686), seguida de AntiLGBTIQ+ ($\beta = 0,552$; OR = 1,736) y AntiFeminism ($\beta = 0,297$; OR = 1,346). La Figura 3 permite visualizar estas diferencias y confirma que la asociación entre narrativa y *othering* no es homogénea.

En conjunto, los resultados respaldan H2. Todas las familias narrativas se relacionan positivamente con el *othering*, pero la construcción del exogrupo aparece especialmente vinculada al eje antiélite representado por AntiEstablishment. Al mismo tiempo, AntiRelEthCulMig muestra una asociación sustantivamente mayor que las restantes narrativas dirigidas contra grupos sociales, lo que introduce una diferenciación relevante dentro de las propias narrativas antiminoritarias.

Finalmente, para analizar si la relación entre *othering* y toxicidad varía según la familia narrativa, se estimó un modelo logístico con toxicidad dicotómica como variable dependiente. El modelo incluye *othering*, las cuatro familias narrativas y sus correspondientes interacciones *othering* $\times$ *narrativa*. Esta especificación permite estimar

cómo cambia el efecto del *othering* sobre la toxicidad dependiendo del contexto narrativo en el que aparece.

Tabla 5. Regresión logística con interacciones para la predicción de toxicidad

Predictor	β	SE	Odds Ratio (OR)
Constant	-1,915***	0,002	0,147
Othering	0,943***	0,004	2,567
AntiEstablishment	0,271***	0,007	1,312
AntiRelEthCulMig	-1,341***	0,014	0,262
AntiFeminism	0,147***	0,007	1,158
AntiLGBTIQ+	0,960***	0,011	2,612
Othering × AntiEstablishment	-0,283***	0,009	0,754
Othering × AntiRelEthCulMig	0,602***	0,018	1,826
Othering × AntiFeminism	-0,089***	0,012	0,915
Othering × AntiLGBTIQ+	-0,520***	0,018	0,595

*** $p < 0,001$.

Nota. En los modelos con interacciones, los efectos principales son condicionales a la categoría de referencia de las variables con las que interactúan.

En ausencia de cualquiera de las cuatro narrativas, el coeficiente de *othering* es positivo y estadísticamente significativo ($\beta = 0,943$; OR = 2,567; $p < 0,001$). Esto significa que, en el escenario de referencia, los mensajes con *othering* presentan unas *odds* de toxicidad aproximadamente 2,57 veces superiores a las de mensajes equivalentes sin *othering*.

Sin embargo, esta relación varía significativamente según la familia narrativa. La interacción Othering × AntiRelEthCulMig es positiva ($\beta = 0,602$; OR = 1,826; $p < 0,001$), lo que indica que la asociación entre *othering* y toxicidad se intensifica respecto al escenario basal cuando aparece esta narrativa. Por el contrario, las interacciones con AntiEstablishment ($\beta = -0,283$; OR = 0,754), AntiFeminism ($\beta = -0,089$; OR = 0,915) y AntiLGBTIQ+ ($\beta = -0,520$; OR = 0,595) son negativas y significativas, indicando que el incremento de toxicidad asociado al *othering* es comparativamente menor en estos contextos que en la categoría de referencia.

Los coeficientes de interacción, no obstante, no representan por sí mismos el efecto del *othering* dentro de cada familia narrativa. Para facilitar su interpretación se calcularon sus efectos condicionales, sumando el coeficiente general de *othering* y el término de interacción correspondiente y transformando posteriormente el resultado en *odds ratios*.

Tabla 6. Efecto condicional del *othering* sobre la toxicidad según familia narrativa

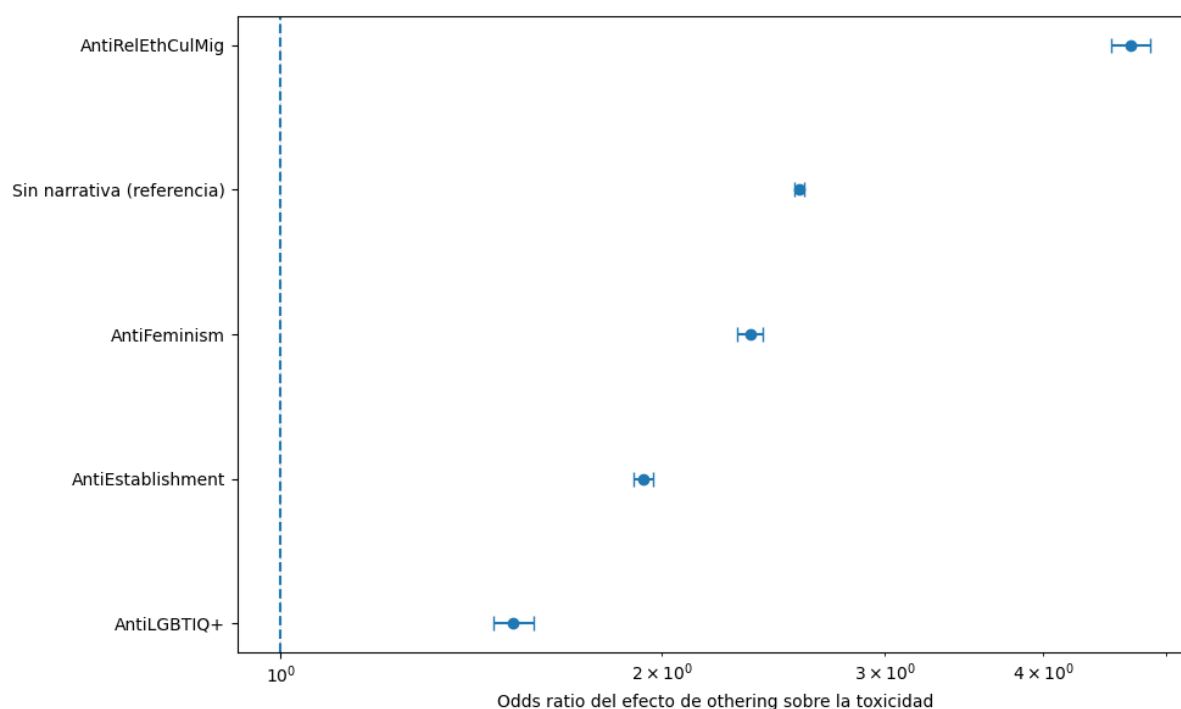
Condición	β (<i>othering</i>)	SE	OR (<i>othering</i>)
Sin narrativa (referencia)	0,943***	0,004	2,567
AntiEstablishment = 1	0,660***	0,009	1,934
AntiRelEthCulMig = 1	1,545***	0,018	4,688
AntiFeminism = 1	0,853***	0,012	2,347
AntiLGBTIQ+ = 1	0,423***	0,019	1,527

*** $p < 0,001$.

Los resultados muestran que el *othering* se asocia con una mayor toxicidad en todas las condiciones analizadas, aunque con magnitudes claramente diferentes. En el escenario de referencia, el *othering* multiplica las *odds* de toxicidad por 2,57. La asociación es algo menor dentro de AntiFeminism (OR = 2,347) y AntiEstablishment (OR = 1,934), y alcanza su magnitud más reducida en AntiLGBTIQ+ (OR = 1,527).

La principal diferencia aparece en AntiRelEthCulMig, donde la presencia de *othering* se asocia con unas *odds* de toxicidad 4,69 veces superiores a las de mensajes de la misma condición narrativa sin *othering*. Se trata, por tanto, de la única familia en la que el efecto del *othering* sobre la toxicidad se intensifica claramente respecto al escenario de referencia.

Figura 4. Efecto condicional del *othering* sobre la toxicidad según familia narrativa



Fuente: Elaboración propia

La Figura 4 permite apreciar esta heterogeneidad. Es importante subrayar que los términos de interacción negativos observados en AntiEstablishment, AntiFeminism y AntiLGBTIQ+ no implican que el *othering* reduzca la toxicidad. En las cuatro familias el efecto condicional permanece positivo y estadísticamente significativo. Las interacciones negativas indican únicamente que el incremento asociado a la presencia de *othering* es menor que en el escenario basal, manteniendo constantes las restantes variables.

Por tanto, en línea con H3, dentro de todas las familias narrativas analizadas, los mensajes que contienen *othering* presentan una asociación significativamente mayor con la toxicidad que aquellos sin *othering*, pero la magnitud de esta relación depende del contexto narrativo. La intensificación es especialmente pronunciada en las narrativas AntiRelEthCulMig, mientras que resulta más moderada en AntiFeminism, AntiEstablishment y, especialmente, AntiLGBTIQ+.

Discusión

Los resultados muestran que las narrativas antagonistas están estrechamente vinculadas con las estrategias de *othering*. Tanto la presencia de alguna narrativa contra minorías o élites como la acumulación de varias en un mismo mensaje incrementan sustantivamente la probabilidad de construir discursivamente una frontera entre «nosotros» y «ellos». Esto sugiere que estas narrativas no operan únicamente como evaluaciones negativas de determinados actores o grupos, sino también como mecanismos de demarcación simbólica. Al activar una lógica de conflicto, favorecen la identificación de responsables, amenazas y grupos enfrentados y, con ello, la construcción de un endogrupo legítimo frente a un exogrupo problematizado.

Desde esta perspectiva, el *othering* aparece como un recurso relevante para organizar discursivamente las atribuciones de responsabilidad, amenaza y pertenencia. Este patrón resulta coherente con las dinámicas de *framing* descritas por Entman (2003), en la medida en que la definición de problemas y responsables contribuye también a establecer distancia simbólica respecto de determinados actores. La relación acumulativa observada en H1b refuerza esta interpretación: la coexistencia de diferentes narrativas antagonistas parece ampliar las líneas de conflicto activadas dentro del mensaje y hacer más probable una delimitación explícita entre endogrupo y exogrupo.

La desagregación por familias narrativas revela, sin embargo, una jerarquía clara. AntiEstablishment presenta con diferencia la asociación más intensa con el *othering*, resultado coherente con la relevancia de la construcción de élites antagonistas en el discurso populista de extrema derecha (Resende y Freire, 2025). Este hallazgo contrasta parcialmente con investigaciones como la de Segado-Boj et al. (2026), que encuentran una presencia especialmente relevante del *othering* dirigido contra minorías durante la campaña de las elecciones europeas de 2024. Una posible explicación de esta diferencia se encuentra en el contexto comunicativo analizado: frente a una campaña electoral delimitada temporalmente, Telegram constituye aquí un espacio de comunicación sostenido en el que la oposición a instituciones, gobiernos, élites y al propio «sistema» puede adquirir un carácter más estructural.

Este patrón puede interpretarse también a partir de la distinción entre *upward* y *downward othering* (Novais y Christofolletti, 2025). La construcción de alteridad no se dirige únicamente hacia grupos socialmente subordinados, sino también hacia actores percibidos como poderosos, como gobiernos, instituciones o élites políticas. En el ecosistema analizado, estos últimos parecen constituir un eje particularmente fértil para la construcción de fronteras antagonistas. Esto sugiere que el *othering* de la extrema derecha no debe entenderse exclusivamente como un mecanismo dirigido contra minorías, sino también como una estrategia de deslegitimación de aquellos actores considerados responsables del orden político existente.

La posición de AntiRelEthCulMig como segunda familia más asociada con el *othering* matiza, no obstante, una interpretación exclusivamente política. Las narrativas relacionadas con migración, religión, etnicidad y cultura parecen especialmente proclives a representar determinados colectivos como un exogrupo diferenciado. Este resultado es coherente con la Teoría de la Amenaza Intergupal (Stephan y Stephan, 2000): las diferencias culturales, religiosas o nacionales pueden ser construidas como amenazas a la identidad, los valores o los recursos del endogrupo, intensificando así la frontera entre «nosotros» y «ellos». La menor asociación encontrada para AntiLGBTIQ+ y, especialmente, AntiFeminism muestra además que las propias narrativas dirigidas contra grupos sociales no funcionan de forma homogénea.

Los resultados de H3 profundizan en esta heterogeneidad. El *othering* se asocia con una mayor toxicidad en todas las familias narrativas, aunque la magnitud de esta relación depende significativamente del contexto. La asociación se intensifica especialmente en AntiRelEthCulMig, donde la presencia de *othering* se relaciona con el mayor incremento de toxicidad observado. En este caso, la construcción de colectivos religiosos, étnicos, culturales o migrantes como un exogrupo diferenciado parece estar más estrechamente conectada con formas explícitas de agresividad o degradación discursiva.

Por el contrario, en AntiEstablishment, AntiFeminism y AntiLGBTIQ+ el efecto del *othering* sobre la toxicidad continúa siendo positivo, pero resulta comparativamente menor que en el escenario de referencia. Esto no significa que estas narrativas sean menos hostiles, sino que su antagonismo puede adoptar formas que no necesariamente se traducen en un incremento equivalente de toxicidad léxica. En AntiEstablishment, por ejemplo, la oposición entre «nosotros» y «ellos» puede expresarse mediante deslegitimación institucional, acusaciones de corrupción, conspiración o traición, sin necesidad de recurrir a insultos o formas abiertamente agresivas.

Esta cuestión remite también a los límites de la medición de la toxicidad. Detoxify permite captar una dimensión relevante de agresividad discursiva explícita, pero no necesariamente todo el repertorio de hostilidad. Estrategias como la insinuación, el eufemismo, la ironía, la deslegitimación o determinadas construcciones aparentemente racionalizadas pueden reforzar intensamente el *othering* sin alcanzar puntuaciones elevadas de toxicidad. Por ello, los resultados de H3 deben interpretarse como evidencia sobre la relación entre *othering* y aquellas manifestaciones de agresividad que el modelo identifica con mayor claridad, y no como una estimación exhaustiva de todas las formas de hostilidad presentes en el discurso.

En conjunto, los resultados respaldan las tres hipótesis y muestran que el *othering* no constituye un mecanismo discursivo uniforme. Su presencia está estrechamente vinculada con las narrativas antagonistas, pero tanto su frecuencia como su relación con formas explícitas de agresividad dependen de la familia narrativa en la que se inserta. Esta heterogeneidad resulta especialmente visible en AntiRelEthCulMig, donde convergen una elevada asociación con el *othering* y la mayor intensificación de la toxicidad.

Conclusiones

Los resultados muestran que el *othering* desempeña un papel estructural en la organización del discurso hostil de extrema derecha en Telegram. Las narrativas antagonistas incrementan claramente su presencia y esta asociación se intensifica cuando varias de ellas concurren en un mismo mensaje. El *othering* puede entenderse así como un mecanismo de demarcación simbólica que no constituye necesariamente odio por sí mismo, pero que puede facilitar repertorios posteriores de deslegitimación, estigmatización o degradación del «otro».

Esta relación presenta, además, una jerarquía clara. AntiEstablishment constituye el principal eje de *othering* en el ecosistema analizado, lo que apunta a la importancia de la confrontación con instituciones y élites en la construcción identitaria de estos espacios. Estas élites pueden funcionar como un enemigo transversal al que se atribuye responsabilidad por otros procesos problematizados por la extrema derecha, como la inmigración, las políticas de igualdad o las agendas LGBTIQ+. No obstante, AntiRelEthCulMig ocupa también una posición especialmente relevante: es la segunda familia más asociada con el *othering* y aquella en la que su presencia produce el mayor incremento de toxicidad.

Los resultados permiten, por tanto, distinguir entre la intensidad de la construcción de fronteras simbólicas y la forma en que estas se traducen en agresividad explícita. Un discurso puede presentar una oposición «nosotros/ellos» muy marcada sin utilizar necesariamente un lenguaje altamente tóxico. Esta distinción resulta especialmente relevante para analizar formas de hostilidad que se presentan mediante argumentos aparentemente racionales, alusiones institucionales, eufemismos o insinuaciones.

Desde una perspectiva aplicada, la detección automatizada del *othering* puede complementar los sistemas centrados en insultos, amenazas u otras manifestaciones explícitas de agresividad. Su identificación puede ayudar a localizar espacios discursivos en los que comienzan a consolidarse fronteras identitarias y repertorios de exclusión incluso antes de que aparezcan expresiones abiertamente tóxicas. En este sentido, el estudio contribuye a aproximarse a una «zona gris» del discurso hostil que los sistemas tradicionales de detección pueden pasar por alto.

La principal limitación se encuentra precisamente en la operacionalización de la toxicidad. La menor sensibilidad de Detoxify frente a la codificación humana puede indicar que determinadas formas de hostilidad quedan fuera de la medición. Sin embargo, esta discrepancia también puede responder a una diferencia conceptual: los anotadores humanos pueden estar empleando una noción más amplia de toxicidad, incorporando formas generales de desprecio, degradación u hostilidad que no coinciden exactamente con

las dimensiones de agresividad lingüística que Detoxify está diseñado para identificar. Por ello, futuras investigaciones deberían diferenciar con mayor precisión entre toxicidad, incivilidad, hostilidad y discurso de odio, así como evaluar modelos específicamente adaptados al español y al discurso político.

Finalmente, sería necesario comprobar hasta qué punto estos patrones se reproducen en otros contextos. La extrema derecha reúne corrientes populistas, ultraconservadoras, identitarias y conspirativas que pueden organizar de manera distinta sus fronteras simbólicas. Los diseños comparativos permitirían determinar si la centralidad del eje AntiEstablishment y la especial relación entre AntiRelEthCulMig, *othering* y toxicidad constituyen rasgos específicos del ecosistema español de Telegram o patrones más generales del discurso contemporáneo de extrema derecha.

En términos generales, el estudio realiza una doble contribución. Sustantivamente, muestra que las narrativas antagonistas están estrechamente vinculadas con la construcción del «otro» y que esta relación varía según los grupos y actores hacia los que se dirige. Metodológicamente, propone una estrategia automatizada y reproducible para detectar mecanismos lingüísticos de demarcación simbólica que pueden preceder, acompañar o facilitar formas más explícitas de hostilidad.

Referencias

Alorainy, W., Burnap, P., Liu, H., & Williams, M. (2018). *The enemy among us: Detecting hate speech with threats based 'othering' language embeddings*. arXiv. <https://arxiv.org/abs/1801.07495>

Al-Rawi, A. (2021). Telegramming hate: Far-right themes on dark social media. *Canadian Journal of Communication*, 46(4), 821–851.

Anjum, & Katarya, R. (2024). Hate speech, toxicity detection in online social media: A recent survey of state of the art and opportunities. *International Journal of Information Security*, 23(1), 577–608. <https://doi.org/10.1007/s10207-023-00755-2>

Assimakopoulos, S., Baider, F. H., & Millar, S. (2017). *Online hate speech in the European Union: A discourse-analytic perspective*. Springer Nature.

Bastos, E. (2020). Estratégias discursivas na construção do espírito nacionalista: A Hungria 'de' Viktor Orbán. *REDIS: Revista de Estudos do Discurso*, (9), 12–42. <https://doi.org/10.21747/21833958/red9a1>

Bennett, W. L., & Kneuer, M. (2024). Communication and democratic erosion: The rise of illiberal public spheres. *European Journal of Communication*, 39(2), 177–196.

Bennett, W. L., & Livingston, S. (2025). Platforms, politics, and the crisis of democracy: Connective action and the rise of illiberalism. *Perspectives on Politics*, 23(4), 1362–1381. <https://doi.org/10.1017/S1537592724002123>

BSC-LT. (2025). *mRoBERTa* [Computer software]. Hugging Face.

Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020). Spanish pre-trained BERT model and evaluation data. In *Proceedings of the Practical ML for Developing Countries Workshop at ICLR 2020 (PML4DC)*.

Cervi, L., Tejedor, S., & Villar, M. (2023). Twitting against the enemy: Populist radical right parties discourse against the (political) “other”. *Politics and Governance*, 11(2), 235–248. <https://doi.org/10.17645/pag.v11i2.6391>

Curley, C., Siapera, E., & Carthy, J. (2022). Covid-19 protesters and the far right on Telegram: Co-conspirators or accidental bedfellows? *Social Media + Society*, 8(4). <https://doi.org/10.1177/20563051221129187>

DiMaggio, A. R. (2022). Conspiracy theories and the manufacture of dissent: QAnon, the ‘Big Lie’, Covid-19, and the rise of rightwing propaganda. *Critical Sociology*, 48(6), 1025–1048. <https://doi.org/10.1177/08969205211073669>

Donovan, J., Lewis, B., & Friedberg, B. (2018). Parallel ports: Sociotechnical change from the alt-right to alt-tech. In M. Fielitz & N. Thurston (Eds.), *Post-digital cultures of the far right: Online actions and offline consequences in Europe and the US* (pp. 49–66). transcript Verlag. <https://doi.org/10.1515/9783839446706-004>

Dowling, M.-E. (2023). Far-right populism in alt-tech: A challenge for democracy? *New Media & Society*, 27(4), 2068–2086. <https://doi.org/10.1177/14614448231205889>

Edilashvili, N. (2025). Anti-immigration rhetoric in Donald Trump’s presidential speeches: A corpus-assisted comparative analysis of US presidential discourse (1946–2021) populist and emotionally charged language versus constructive, policy-oriented framing. *Przegląd Politologiczny*, (2), 141–156.

Entman, R. M. (2003). *Projections of power: Framing news, public opinion, and U.S. foreign policy*. University of Chicago Press.

García-Marín, J., Fernández-García, B., & Luengo, O. (2025). Otherness and othering. In A. Nai, M. Grömping, & D. Wirz (Eds.), *Elgar encyclopedia of political communication* (Vol. 3, pp. 131–134). Edward Elgar Publishing. <https://doi.org/10.4337/9781035301447.vol3.00036>

Hanu, L., & Unitary team. (2020). *Detoxify* [Computer software]. GitHub.

Indelicato, M. E., & Magalhães Lopes, M. (2024). Understanding populist far-right anti-immigration and anti-gender stances beyond the paradigm of gender as ‘a symbolic glue’: Giorgia Meloni’s modern motherhood, neo-Catholicism, and reproductive racism. *European Journal of Women’s Studies*, 31(1), 6–20. <https://doi.org/10.1177/13505068241230819>

Lonami. (2024). *Telethon* (Version 1.41.1) [Computer software]. GitHub.

Machuy, C. (2025). The enemy on the left: The construction of the “evil communist” in Brazilian far-right discourse. In R. A. Novais & R. Christofolletti (Eds.), *The Palgrave handbook on right-wing populism and otherness in global perspective*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-77868-1_7

Marwick, A., Clancy, B., & Furl, K. (2022). Far-right online radicalization: A review of the literature. *The Bulletin of Technology & Public Life*. <https://doi.org/10.21428/bfcb0bff.e9492a11>

Molla, T. (2024). Racial othering: Structural roots and anti-racist actions from below. *SN Social Sciences*, 4(7). <https://doi.org/10.1007/s43545-023-00806-4>

Morales, E., Hodson, J., O'Meara, V., Gruzd, A., & Mai, P. (2025). Online toxic speech as positioning acts: Hate as discursive mechanisms for othering and belonging. *New Media & Society*. <https://doi.org/10.1177/14614448251338493>

Mudde, C. (2019). *The far right today*. Polity.

Montani, I., Honnibal, M., Boyd, A., Van Landeghem, S., & Peters, H. (2023). *explosion/spaCy: v3.7.2: Fixes for APIs and requirements (v3.7.2)* [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.10009823>

Nartey, M. (2021). Marginality and otherness: The discursive construction of LGBT issues/people in the Ghanaian news media. *Media, Culture & Society*, 44(4), 785–801. <https://doi.org/10.1177/01634437211045552>

Novais, R. A., & Christofolletti, R. (Eds.). (2025). *The Palgrave handbook on right-wing populism and otherness in global perspective*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-77868-1>

Pérez-Díaz, P. L., & Arroyas Langa, E. (2025). El populismo disruptivo de Javier Milei: Análisis de su estrategia discursiva en redes sociales. *European Public & Social Innovation Review*, 10, 1–23. <https://doi.org/10.31637/epsir-2025-2064>

Pirro, A. L. P. (2023). Far right: The significance of an umbrella concept. *Nations and Nationalism*, 29(1), 101–112. <https://doi.org/10.1111/nana.12860>

Pradel, F., Zilinsky, J., Kosmidis, S., & Theocharis, Y. (2024). Toxic speech and limited demand for content moderation on social media. *American Political Science Review*, 118(4), 1895–1912.

Resende, E., & Freire, C. M. (2025). Populism, enemies and threat construction, and otherness: How populists use othering rhetoric to get reelected to remain in power. The case of Viktor Orbán 2028 election. In R. A. Novais & R. Christofolletti (Eds.), *The Palgrave handbook on right-wing populism and otherness in global perspective*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-77868-1_12

Rogers, R. (2020). Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication*, 35(3), 213–229. <https://doi.org/10.1177/0267323120922066>

Salvatore, S., Mannarini, T., Avdi, E., Battaglia, F., Cremaschi, M., Fini, V., Forges Davanzati, G., Kadianaki, I., Krasteva, A., Kullasepp, K., Matsopoulos, A., Mølholm, M., Redd, R., Rochira, A., Russo, F., Santarpia, A., Sammut, G., Valmorbida, A., & Veltri, G. A. (2019). Globalization, demand of sense and enemization of the other: A psychocultural analysis of

European societies' sociopolitical crisis. *Culture & Psychology*, 25(3), 345–374.
<https://doi.org/10.1177/1354067X18779056>

Scrinzi, F. (2025). Populist radical right frames of gender and sexuality in France and Italy: Targeting feminists and other enemies of the people. *Social Politics: International Studies in Gender, State & Society*. <https://doi.org/10.1093/sp/jxaf056>

Segado-Boj, F., Martín-Quevedo, J., González-Aguilar, J.-M., & Antona-Jimeno, T. (2026). Political and social hate speech: Differences on emotional and rhetorical features on TikTok messages. *Frontiers in Political Science*, 7. <https://doi.org/10.3389/fpos.2025.1712927>

Southern, R., & Harmer, E. (2019). Othering political women: Online misogyny, racism and ableism towards women in public life. In K. Lumsden & E. Harmer (Eds.), *Online othering*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-12633-9_8

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–45). Psychology Press. <https://doi.org/10.4324/9781410605634>

Stewart, N. K., Al-Rawi, A., Celestini, C., & Worku, N. (2023). Hate influencers' mediation of hate on Telegram: "We declare war against the anti-white system". *Social Media + Society*, 9(2). <https://doi.org/10.1177/20563051231177915>

Svatoňová, E., & Doerr, N. (2024). How anti-gender and gendered imagery translate the Great Replacement conspiracy theory in online far-right platforms. *European Journal of Politics and Gender*, 7(1), 83–101. <https://doi.org/10.1332/25151088Y2023D000000006>

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–37). Brooks/Cole.

Urman, A., & Katz, S. (2020). What they do in the shadows: Examining the far-right networks on Telegram. *Information, Communication & Society*, 25(7), 904–923. <https://doi.org/10.1080/1369118X.2020.1803946>

Van Dijk, T. A. (1993). *Elite discourse and racism*. SAGE Publications. <https://doi.org/10.4135/9781483326184>

Vaughan, M., & Heft, A. (2023). Anti-elitism in the European radical right in comparative perspective. *JCMS: Journal of Common Market Studies*, 61, 76–94. <https://doi.org/10.1111/jcms.13347>

Vidgen, B., Thrush, T., Talat, Z., & Kiela, D. (2021). Learning from the worst: Dynamically generated datasets to improve online hate detection. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (Vol. 1, pp. 1667–1682).

Vollhardt, J. R. (2012). Collective victimization. In L. R. Tropp (Ed.), *The Oxford handbook of intergroup conflict* (pp. 136–157). Oxford University Press.

Willaert, T., Peeters, S., Seijbel, J., & Van Raemdonck, N. (2022). Disinformation networks: A quali-quantitative investigation of antagonistic Dutch-speaking Telegram channels. *First Monday*, 27(5). <https://doi.org/10.5210/fm.v27i5.12533>

Winberg, O. (2017). Insult politics: Donald Trump, right-wing populism, and incendiary language. *European Journal of American Studies*, 12(2).

Wodak, R. (2021). *The politics of fear*. SAGE Publications. <https://doi.org/10.4135/9781529739664>

Young, J. (2025). Dogwhistles, discrimination, humour and the law: Regulating implicit messaging. *Open Library of Humanities*, 11(2). <https://doi.org/10.16995/olh.19789>

Zamri, N. A. K., Mohamad Nasir, N. A., Hassim, M. N., & Ramli, S. M. (2023). Digital hate speech and othering: The construction of hate speech from Malaysian perspectives. *Cogent Arts & Humanities*, 10(1), Article 2229089. <https://doi.org/10.1080/23311983.2023.2229089>