

# Polarización y discurso de odio con sesgo de género asociado a la política: análisis de las interacciones en Twitter

## Polarization and hate speech with gender bias associated with politics: analysis of interactions on Twitter

Blanco-Alfonso, I., Rodríguez-Fernández, L. y Arce-García, S.<sup>1</sup>

Recibido: 4-03-2022 – Aceptado: 05-07-2022

<https://doi.org/10.26441/RC21.2-2022-A2>

**RESUMEN:** La propagación de discursos de odio a través de las redes sociales contribuye a intoxicar la esfera pública y a mermar la calidad de las democracias liberales. Este tipo de discursos es singularmente virulento contra la clase política y contra el feminismo. Partiendo de esta realidad, la presente investigación tratará de delimitar el sesgo de género en el discurso de odio en el ámbito político. ¿Reciben las mujeres políticas más agresiones verbales que sus homólogos varones, no por su condición de políticas, sino por su condición de mujeres? ¿Concentran las mujeres políticas más polaridad emocional que los hombres en las menciones que reciben en Twitter? A través del análisis del discurso operado con técnicas de PLN para la de detección de las emociones y de la minería de texto sobre un corpus de 3.483.232 de tuits recolectados de 20 cuentas de políticos españoles, se confirma que los mensajes que reciben las mujeres políticas concentran mayor polaridad emocional que los de los hombres, pero no más odio, que es un poco superior en los hombres. También se confirma que las expresiones sexistas y misóginas son empleadas para denigrar a las mujeres y, por extensión, al feminismo, lo que convierte al discurso del odio en un tipo de desorden informativo.

**Palabras clave:** discurso del odio; desinformación; desórdenes informativos; género; sexismo; emociones; política.

**ABSTRACT:** The spread of hate speech through social media contributes to poisoning the public sphere and undermining the quality of liberal democracies. This type of discourse is particularly virulent against the political class and against feminism. Taking this reality as a starting point, this research will attempt to identify the gender bias in hate speech in the political sphere: do female politicians receive more verbal attacks than their male counterparts, not because they are politicians, but because they are women? Do female politicians receive more emotional polarity in the mentions they receive on Twitter than their male counterparts? Through discourse analysis using PLN techniques for emotion detection and text mining on a corpus of 3,483,232 tweets collected from 20 accounts of Spanish politicians, it is found that the messages received by women politicians concentrate more emotional polarity than men's, but not more hatred, which is slightly higher in men. It also confirms that sexist and misogynist expressions are used to denigrate women and, by extension, feminism, which makes hate speech a type of information disorder.

**Keywords:** hate speech; disinformation; information disorders; gender; sexism; emotions; politics.

---

<sup>1</sup> **Ignacio Blanco-Alfonso** es Catedrático de Periodismo en la Universidad CEU San Pablo, premio extraordinario de doctorado, con tres sexenios de investigación. Director de la revista *Doxa Comunicación*. Su línea de investigación se centra en la vulnerabilidad digital y la alfabetización mediática para una sociedad inclusiva. [iblanco@ceu.es](mailto:iblanco@ceu.es), <https://orcid.org/0000-0002-2595-464X>

**Leticia Rodríguez-Fernández** es Doctora en Ciencias de la Información. Profesora e investigadora en la Universidad de Cádiz. Anteriormente se desempeñó en la Universidad Antonio Nebrija donde dirigía el Grado en Comunicación Corporativa, Protocolo y Organización de Eventos. Entre sus líneas de investigación se encuentra la comunicación corporativa y la propaganda. [leticia.rodriguez@uca.es](mailto:leticia.rodriguez@uca.es), <https://orcid.org/0000-0002-7472-5472>

**Sergio Arce-García** es Doctor en Comunicación con premio extraordinario, con un sexenio de investigación. Profesor e investigador en la Universidad Internacional de La Rioja. Su línea de investigación se centra en las redes sociales y medios digitales mediante estudios masivos. [sergio.arce@unir.net](mailto:sergio.arce@unir.net), <https://orcid.org/0000-0003-0578-9787>

## 1. Introducción. El discurso de odio en Twitter y su relación con la política

La toxicidad de los mensajes en las redes sociales es un problema complejo para la sociedad, pues afecta a determinados colectivos y contribuye a debilitar la democracia. Aunque no hay una definición universal sobre el discurso de odio, dadas las implicaciones éticas y jurídicas que tendría la misma, entendemos que se refiere a aquellas expresiones que fomentan los prejuicios o la intolerancia y contribuyen indirectamente a que se genere un clima de hostilidad que pueda propiciar, eventualmente, actos discriminatorios o ataques violentos (Gagliardone et al. 2015). Parekh (2006) observa que estos discursos señalan a un grupo de individuos sobre la base de ciertas características, estigmatiza a su objetivo atribuyéndole un conjunto de cualidades constitutivas que se consideran altamente indeseables y no necesariamente tiende a apelar a la violencia, sino que puede también utilizar formas más refinadas como chistes, insinuaciones o imágenes ambiguas.

El discurso de odio ha encontrado en las redes sociales un altavoz que fomenta la adhesión y la viralidad, siendo Twitter, por su vinculación con la política, una plataforma especialmente sensible. Así se evidencia en los numerosos trabajos académicos (Ramírez-García et al., 2022) que abordan su relación con los discursos de odio contra la migración (Arcila-Calderón et al., 2022), por cuestiones de ideología política (Amores et al., 2021) o de género (Piñeiro-Otero; Martínez-Rolán, 2021); así como en los intentos, aún insuficientes, de la propia red de eliminar este tipo de contenidos.

Los discursos de odio también han sido estudiados por su capacidad de polarizar en las campañas electorales (Lim, 2017) y su vinculación con el populismo. En Twitter, esta polarización se observa a través de la conversación social. Smith et al. (2014) recogieron distintos tipos de arquetipos conversacionales, destacando las multitudes polarizadas (*polarized crowds*) en las que se suelen observar dos grandes grupos con poca conexión entre ellos que debaten temas políticos muy divisivos y acalorados. Las multitudes polarizadas no discuten entre ellos, se ignoran mientras señalan diferentes recursos web y usan diferentes *hashtags*. Así, se evita el diálogo y se busca insistir en las ideas y argumentos que sostienen los mismos, creando cámaras eco que continúan alimentando las diferencias.

España es el país más polarizado de Europa (Gidron et al., 2018) con una alta polarización ideológica, pero baja polarización afectiva (Garrido et al., 2021). El 68% de los ciudadanos considera que los partidos tradicionales y los políticos no se preocupan de personas como ellos (Ipsos, 2021). En este sentido, se ha observado que los niveles de polarización ciudadana suelen estar relacionados con la pertenencia a determinados grupos sociales y consumos informativos políticos (Robles et al., 2022), circunstancia que podría darse por el efecto contagio y como consecuencia de la polarización recibida por agentes políticos. Recientes investigaciones realizadas sobre la actividad de los parlamentarios españoles en Twitter evidencian el uso de estrategias discursivas como la acusación de falsedad a otros representantes, que contribuyen a generar desconfianza institucional y apatía entre la ciudadanía (Campos-Domínguez et al., 2022). Por otra parte, Guerrero-Solé y Philippe (2020) observaron que la toxicidad de los diputados españoles en Twitter se redujo durante el estado de alarma, en comparación con la toxicidad antes y después. En este sentido, se identificaba a Vox como el partido emisor de mayor cantidad de mensajes tóxicos durante el estado de alarma y se observaron también diferencias relacionadas con el género pues los diputados hombres publicaron mayor proporción de mensajes tóxicos que las mujeres.

Los discursos de odio también funcionan como una de “las tácticas discursivas más importantes utilizadas en los movimientos populistas para promover sus acciones y agenda” (Ramírez-Plascencia, Alonzo-González y Ochoa-Amezquita, 2022). Así, el feminismo junto con el colectivo LGTBI y la migración se han convertido en el centro del discurso de odio de los partidos de extrema derecha

europea. En el caso español, los ataques se centran en desprestigiar al feminismo como pensamiento político, en atacar a aquellas representantes políticas que se identifican como feministas, y en negar y cuestionar la violencia de género y su legislación (Igareda et al., 2021). Cabe destacar también que la temática “mujer” les genera en Twitter un buen retorno en actividad y, como se demostró en las elecciones andaluzas de 2018, contribuye a vincular temas como la supuesta vinculación de la violencia de género con la inmigración (Luque-Ortíz y Cano-Alarcón, 2020).

Estas narrativas digitales motivan que organizaciones como Amnistía Internacional incidan en la creciente violencia de género digital, cuestión que se ha convertido en un problema global. Según sus pesquisas, el 7,1% de los tuits enviados a mujeres y recogidos en su muestra de estudio eran “abusivos”. Esto equivale a 1,1 millones de tuits que mencionan a 778 mujeres. Se añade, además, que las mujeres de color (negras, asiáticas, latinas y mestizas) tienen un 34% más de probabilidades de ser mencionadas en este tipo de contenidos abusivos.

### 1.1. El discurso de odio y su relación con la desinformación con sesgo de género

Según Williams (2021), la construcción del discurso de odio se articula a través de cinco elementos fundamentales: (1) la infracción de las reglas; (2) induciendo a la vergüenza en las víctimas; (3) induciendo el miedo en las víctimas a través de amenazas e intimidación; (4) intentando deshumanizar a la víctima comparándolos con insectos, alimañas o primates; y (5) a través de la desinformación en personas o en grupos a los que pertenece.

Sobre este último punto, cabe destacar que normalmente las campañas de desinformación recurren en sus tácticas narrativas a grupos de población sobre los que focalizan el odio y el malestar, como se evidenció con la minoría rohingya en Myanmar (Moreno y Calvillo, 2021). La simplificación del enemigo único es un principio clásico de la propaganda, que se extiende también a la propaganda digital (Rodríguez-Fernández, 2021) y alienta la creación y exageración de historias que puedan suponer una amenaza para la sociedad.

En España, casi diez millones de españoles recibieron desinformación política por WhatsApp durante la campaña para las elecciones generales de 2019 y gran parte de este contenido era discurso de odio contra la inmigración (14%), el feminismo y el colectivo LGBTQ (10%) (Avaaz, 2019). Igualmente, Blanco-Alfonso et al. (2021) y Bernal-Triviño y Clares-Gavilán (2019) confirman que la política es el principal argumento de los bulos desmentidos, e identifican la inmigración, el racismo, el feminismo y la violencia de género como otros temas prominentes de las noticias falsas.

Sobre estos supuestos, se puede desarrollar una línea de investigación sobre “desinformación de género” (Occeñola, 2018), entendida como “la difusión de información falsa o engañosa que ataca a las mujeres (especialmente líderes políticas, periodistas y figuras públicas), basando el ataque en su identidad como mujeres” (Sessa, 2020). Esta desinformación mezcla viejas actitudes sexistas con el anonimato y el alcance de las redes sociales con el objetivo de destruir la reputación de las mujeres y apartarlas de la vida pública (Jankowicz, 2017). Marianna Spring (2021), reportera de la BBC especializada en desinformación, demostró a través de un experimento social que las mujeres tienen más probabilidades de sufrir este tipo de abuso que los hombres y, a menudo, se combina con racismo y homofobia. Con un perfil falso, antivacunas y conspiranoico, interactuó en cinco redes sociales distintas, con el contenido recomendado por las mismas, pero sin enviar ningún mensaje de odio. Entre las conclusiones se observa que Facebook e Instagram recomendaban en mayor grado contenido antimujeres, que a veces involucraban violencia sexual.

Según *The Economist Intelligence Unit* (2020), sobre una muestra de 4.500 mujeres en 45 países, el 67% identificó el uso de tácticas de desinformación con intención de desacreditar o dañar a las mujeres,

el 65% había detectado discurso de odio con presencia de lenguaje sexista, y el 48% advertía del uso de *astroturfing* para incrementar la difusión de contenido dañino contra la mujer. Estos desórdenes informativos son, además, una estrategia para socavar los valores democráticos occidentales, ya que la misoginia es un marco de referencia que contribuye a mantener a flote la industria de las noticias falsas (Pardy, 2018).

Esta incidencia ha motivado que servicios de *fact-checking* como Maldita.es cree secciones de verificación específicas como Maldito Feminismo. Del estudio de sus contenidos se desprende que, en España, los objetivos fundamentales de la desinformación con sesgo de género son dañar al feminismo con propósito político para socavar su credibilidad y dignidad (Herrero-Diz, Pérez-Escolar y Plaza Sánchez, 2020). Objetivos que han sido también recogidos en otros países como EE. UU. en los que se relaciona la misoginia, la islamofobia y el pensamiento supremacista blanco como argumentarios fundamentales de la extrema derecha en internet (Marwick y Lewis, 2017). EU Disinfo Lab (2020) advierte de que estas narrativas misóginas impulsan determinadas agendas sociales o políticas, a través de la representación de las mujeres como enemigas y oponentes en el debate público o como víctimas (Sessa, 2020).

El acoso y las campañas de desprestigio se incrementan en determinados perfiles profesionales. El 73% de mujeres periodistas reconoce haber experimentado violencia *online* (n=714 reporteras/113 países) y el 41% ha sido blanco de campañas orquestadas de desinformación. Las periodistas, defensoras de derechos humanos e implicadas en política son las más vulnerables al recibir, en primer lugar, ataques por razón de su género (47%), seguidos por cuestiones sobre política y elecciones (44%), y derechos y política social (31%) (Posetti et al., 2021).

Igualmente, las mujeres políticas suelen ser objeto de una tendencia creciente de desprestigio. En 2016, el 41,8% de parlamentarias europeas (n=55) reconocía haber sido objeto de difusión en los medios sociales de imágenes humillantes o con connotaciones sexuales de la interesada (Unión Interparlamentaria, 2016). En 2018, el porcentaje ascendía a 58,2% (n=123), llegando hasta el 76,2% en el caso de las entrevistadas menores de 40 años (Unión Interparlamentaria, 2018). En España, y atendiendo a Twitter, las mujeres de los ámbitos de la comunicación y de la política concentran el 90% de los insultos y discursos de odio por cuestiones de género. En el caso de las políticas, se recogen una media de 15 apelaciones negativas diarias (Piñeiro-Otero y Martínez-Rolán, 2021).

Señalan Stabile et al. (2019) que estas historias falsas evidencian narrativas de género que apoyan los estereotipos de que las mujeres no son aptas para los puestos de liderazgo, y demonizan o trivializan a las mujeres, dependiendo de su grado de poder percibido. Recientemente, en Alemania, durante las últimas elecciones presidenciales (2021), la única candidata mujer, Annalena Baerbock, del partido Los Verdes, fue objeto del 70% de la desinformación analizada. Al ser la única candidata femenina, se concluye que la mayoría de los contenidos falsos se dirigieron a desacreditar su propuesta por cuestiones de género (DW, 2021).

Cabe destacar, por otra parte, y tomando como referencia los datos de ONU Mujeres (2021) que solo en 22 países hay jefas de Estado o de Gobierno, y 119 países nunca han sido presididos por mujeres. La organización calcula que la igualdad de género en las más altas esferas de decisión no se logrará en 130 años. La desinformación y las prácticas de acoso resultan efectivas pues merman progresivamente la proyección de liderazgo y su visibilidad en redes sociales (Ferruz-González et al., 2020). El 50% de las mujeres (n=4.500) observan internet como un lugar inseguro para compartir sus pensamientos, lo que conlleva la autocensura y la subrepresentación de mujeres en discusiones sociales, políticas y económicas (*The Economist Intelligence Unit*, 2020). Sin embargo, las prácticas de seguimiento y curación de contenido por parte de las redes sociales resultan insuficientes, pues se destinan mayores recursos y se presta más atención a otros contenidos falsos. Según datos del Centro para Contrarrestar el Odio

Digital, el 97% de 330 cuentas que habían enviado abusos misóginos en Twitter e Instagram permanecieron activas en el sitio después de ser denunciadas (Spring, 2021). Igualmente, tampoco se recogen respuestas o acciones concretas por parte de los gobiernos (Jankowicz, 2017).

## 2. Objetivos e hipótesis

Asumiendo, por lo tanto, la existencia de un sesgo de género cada vez más intenso y polarizado en los discursos de odio la presente investigación tiene como objetivos:

- O1: analizar el sesgo de género en las menciones que reciben los políticos españoles en Twitter, para establecer si existen diferencias discursivas en función del sexo de la persona y/o tema del que se habla.
- O2: hacer una aportación significativa al estudio del discurso del odio como desorden informativo, proporcionando un repertorio de los insultos más recurrentes en lengua española contras las mujeres políticas y su relación con el sesgo de género.

Las dos hipótesis de partida son:

- H1: Las agresiones verbales recibidas por las mujeres políticas españolas presentan un componente sexista que no se percibe en las agresiones recibidas por sus homólogos varones.
- H2: La carga emocional de los mensajes que reciben las mujeres políticas españolas está más polarizada que los mensajes que reciben los políticos hombres.

Para demostrarlo, se llevará a cabo un análisis del discurso operado con técnicas complementarias:

- Por un lado, el análisis de la polaridad y de las emociones, que permitirá medir el grado de polarización del discurso a partir de los sentimientos dominantes en los mensajes que reciben los políticos en Twitter.
- En este sentido, se tendrá en cuenta la probabilidad de que los mensajes con mayor carga de odio hayan podido ser emitidos por bots, lo que demostraría un grado de artificialidad en el discurso que estaría deformando el verdadero sentir social.
- Por otro lado, la minería de texto, que permitirá analizar el léxico empleado para descalificar a las mujeres políticas españolas y relacionarlo con los temas con que son vinculados en la conversación pública en la red social Twitter.

## 3. Material y métodos

La investigación se ha centrado en el estudio de la conversación pública en la red social Twitter. Se trata de un estudio cuantitativo y cualitativo de interpretación final de grandes volúmenes de datos y de información, basado en el uso de técnicas de análisis masivo de datos para su tratamiento y eliminación de posibles subjetividades.

Para garantizar la representatividad de los resultados, se han seleccionado dos hombres y dos mujeres con alta exposición mediática y con cuenta activa de Twitter, de cada uno de los cinco partidos políticos de ámbito nacional con representación parlamentaria: Ciudadanos, PP, PSOE, UP y VOX. En total, se han analizado los perfiles de Twitter de 20 políticos españoles (10 hombres y 10 mujeres) durante todo el mes de octubre de 2021. Como se muestra en la Tabla 1, se han extraído 3.483.232 tuits (1.144.198 de las cuentas de las 10 mujeres y 2.339.034 de las cuentas de los 10 hombres).

**Tabla 1.** Perfiles políticos analizados y tuits recogidos

Partido	Nombre	Sexo	Nick	Seguidores	Tuits recogidos
Cs	Inés Arrimadas	M	@InesArrimadas	690,5k	53.268
	Begoña Villacís	M	@begonavillacis	198,1k	53.306
	Juan Marín	H	@JuanMaris_Cs	35,2k	10.328
	Edmundo Bal	H	@BalEdmundo	51,4k	28.727
PSOE	Adriana Lastra	M	@AdriLastra	118,3k	14.968
	Nadia Calviño	M	@NadiacCalvino	72k	17.468
	Miquel Iceta	H	@miqueliceta	118,2k	41.979
	Pedro Sánchez	H	@sanchezcastejon	1,6M	773.385
PP	Cayetana Álvarez	M	@cayetanaAT	395,4k	22.432
	Isabel Díaz Ayuso	M	@IdiazAyuso	623,3k	119.736
	Pablo Casado	H	@pablocasado	556k	409.434
	J.L. Martínez Almeida	H	@AlmeidaPP_	324,5k	93.386
UP	Yolanda Díaz	M	@Yolanda_Diaz_	377,2k	406.662
	Irene Montero	M	@IreneMontero	620,8k	271.675
	Pablo Echenique	H	@PabloEchenique	591,5k	328.008
	Alberto Garzón	H	@agarzon	1,1M	51.083
Vox	Rocío Monasterio	M	@monasterioR	291k	61.640
	Macarena Olona	M	@Macarena_Olona	321,1k	123.043
	Santiago Abascal	H	@Santi_ABASCAL	630,5k	356.559
	Iván Espinosa de los M.	H	@ivandlm	347,1k	246.145
<b>TOTAL</b>					<b>3.483.232</b>

Fuente: elaboración propia

La extracción de la información se ha efectuado en diciembre de 2021 mediante software R en su entorno virtual RStudio 1.2.5033, a través del interfaz de aplicaciones (API) 2.0 de la red social Twitter. Esta API está preparada para empresas y usos académicos, lo que le confiere una gran fiabilidad al permitir la consulta histórica con mayores privilegios que el acceso gratuito. Como criterio de búsqueda, se recogieron todos los mensajes (tanto los directos como los retuits) que mencionaran el *nickname* de alguna de las cuentas de los políticos seleccionados y que estuvieran escritos en español durante el mes completo de octubre de 2021, obteniendo un total de 3.483.232 mensajes. Con estos criterios, el 31,5% de los mensajes recolectados fueron directos, y el 68,5% fueron retuits. Para su recogida en API 2.0 se empleó la librería *academictwitter* (Barrie y Chung-ting, 2021).

### 3.1. Triangulación metodológica

Dado que la investigación pretende, por un lado, delimitar el sesgo de género en el discurso de odio contra la clase política y, por otro, analizar el léxico recurrente en las agresiones verbales contra las mujeres, la investigación se ha efectuado con las siguientes técnicas metodológicas:

- **Análisis de polaridad y emociones.** Se ha llevado a cabo a partir del corpus léxico desarrollado por el *National Research Council of Canada* (NRC) de más de 14.000 palabras en su versión 0.9.2 en español (Mohammad y Turney, 2010, 2013), a través del uso de la

librería Syuzhet (Jockers, 2017) en R. A partir de este corpus, es posible determinar el grado de polaridad de cada mensaje (positiva o negativa) sobre las siguientes emociones básicas: ira, anticipación, aversión, miedo, alegría, tristeza, sorpresa y confianza (Plutchik, 1980; Sauter et al., 2010). Este método permite comparar el léxico (sustantivos y verbos) con valores prefijados, que aumentan o disminuyen en función de otras palabras con significado, adjetivos y adverbios a su alrededor, y consignar un valor numérico global de intensidad para cada mensaje. El valor numérico alcanzado viene dado por el algoritmo al relacionar dichas palabras en un mismo mensaje. Los valores de polaridad podrán ser negativos o positivos, es decir, que muestran un discurso desfavorable o favorable; mientras que las emociones siempre tendrán un valor positivo de mayor o menor intensidad. Por lo tanto, esta técnica, basada en el Procesamiento del Lenguaje Natural permite determinar el grado de subjetividad positiva o subjetividad negativa de cada tuit, así como su intensidad. Hay que indicar, no obstante, que, si bien la herramienta está muy avanzada en inglés, no lo está tanto en otros idiomas (en torno al 70% de acierto en lenguas latinas) (Mohammad, 2016).

- **Detección y cuantificación de odio.** Se ha efectuado a partir del corpus léxico Hurltex en su versión en español de 2018, que contiene 5.007 palabras asociadas con odio en los mensajes (Bassignana, Basile y Patti, 2018). A través de la librería Syuzhet, mencionada en el punto anterior, se ha medido la presencia e intensidad de odio mediante Procesamiento de Lenguaje Natural. El resultado ha sido la identificación de un primer corpus de 423.461 tuits (12,15% del total) con algún valor de odio. Su cuantificación global, al utilizar la misma librería, funciona de un modo similar al cálculo de la polaridad y las emociones, es decir, se consigue por la medición de las palabras junto con las que están a su alrededor y su modulación mediante adjetivos y adverbios. Para evitar usos coloquiales, ironías o dobles sentidos, en los estudios de odio se determinarán aquellas que estén únicamente en el cuartil más elevado, que en este caso se corresponde con un nivel superior a 3.

Las gráficas de polaridad y de odio se han efectuado mediante diagramas de caja, donde la línea gruesa representa la mediana del valor; la parte inferior y la parte superior de la caja representan el primer y tercer cuartil. Por su parte, la línea continua que sale de cada caja puede alcanzar un valor de hasta 1,5 del rango intercuartílico, mientras que los puntos sueltos representan datos atípicos desde el punto de vista estadístico.

- **Minería de texto.** De acuerdo con Liddy (1998), las etapas del proceso consisten en preparación del texto, procesamiento y posterior análisis. Siguiendo este orden, de entre los tuits en los que se ha identificado odio según la detección algorítmica (423.461 tuits), se seleccionaron solo los mensajes que estuvieran por encima del tercer cuartil de mayor nivel de intensidad (por encima de valor 3). Así, se obtuvieron 372.777 tuits distribuidos de esta forma: 108.346 de las cuentas hacia las 10 mujeres (9,47% del total de tuits recibidos por las mujeres) y 264.431 hacia las cuentas de los 10 hombres (11,31% del total recibidos por los hombres). Esta criba ha permitido a los investigadores focalizarse en el análisis de los mensajes detectados como de odio intenso, y eludir los usos esporádicos e irónicos. Una vez obtenido este corpus, mediante R se eliminaron las palabras que no aportaban significado (denominadas *stopwords*), y aplicando el algoritmo *Snowball* en forma *stemming*, se determinaron los morfemas de cada palabra, agrupando las más usadas mediante análisis de clúster con las que más se relacionan (Arce-García y Menéndez-Menéndez, 2018). La última fase ha consistido en elaborar un repertorio con las agresiones verbales identificadas en una muestra suficiente y representativa de los mensajes que contenían un odio más intenso (463 recibidos por las políticas y 471 por los políticos) y, mediante minería de texto, conocer con qué otras palabras se relacionan. Dicho repertorio ha permitido conocer la naturaleza léxica y semántica de las

agresiones recibidas por las mujeres y por los hombres; al cruzarlas con las palabras con las que se vinculan, se ha podido establecer la relación con el sesgo de género.

- **Análisis de probabilidad de comportamiento bot.** Se ha efectuado mediante la librería R *Tweetbotornot2* de Kearney (2020) que, en su segunda versión, analiza 50 variables de comportamiento y naturaleza de cada cuenta de Twitter. Está considerado entre los mejores de los empleados en Ciencias Sociales (Martini et al., 2021).

## 4. Análisis y resultados

### 4.1. Estudio de la polaridad y las emociones

El análisis de las 8 emociones básicas de Plutchik detectadas por el algoritmo en el corpus de mensajes que conforman la investigación (n=3.483.232) presenta una valencia positiva (por encima de cero) en todos los casos, aunque con desigual intensidad. Por orden decreciente, el promedio de la polaridad emocional es el siguiente: confianza=0,6117; miedo=0,4048; tristeza=0,3796; ira=0,2852; aversión=0,2677; anticipación=0,2641; alegría=0,2344; sorpresa=0,1569.

Al observar estas emociones en los tuits extraídos de las cuentas de los políticos analizados, se recoge mayor polaridad emocional en los mensajes destinados a las mujeres (0,162), que en los de los hombres (0,105). Los promedios indican proximidad entre ambos grupos; sin embargo, llama la atención que en 7 de las 8 emociones básicas detectadas en los mensajes destinados a las mujeres se percibe más carga emocional que en los de los hombres, que solamente están ligeramente por encima en el factor ira (0,2852 en los hombres y 0,2826 en las mujeres).

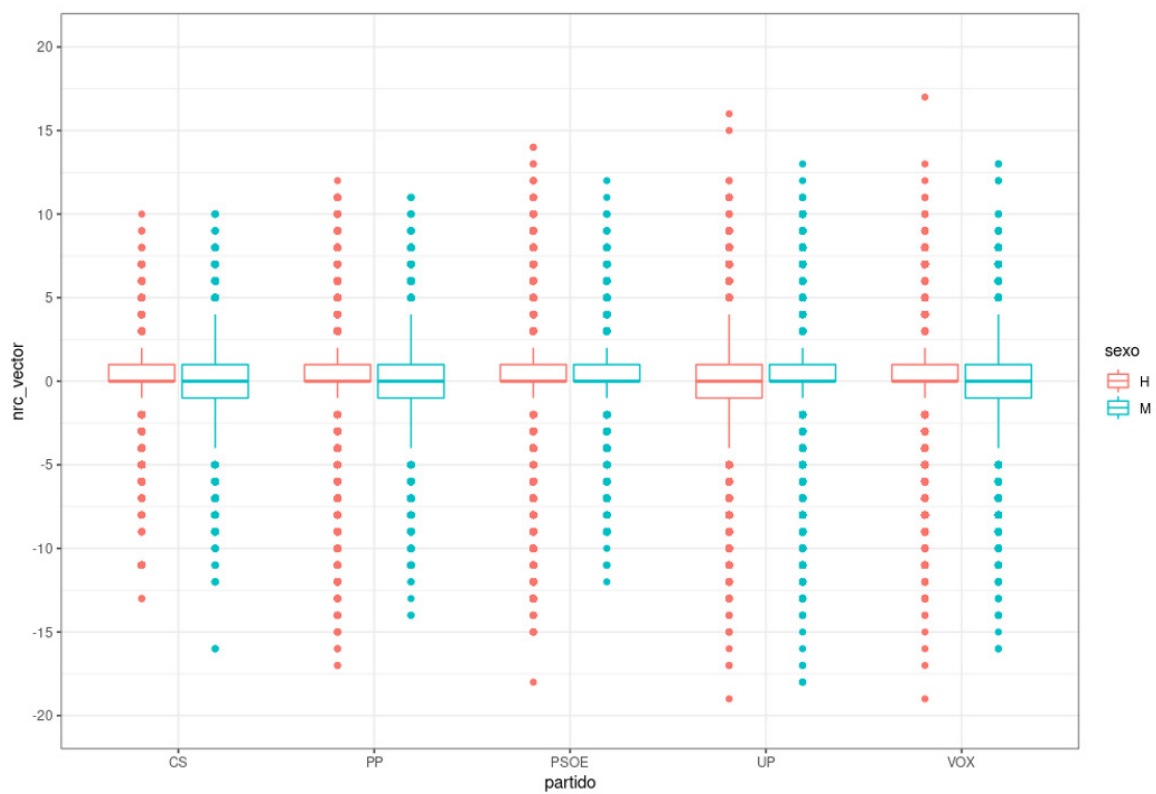
La polarización de estas emociones respecto a la mediana (=0) alcanza un mínimo de -18 y un máximo de 13 en el caso de las mujeres; y un mínimo de -19 y un máximo de 17 en el caso de los hombres. En el caso de las mujeres, Inés Arrimadas (Cd's) es la política que alcanza el vector mínimo de polaridad, mientras que en el máximo empatan Macarena Olona (Vox) y Yolanda Díaz (UP), con 13 cada una. En el caso de los hombres, el vector mínimo de polaridad lo obtienen Iván Espinosa de los Monteros (Vox) y Pablo Echenique, ambos con -19; y el máximo, Santiago Abascal con 17.

Respecto a la polaridad emocional se observan dos fenómenos particulares: por un lado, que solo 5 de las 20 personas analizadas concentran los mayores grados de polaridad; por otro, que estas 5 personas son todas mujeres. Como se puede observar en la siguiente tabla, Macarena Olona (Vox) obtiene la máxima polaridad en ira, miedo y tristeza; seguida de Cayetana Álvarez de Toledo (PP), que alcanza la máxima puntuación en aversión y confianza; el factor de anticipación lo encabeza Yolanda Díaz (UP); el de alegría, Isabel Díaz Ayuso (PP); y el de sorpresa, Inés Arrimadas (Cds). Se han identificado 4 casos en los que el vector medio de polaridad es negativo, es decir, que en los tuits de las cuentas de esas personas hay más mensajes con carga emocional negativa (desfavorable) que con carga emocional positiva (favorable). Se trata de Macarena Olona (Vox), Inés Arrimadas (Cd's), Irene Montero (UP) y Santiago Abascal (Vox).

Los valores agrupados por sexo y partido pueden observarse en la Figura 1, donde se aprecia que la gran mayoría de polaridad es cercana a la neutralidad (valor 0 de la figura), pero donde existen numerosos casos atípicos de gran intensidad (representados por puntos en la figura), al igual que la desviación de 1,5 con respecto al rango intercuartílico.



**Figura 1.** Análisis de polaridad por sexo y partido



Fuente: elaboración propia

Nota: diagrama de caja donde se representan los niveles de vector de polaridad por partido y sexo (H -rojo-: hombre, M -azul-: mujer)

Por lo tanto, un primer hallazgo de la investigación es que las 5 personas que concentran más carga emocional en el total de mensajes del corpus son todas mujeres, lo que viene a confirmar la segunda hipótesis (H2: La carga emocional de los mensajes que reciben las mujeres políticas españolas está más polarizada que los mensajes que reciben los políticos hombres).

Por otra parte, al trasladar el análisis de la polaridad emocional a las 5 formaciones políticas estudiadas, observamos en la Tabla 2 que Vox concentra la máxima intensidad de ira, aversión, miedo, tristeza y sorpresa; mientras que el PP se impone en anticipación y confianza; y UP, en alegría. Ciudadanos y PSOE quedan fuera de los valores máximos de polaridad emocional.

**Tabla 2.** Valores máximos de polaridad emocional en mujeres y partidos (mediana=0)

Mujeres/Partidos	Ira	Anticipación	Aversión	Miedo	Alegría	Tristeza	Sorpresa	Confianza
Macarena Olona (VOX)	0,5375	0,1827	0,4877	0,7865	0,1841	0,6108	0,2019	0,5635
Cayetana Álvarez (PP)	0,2488	0,3266	0,5406	0,3531	0,1246	0,2488	0,0508	0,9350
Yolanda Díaz (UP)	0,2175	0,3580	0,1475	0,3197	0,3241	0,2917	0,1666	0,5887
Isabel Díaz Ayuso (PP)	0,2426	0,3451	0,2673	0,4398	0,3537	0,3858	0,2390	0,5392
Inés Arrimadas (Cd's)	0,4274	0,2172	0,2812	0,5632	0,1889	0,4412	0,2422	0,6280
VOX	0,3738	0,2453	0,3399	0,5212	0,2127	0,4094	0,1692	0,5330
PP	0,2631	0,2957	0,3242	0,3846	0,2382	0,3252	0,1429	0,6195
UP	0,2788	0,2956	0,2306	0,3498	0,2882	0,3599	0,1481	0,5500
PSOE	0,2239	0,2385	0,2029	0,3228	0,2251	0,3601	0,1417	0,5609
Cd's	0,2799	0,2450	0,1991	0,4218	0,1990	0,3441	0,1535	0,6018

Fuente: elaboración propia

Así mismo, las personas que generan mayor polarización emocional en sus respectivos partidos son las siguientes de la Tabla 3:

**Tabla 3.** Políticos que generan más polarización emocional

Partidos	Ira	Anticipación	Aversión	Miedo	Alegría	Tristeza	Sorpresa	Confianza
VOX	Macarena Olona	Iván Espinosa	Macarena Olona	Macarena Olona	Iván Espinosa	Macarena Olona	Santiago Abascal	Macarena Olona
PP	Pablo Casado	Isabel Díaz Ayuso	Cayetana Álvarez	Isabel Díaz Ayuso	Isabel Díaz Ayuso	Pablo Casado	Isabel Díaz Ayuso	Cayetana Álvarez
UP	Pablo Echenique	Yolanda Díaz	Pablo Echenique	Pablo Echenique	Alberto Garzón	Pablo Echenique	Yolanda Díaz	Yolanda Díaz
PSOE	Pedro Sánchez	Miquel Iceta	Pedro Sánchez	Pedro Sánchez	Adriana Lastra	Nadia Calviño	Miquel Iceta	Nadia Calviño
Cd's	Inés Arrimadas	Edmundo Bal	Inés Arrimadas	Inés Arrimadas	Begoña Villacís	Inés Arrimadas	Inés Arrimadas	Edmundo Bal

Fuente: elaboración propia

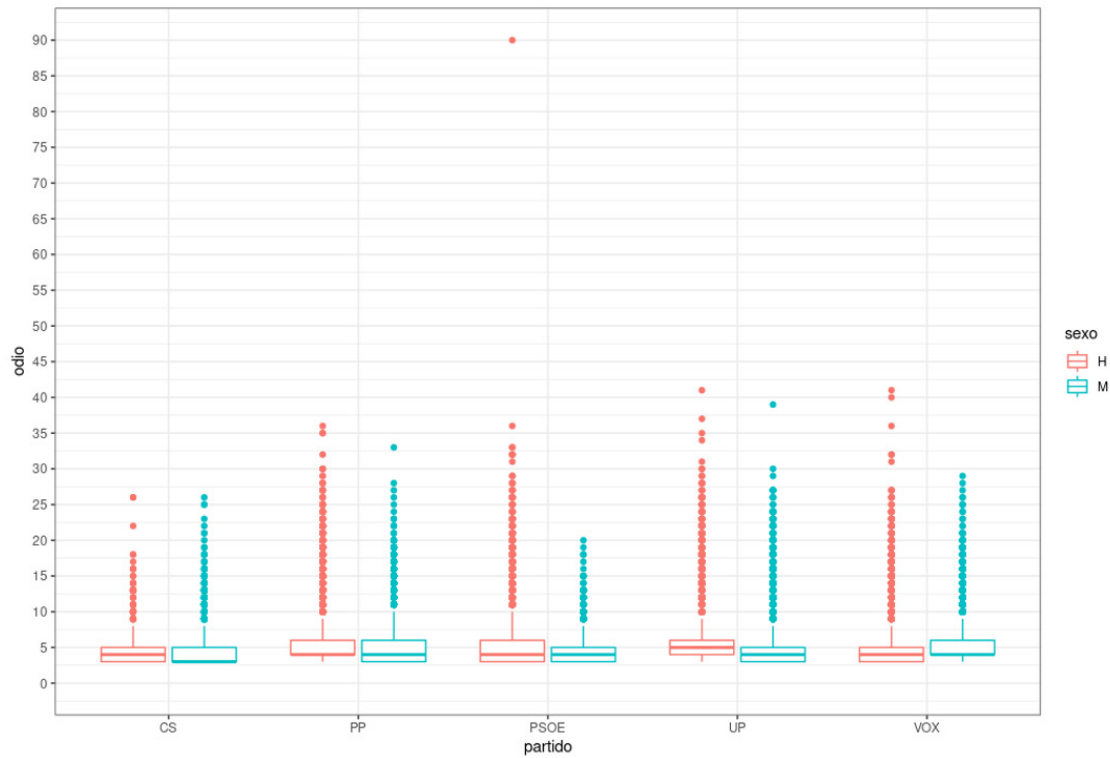
#### 4.2. Estudio del léxico (minería de texto)

La investigación sobre el odio detectado en el corpus de mensajes que integran este trabajo (3.483.232 tuits) arroja unos resultados esclarecedores para comprender la dinámica de las agresiones verbales en Twitter contra los representantes políticos. Al aplicar el léxico Hurltlex al corpus de la investigación, hemos identificado 372.777 tuits con presencia de odio intenso (10,7% del total). La variación de la intensidad de odio entre los mensajes recibidos por las mujeres y los recibidos por los hombres es inferior al 2%. Por lo tanto, los datos obtenidos no permiten afirmar que las mujeres políticas reciban más mensajes de odio que los hombres políticos.

Según puede observarse en la Figura 2, los hombres concentran los mensajes con mayor intensidad de odio según el algoritmo empleado, así como los políticos de UP y de Vox, a excepción de un único mensaje contra Pedro Sánchez cargado de insultos y odio. Al observar los valores medios del 25% al 75% del siguiente diagrama de cajas, puede apreciarse una mayor cantidad de mensajes de odio hacia los hombres de UP y hacia las mujeres de Vox.

Para averiguar si pudiera existir una artificialidad en los mensajes y su relación con el odio, se determinó la probabilidad de que fueran mensajes emitidos por bots. Para ello, se separaron todos los mensajes que tuvieran una intensidad igual o superior a un valor 12, lo que supuso un total de 11.518 tuits (de los cuales 1.157 fueron retuits) por parte de 7.821 cuentas (no se recogieron más por limitaciones de capacidad de los sistemas informáticos empleados). El algoritmo detectó un total de 2.253 cuentas con una probabilidad superior al 50% de ser un bot; de estas, 2.250 superaban el 85% de probabilidad de ser un bot. Esto supone que el 28,81% de las cuentas analizadas serían bots, y que estos bots emitieron 3.732 mensajes en total. Se llega a detectar, incluso, que muchas de esas cuentas han sido suspendidas por la propia red Twitter (p. e. @Tierraymadera2, @MiguelA99267038, @CubalibreLibre o @M31237610) o han sido mayoritariamente borradas meses después.

**Figura 2.** Análisis de odio por sexo y partido

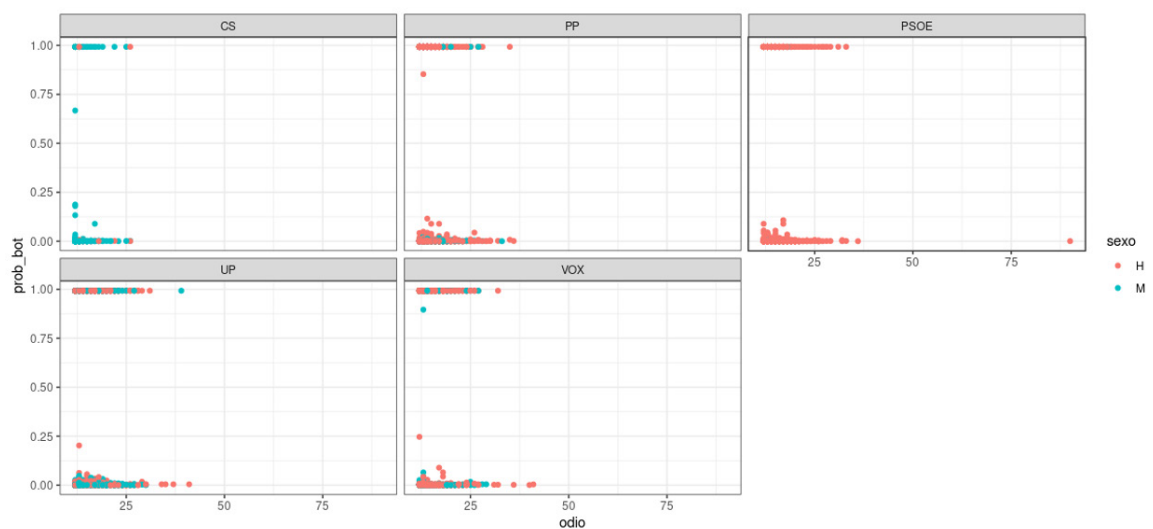


Fuente: elaboración propia

Nota: diagrama de caja donde se representan los niveles de intensidad de odio por partido y sexo (H -rojo-: hombre, M -azul-: mujer)

La representación en la Figura 3 muestra el cruce de dos variables: la probabilidad de ser un bot con la intensidad del discurso de odio por partido y sexo. Esto permite apreciar ese reparto de odio casi equitativo entre cuentas bot y las que no. Se puede observar cómo el odio extremadamente intenso hacia mujeres aparece principalmente en Ciudadanos, es significativo en Unidas Podemos, más escaso en el PP y en Vox, e inexistente hacia mujeres en el PSOE.

**Figura 3.** Representación de probabilidad bot con respecto nivel de odio



Fuente: elaboración propia

Nota: se representan los niveles de intensidad de odio con respecto a la probabilidad de ser bot por partido y sexo (H -rojo-: hombre, M -azul-: mujer)

Un análisis singularizado más específico de los tuits que mayor odio acumulan (463 de los recibidos por las mujeres y 471 por los hombres) ha permitido confeccionar un listado exhaustivo y concreto de los agravios recibidos por cada colectivo, cuyo primer resultado es que el 20% de los insultos representa a más el 80% del total, cumpliéndose el principio de Pareto. Para una mayor claridad expositiva, hemos agrupado estas agresiones verbales en tres categorías semánticas: insultos genéricos, insultos de carácter político e insultos con sesgo de género.

Como se puede observar en la Tabla 4, los descalificativos más recurrentes son insultos genéricos de uso común que, frecuentemente, aparecen acompañados de potenciadores semánticos (muy, más que, el más...) y combinados entre sí (rata miserable, mentiroso de mierda, sectaria analfabeta, etc.).

**Tabla 4.** Insultos genéricos más recurrentes

Insultos a mujeres	Frecuencia	Porcentaje	Insultos a Hombres	Frecuencia	Porcentaje
Delincuente	8.357	7,71%	Delincuente	6.708	2,54%
Ignorante	1.243	1,15%	Rata	6.688	2,53%
Gilipollas	898	0,83%	Basura	6.224	2,35%
Miserable	880	0,81%	Mentiroso	6.122	2,32%
Cobarde	744	0,69%	Mierda	5.487	2,08%
Gentuza	666	0,61%	Tonto	5.174	1,96%
Idiota	665	0,61%	Ridículo	4.144	1,57%
Rata	326	0,30%	Criminal	3.969	1,50%
Analfabeta	255	0,24%	Traidor	3.378	1,28%
Sectaria	239	0,22%	Asesino	3.349	1,27%
Incompetente	220	0,20%	Corrupto	3.230	1,22%
Subnormal	219	0,20%	Miserable	3.147	1,19%

Fuente: elaboración propia

En cuanto a los insultos de carácter político, sobresalen agresiones que se escuchan con frecuencia en boca de los propios líderes políticos, como fascista o terrorista, lo cual evidencia, por un lado, la permeabilidad social del lenguaje ofensivo utilizado por los representantes políticos; y, por otro, la devaluación semántica de conceptos que, aun siendo excepcionales (como franquista o nazi), se han convertido en lugar común de la dialéctica partidista.

Los insultos de carácter político dirigidos a las mujeres fueron, por orden decreciente: “fascista” (y sus variantes: “escoria fascista”, etc.=2.765); “comunista” (y variantes: “socialcomunista”, “neofachacommunista”, etc.=1.004); “facha” (=508), “etarra” (=214) y “nazi” (=186). En el caso de los hombres, se detectaron “terrorista” (=6.967); “fascista” (y sus variantes: “fascista de mierda”, “puto fascista”, “facha”, etc.=6.073); “socialista” (de modo despectivo=4.352); “comunista” (de modo despectivo=3.564); “nazi” (y variantes: “puto nazi”, “catalanazis” =1.076), y “franquista” (de modo despectivo=827).

En cuanto al sesgo de género, se han detectado agresiones verbales vinculadas con la misoginia y el sexismo en el bloque de mensajes recibidos por las mujeres políticas; no así en el de los hombres. Esto significa que, si bien no se puede afirmar que las mujeres concentren más mensajes de odio que los hombres, en el caso de las mujeres las agresiones verbales sí que incluyen un componente sexista, lo que confirma la primera hipótesis (H1: Las agresiones verbales recibidas por las mujeres políticas españolas presentan un componente sexista que no se percibe en las agresiones recibidas por sus homólogos varones).

Los dos grupos léxicos más recurrentes son las humillaciones de contenido sexista y las agresiones referidas al feminismo. Los insultos con sesgo de género más comunes desde el punto de vista sexista fueron “concubina” (=22), “sumisa” (=18) y “chupa pollas/chupa rabos” (=3). También se han detectado expresiones que buscan ridiculizar (no la follaba ni su marido=1), y expresiones odiosas como “vaginoplastia” (=1), “bukake” (=1) y “prostituta fascista” (=1).

Los insultos ideológicos contra el feminismo fueron: “feminista” (en tono despectivo=720), “ninistra” (=186), “feminazi” (=126), “autoridad” (=75), “ministerio de desigualdad” (=32), “femiloca” (=15) y “menistra” (=6). Llamamos la atención sobre la ausencia de la palabra “puta” pues, si bien no aparece en los mensajes recolectados con mayor nivel de odio como un insulto directo, sí lo hace ligada a expresiones cotidianas: “ni puta idea”, “puta vergüenza”, “puta calle”, “puta cuenta” o “puta vida”. Esto se puede deber a la intención de eludir la censura al tratarse de una palabra fácilmente identificable por los algoritmos de detección y de reacción rápida en contra. Sí se ha detectado la expresión “hijo/a de puta” en 274 ocasiones sobre el total de mensajes, y ninguna vez en los tuits de mayor odio.

Como se puede observar, abundan las construcciones lingüísticas basadas, por un lado, en derivados con sufijación valorativa despectiva (p. e., “gentuza”), y, por otro, la composición atributiva sintagmática (en casos como “feminista de mierda”), y léxica (en casos como “cerebroasno”).

En general, son bastante recurrentes, tanto en mujeres como en hombres, la composicionalidad (p. e. “chupa pollas”) y el neologismo despectivo procedente del inglés formado por acronimia, como “feminazi” ([femi]nist + nazi[onalsocialist]) utilizado, según la Real Academia Española (Figura 4), para designar a una feminista radicalizada<sup>2</sup>.

**Figura 4.** Definición de la voz feminazi según la RAE



Fuente: elaboración propia

Comparten el mismo origen morfológico expresiones como “ninistra” ([ni]estudia [ni]trabaja + [mini]stra) o “autoridad” ([autoridad] + morfema gramatical de género femenino [a]), lapsus gramatical cometido por Yolanda Díaz durante su intervención en el 12.º Congreso Confederado del sindicato Comisiones Obreras (CC. OO.).

En cuanto a las agresiones verbales dirigidas a los hombres, se han recolectado expresiones como “violador” (=476 veces), “misógino” (=224 veces), “homófobo” (=210 veces), “pederasta” (=119 veces), y, ocasionalmente, “dentro del armario” (=1 vez) o “menistro” (2 veces en alusión a Miquel

<sup>2</sup> Para una comprensión más profunda del neologismo “feminazi” y del discurso semiótico y simbólico que subyace en su presencia en la red social Twitter, véase el estudio de caso de la campaña “StopFeminazis”, llevado a cabo por Sara Núñez Puente y Diana Fernández Romero (2018). Discurso visual posmachista en Twitter: análisis multimodal de la iconicidad de la victimización. *Digitos*, 4, 131-148. <http://dx.doi.org/10.7203/rd.v0i4.127>

Iceta). Cabe destacar que, en menor grado, también se recogen alusiones a mujeres en conversaciones cruzadas.

Así mismo, se observa la presencia de dos subconjuntos de agresiones verbales concentradas en la persona del diputado Pablo Echenique: por un lado, las referidas a su discapacidad física (p. e. “rata motorizada”, “basura motorizada” o “engendro”), y, por otro, las de contenido xenófobo (p. e. “argentino de mierda” o “parásito de las arcas españolas”).

Por último, se han detectado numerosas alteraciones gráficas de nombres propios que, sin dejar de ser reconocibles, expresan connotaciones negativas. Es el caso de “potemos” (Podemos), “congreso de los imputados” (Congreso de los Diputados), “corruPPto” (en alusión al Partido Popular), “Perro Sánchez” (Pedro Sánchez), “cum fraude” (cum laude, en alusión a los supuestos títulos universitarios fraudulentos obtenidos por determinados políticos).

## 5. Discusión y conclusiones

Las agresiones verbales contra mujeres y hombres políticos encontradas en el marco de esta investigación deben ser considerados actos de habla (Searle, 2009), es decir, locuciones emitidas con una intención, en un contexto de comunicación pública como es la red social Twitter, sobre la base de una evidente actitud prejuiciosa. La mayor parte de las agresiones detectadas equivalen al segundo peldaño de la pirámide del odio, donde se concentran los “actos de prejuicio” (lenguaje no inclusivo, comentarios desconsiderados, microagresiones, contar bromas/chistes denigrantes, apropiación cultural, anulación y/o exclusión social, insultos, ridiculización, intimidación, epítetos, deshumanización) (ADL, 2017).

No cabe duda de que los discursos analizados están presididos por un elevado índice de odio que es similar en hombres y en mujeres (un 10,7% del total de tuits analizados contenían odio intenso). Sin embargo, no se ha descubierto un mayor encarnizamiento contra las mujeres por el hecho de ser mujeres, aunque sí se ha detectado una variación sexista y misógina significativa en las agresiones contra las mujeres. Estos hallazgos son consistentes con investigaciones previas (Herrero-Diz et al., 2020; Piñeiro-Otero y Martínez-Rolán, 2021; Sarapura, 2021), lo que vendría a reforzar los resultados de este trabajo.

De aquí se puede concluir que las mujeres reciben un importante volumen de insultos por su condición de mujeres más que por su condición de políticas, como se ha observado también en estudios cuantitativos realizados en otros países (*The Economist Intelligence Unit*, 2020). La descalificación permanente de la mujer con palabras sexistas y misóginas acaba creando una ontología de significados que permea poco a poco en la esfera pública hasta convertirla en un espacio tóxico carente de las más elementales reglas de civismo y convivencia.

Igualmente, son ellas las que soportan mayor carga emocional en las interacciones con los usuarios, lo que, si bien no está ligado con el odio, sí sugiere que las mujeres despiertan en la audiencia respuestas emocionales más intensas.

Por otra parte, el discurso del odio contra las mujeres puede ser considerado un tipo de desorden informativo en la medida en que busca mermar su influencia pública y afectar al feminismo como causa. Los ataques al feminismo observados en las agresiones analizadas en esta investigación establecen una narrativa que condiciona la agenda sobre otras cuestiones políticas (Herrero-Diz, Pérez-Escolar y Plaza Sánchez, 2020; Sessa, 2020). En su desarrollo, se emplean neologismos despectivos como “feminazi”, recogido más de 700 veces en este trabajo; acrónimos derivativos como “ninistra”, que restan autoridad y relevancia al trabajo femenino, así como composiciones atributivas sintagmáticas como “feminista de mierda” o “gentuza de género”.

Llama la atención que la palabra “puta” apenas se ha detectado en el corpus como agresión directa, y sí otros insultos menos habituales como “concubina”. Dado que se trata de una palabra fácilmente detectable por los algoritmos de control de la conversación en las redes sociales, su ausencia podría estar indicando que estamos ante mensajes fabricados artificialmente por bots, configurados para eludir la censura. Los hallazgos de esta investigación parecen sugerir esta hipótesis, ya que en el bloque de mensajes con un índice de odio superior a 12 se ha descubierto que 2.250 cuentas superaban el 85% de probabilidad de ser un bot. Esto supone que el 28,81% de las cuentas que más odio emitieron serían bots, que emitieron un total de 3.732 mensajes.

Las limitaciones técnicas de los sistemas informáticos empleados en esta investigación han impedido ampliar la búsqueda de bots al total del corpus de la investigación, por lo que no estamos en disposición de calcular qué porcentaje de cuentas podrían estar siendo operadas artificialmente. Queda así planteada una de las limitaciones de este estudio, que a su vez invita a profundizar en una línea de investigación que relacione la propagación de discursos de odio con la presencia de bots en las redes sociales. De ser así, futuros trabajos tendrían que ayudar a discernir en qué medida el fomento del odio en las redes responde a estrategias sistemáticas de polarización política y radicalización social, más que a la actitud individual de cierto número de individuos que usan las redes para insultar a los políticos.

La limitación técnica también impacta en la detección del odio en los mensajes analizados. En español es un sistema muy novedoso y, aunque existen proyectos en diversas universidades españolas, todavía no son públicos. Para superar esta carencia, la presente investigación se ha basado en una de las primeras soluciones algorítmicas de detección de odio en español, que presenta como mayor problema la falta de calificación de intensidad de los distintos insultos considerados por separado. Sin embargo, este algoritmo sí que tiene en cuenta la conjunción de diversas palabras que se potencian entre sí. Además, la biblioteca de odio de partida, Hurltlex, ha conseguido un gran porcentaje de éxito en español, según sus autores. Las bibliotecas utilizadas de emociones algorítmicas en lenguas latinas (NRC) tienen una fiabilidad menor que en inglés (superior al 80% de certeza), pero que, no obstante, rondan el 70% (Mohammad, 2016).

Hechas estas salvedades, la presente investigación permite asegurar que los discursos de odio contra los políticos españoles representan el 12,15% del corpus, y que los mensajes con mayor intensidad de odio suponen el 10,7% del total. No se puede subestimar, por tanto, la gravedad de estas interacciones que convierten la conversación social en torno a los políticos en un escenario altamente polarizado y sensible. Esta realidad sugiere la necesidad de seguir investigando los discursos de odio, su alcance y su impacto social. En la medida en que el sexismo y la misoginia aparecen recurrentemente en las agresiones verbales contra las mujeres políticas, hay que avanzar en el estudio de la relación del discurso de odio con las estrategias de desinformación y los desórdenes informativos.

En un efecto espejo, tampoco se puede obviar que la virulencia de los ataques verbales analizados en esta investigación invita a una reflexión inaplazable sobre la responsabilidad de los propios políticos en la radicalización y polarización del clima social. Las agresiones identificadas forman parte cotidiana de la dialéctica partidista; con demasiada frecuencia, descalificaciones como fascista o terrorista son escuchadas en sede parlamentaria. En este sentido, no cabe alarmarse cuando el habla popular se contamina del mismo lenguaje abusivo que emplean sus propios representantes políticos.

La misma responsabilidad cabe solicitar a los propietarios de las plataformas que sustentan redes sociales como Twitter. El hecho de que ciertos perfiles promotores de mensajes de odio misóginos y sexistas sigan operativos meses después de ser denunciados (Spring, 2021), alimenta la impunidad de los agresores y deja la vía expedita para la propagación de conductas antidemocráticas y de discursos de odio que merman la convivencia y la paz social.

Las redes sociales son, simultáneamente, extensión y complemento de la vida física. Por ello es necesario insistir en la idea de que lo que sucede en la esfera virtual adquiere naturaleza real y performativa en la vida de las personas.

## Apoyos

Este artículo forma parte del proyecto “Nuevos escenarios de vulnerabilidad digital: alfabetización mediática para una sociedad inclusiva” (PROVULDIG-2), financiado por la Comunidad de Madrid y el Fondo Social Europeo (ref. H2019/HUM-5775).

## Bibliografía

- Amnistía Internacional (2021). *Troll patrol*. <https://bit.ly/3nhoBBr>
- Amores, J. J., Blanco-Herrero, D., Sánchez-Holgado, P., & Frías-Vázquez, M. (2021). Detectando el odio ideológico en Twitter. Desarrollo y evaluación de un detector de discurso de odio por ideología política en tuits en español. *Cuadernos.Info*, 49, 98–124. <https://doi.org/10.7764/cdi.49.27817>
- Anti-defamation League (ADL) (2017). *Pyramid of Hate*. <https://bit.ly/3A0MyV9>
- Arce-García, S., & Menéndez-Menéndez, M.I. (2018). Aplicaciones de la estadística al framing y la minería de texto en estudios de comunicación. *Información, cultura y sociedad*, 39, 61-70. <https://doi.org/10.34096/ics.i39.4260>
- Arcila-Calderón, C., Sánchez-Holgado, P., Quintana-Moreno, C., Amore, J-J., Blanco-Herrero, D. (2022). Discurso de odio y aceptación social hacia migrantes en Europa: Análisis de tuits con geolocalización. *Comunicar*, 71, 21-35. <https://doi.org/10.3916/C71-2022-02>.
- Avaaz (2019). *Whatsapp Social Media's dark web. How the messaging service is being flooded with lies and hate speech ahead of the spanish elections*. Avaaz.org. <https://bit.ly/3FIViyY>
- Barrie, C., & Chun-ting, H. (2021). AcademictwitteR: an R package to access the Twitter Academic Research Product Track v2 API endpoint. *Journal of Open Source Software*, 6(62), 3272. <https://doi.org/10.21105/joss.03272>.
- Bassignana, E., Basile, V., & Patti, V. (2018). *Hurtlex: A multilingual Lexicon of Words to Hurt*. Proceedings of the Fifth Italian Conference on Computational Linguistics CLiC-it 2018. Collana dell'Associazione Italiana di Linguistica Computazionale. <https://doi.org/10.4000/books.aaccademia.3085>
- Bernal-Triviño, A., & Clares-Gavilán, J. (2019). Uso del móvil y las redes sociales como canales de verificación de fake news. El caso de Maldita.es. *Profesional de la información*, 28(3), <https://doi.org/10.3145/epi.2019.may.12>
- Blanco-Alfonso, I., Chaparro-Domínguez, M., & Repiso, R. (2021). El fact-checking como estrategia global para contener la desinformación. *Estudios sobre el Mensaje Periodístico*, 27(3), 779-791. <https://doi.org/10.5209/esmp.76189>
- Burnap, P., & Williams, M.L. (2015). Cyber hate speech on Twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & Internet*, 7(2), 223-242. <https://doi.org/10.1002/poi3.85>
- Campos-Domínguez, E., Esteve-Del-Valle, M. Renedo-Farpón, C. (2022). Retóricas de desinformación parlamentaria en Twitter. *Comunicar*, 72, 47-58. <https://doi.org/10.3916/C72-2022-04>
- Cid, G. (2019). Los ‘deepfakes’ ya son un negocio millonario: cómo evitar que te engañen. *El Confidencial*. <https://bit.ly/3wnOpjb>
- DW (2021). Alemania: desinformación y noticias falsas asedian la campaña electoral. *DW.com* <https://bit.ly/31H21e6>
- Ferruz-González, S.A., Rodríguez-Fernández, L., & Rubio-Romero, J. (2020). Visibilidad del liderazgo femenino en Twitter: aproximación sobre el ranking Top 100. Mujeres líderes en España. *Revista Mediterránea de Comunicación*, 12(2), 45-64. <https://doi.org/10.14198/MEDCOM.18839>



- Gagliardone, I., Gal, D., Alves, T., & Martínez, G. (2015). *Countering Online Hate Speech. Programme in Comparative Media Law and Policy*. University of Oxford.
- Garrido, A., Martínez Rodríguez, M. A., & Mora Rodríguez, A. (2021). Polarización afectiva en España. *Más Poder Local*, 45, 21-40. <https://bit.ly/3HM8uoF>
- Gidron N., Adams, J., & Horne, W. (2018). How Ideology, Economics and Institutions Shape Affective Polarization in Democratic Polities. *American Political Science Association*. <https://bit.ly/2ZrNGS4>
- Gil Pascual, J. A. (2021). *Minería de textos con R. Aplicaciones y técnicas estadísticas de apoyo*. UNED.
- Guerrero-Solé, F., & Philippe, O. (2020). La toxicidad de la política española en Twitter durante la pandemia de la COVID-19. *Hipertext.net*, 21, 133-139. <https://doi.org/10.31009/hipertext.net.2020.i21.12>
- Herrero-Diz, P., Pérez-Escobar, M., & Plaza Sánchez, J.F. (2020). Desinformación de género: análisis de los bulos de Maldito Feminismo. *Icono 14*, 18(2), 188-216. <https://doi.org/10.7195/ri14.v18i2.1509>
- Igareda-González, N., Alberto-Pascale, A., Fejös, A., Zentai, V., Cankurt, I., Sette, R., Sicurella, S., Schulz, A., Wilhelm, C., Bladini, M., Helmstad, A., & Svensson, E-M. (2021). *Map of the main political parties using hate speech against the “ideology of gender” through the social networks and internet in Europe*. <https://bit.ly/3HV3L4A>
- Ipsos (2021). *Sentimiento de sistema roto en 2021. Populismo, anti-elitismo y nativismo*. <https://bit.ly/3xmAHxs>
- Jankowicz, N. (2017). How Disinformation Became a New Threat to Women. *codastory.com*. <https://bit.ly/3x3ukPh>
- Jockers, M. (2017). *Syuzhet, extracts sentiment and sentiment-derived plot arcs from text*. <https://bit.ly/3bmyTxx>
- Kearney, M. (2020). *Detect Twitter Bots: tweetbotornot2*. <https://bit.ly/3HZtpoM>
- Liddy, E.D. (1998). Enhanced text retrieval using natural language processing. *Bolletín of the American Society for Information Science and Technology*, 24(4), 14-16. <http://doi.org/10.1002/bult.91>
- Lim, M. (2017). Freedom to hate: social media, algorithmic enclaves, and the rise of tribal nationalism in Indonesia. *Critical Asian Studies*, 49(3), 411-427. <https://doi.org/10.1080/14672715.2017.1341188>
- Luque-Ortiz, S., & Cano-Alarcón, M. (2020). La violencia de género en Twitter según Vox en las elecciones autonómicas de Andalucía. *Ámbitos. Revista Internacional de Comunicación*, 51, 116-133. <https://doi.org/10.12795/Ambitos.2021.i51.08>
- Martini, F., Samula, P., Keller, T.-R., & Klinger, U. (2021). Bot, or not? Comparing three methods for detecting social bots in five political discourses. *Big Data & Society*, 8(2), 1-13. <https://doi.org/10.1177/20539517211033566>
- Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online*. Data & Society. <https://bit.ly/3CS4v70>
- Mohammad, S. (2016). Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. *Emotion Measurement*, 2016, 201-237. <https://doi.org/10.1016/B978-0-08-100508-8.00009-6>
- Mohammad, S., & Turney, P. (2010). *Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon*. Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text. June 2010. LA, California.
- Mohammad, S., & Turney, P. (2013). Crowdsourcing a Word-Emotion Association Lexicon. *Computational Intelligence*, 29(3), 436-465. <https://doi.org/10.1111/j.1467-8640.2012.00460.x>
- Moreno Cantano, A.C., & Calvillo Cisneros, J.M. (2021). Propaganda del odio contra los Rohingya. Estudio del caso de Facebook y Twitter. Pineda, A., & Sánchez Gutierrez, B. (eds.), *Comunicación política en el mundo digital: tendencias actuales en propaganda, ideología y sociedad* (pp. 792-814). Dykinson.
- Núñez Puente, S., & Fernández Romero, D. (2018). Discurso visual posmachista en Twitter: análisis multimodal de la iconicidad de la victimización. *Dígitos*, 4, 131-148. <http://dx.doi.org/10.7203/rd.v0i4.127>
- Oceño, P. (2018). Fake news real women: Disinformation gone macho. When disinformation is weaponized against women, misogyny rears its ugly head. *Rappler*. <https://bit.ly/3CV5rYi>

- ONU Mujeres (2021). Hechos y cifras: Liderazgo y participación política de las mujeres. *Unwomen.org*. <https://bit.ly/3F1NVIE>
- Pardy, N. (2018). How The Fake News Industry Weaponizes Women. *Refinery29*. <https://r29.co/3D0CpXo>
- Parekh, B. (2006). Hate speech. Is there a case of banning? *Public Policy Research*, 12(4), 213-223. <https://doi.org/10.1111/j.1070-3535.2005.00405.x>
- Piñeiro-Otero, T., & Martínez-Rolán, X. (2021). Eso no me lo dices en la calle. Análisis del discurso del odio contra las mujeres en Twitter. *Profesional De La Información*, 30(5), 1-17. <https://doi.org/10.3145/epi.2021.sep.02>
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. Plutchik, R. & H. Kellerman (eds.), *Emotion: Theory, Research, and Experience* (pp. 3-33). Academic Press.
- Posetti, J., Shabbir, N., Maynard, D., Bontcheva, K., & Aboulez, N. (2021). Online violence Against Women Journalists: A Global Snapshot of Incidence and Impacts. *UNESCO*. <https://bit.ly/3ocsSWQ>
- Ramírez-García, A., González-Molina, A., Gutiérrez-Arenas, M-P., & Moyano-Pacheco, M. (2021). Interdisciplinariedad de la producción científica sobre el discurso del odio y las redes sociales: Un análisis bibliométrico. *Comunicar*, 72, <https://doi.org/10.3916/C72-2022-10>
- Ramírez Plascencia, D., Alonzo González, R. M., & Ochoa Amezcua, A. (2022). Odio, polarización social y clase media en Las Mañaneras de López Obrador. *Doxa Comunicación*, 35. <https://doi.org/10.31921/doxacom.n35a1505>
- Rivas Venegas, M. (2021). Contra las mujeres: el discurso misógino de Vox. “Palabras gruñido” del nacionalpopulismo español. *Revista de Investigaciones Feministas*, 12(1), 67-77. <https://dx.doi.org/10.5209/infe.69585>
- Robles, J.M., Guevara, J.A., Casas-Mas, B., & Gómez, D. (2022). Cuando la negatividad es el combustible. Bots y polarización política en el debate sobre el COVID-19. *Comunicar*, 71, 63-75. <https://doi.org/10.3916/C71-2022-05>
- Rodríguez-Fernández, L. (2021). *Propaganda digital: comunicación en tiempos de desinformación* UOC.
- Sarapura Sarapura, M. M. (2021). El discurso del odio sexista a través de las redes sociales como reacción al himno “Un violador en tu camino”. *Razón y Palabra*, 24(111), 116-140. <https://doi.org/10.26807/rp.v25i111.1807>
- Sauter, D.A., Eisner, F., Ekman, P., & Scott, S. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, 107(6), 2408-2412. <https://doi.org/10.1073/pnas.0908239106>
- Searle, J. (2009). *Actos de habla*. Cátedra.
- Sessa, M.G. (2020). Misogyny and Misinformation: An analysis of gendered disinformation tactics during the COVID-19 pandemic. *Disinfo.eu*. <https://bit.ly/3EY8QpP>
- Smith, M. A., Rainie, L., Shneiderman, B., & Himelboim, I. (2014). Mapping Twitter topic networks: From polarized crowds to community clusters. *The Pew Research Center*. <https://pewrsr.ch/3xUK203>
- Spring, M. (2021). La periodista de la BBC brutalmente troleada en redes sociales (y las impactantes revelaciones de su experimento). *BBC.com*. <https://bbc.in/3qfMqfN>
- Stabile, B., Grant, A., Purohit, H., & Harris, K. (2019). Sex, Lies, and Stereotypes: Gendered Implications of Fake News for Women in Politics. *Public Integrity*, 21, 491-502. <https://bit.ly/31LSKBx>
- The Economist Intelligence Unit (2020). *Measuring the prevalence of online violence against women*. <https://bit.ly/3qfA8ns>
- Unión Interparlamentaria (2016). Sexism, harassment and violence against women parliamentarians. *ipu.org*. <https://bit.ly/3o6nIf4>
- Unión Interparlamentaria (2018). Sexism, harassment and violence against women in parliaments in Europe. *ipu.org*. <https://bit.ly/308aw12>
- Williams, M. (2021). *The Science of Hate*. Faber&Faber.