

Una mirada dialectològica al Observatorio de Palabras de la RAE vía Twitter

JULIA ELENA PARDO
Universitat Pompeu Fabra
jpardo.al@gmail.com

Julia Elena Pardo és graduada en Enginyeria per la Universitat de Buenos Aires i màster en Lingüística Teòrica i Aplicada per la Universitat Pompeu Fabra. El seu treball de fi de màster sobre dialectologia computacional i prescriptivisme popular va brindar les bases metodològiques per a aquest treball de terminologia. Els seus interessos se centren en la variació lingüística i la lingüística popular, usant la ciència de dades com a eina principal.



Resumen

El Observatorio de Palabras recoge términos —fundamentalmente, neológicos— que no aparecen en el Diccionario de la lengua española y que han generado dudas de los hispanohablantes a la Real Academia Española (RAE). Su función normativa, sin embargo, se presta a confusión. Con el objeto de esclarecerla, este artículo toma algunos de estos términos y analiza la evolución de su uso y su variación diatópica en el ámbito panhispánico, utilizando Twitter como corpus dialectológico. Se satisfacen los criterios de neologicitad y se sugiere el acercamiento de la Academia al hablante. PALABRAS CLAVE: RAE; Twitter; neología; dialectología panhispánica

Resum

Una mirada dialectològica a l'Observatorio de Palabras de la RAE via Twitter

L'Observatorio de Palabras recull termes —fonamentalment, neològics— que no apareixen en el Diccionario de la lengua española i que han generat dubtes dels hispanoparlants a la Real Academia Española (RAE). Tanmateix, la seva funció normativa porta a confusió. A fi d'esclarir-la, aquest article pren alguns d'aquests termes i n'analitza l'evolució pel que fa a l'ús i la variació diatòpica en l'àmbit panhispànic, utilitzant Twitter com a corpus dialectològic. Se satisfan els criteris de neologicitat i se suggereix l'acostament de l'Acadèmia al parlant. PARAULES CLAU: RAE; Twitter; neologia; dialectologia panhispànica

Abstract

A dialectological look at RAE's Observatorio de Palabras on the basis of Twitter

The Observatorio de Palabras («Word Observatory») collects terms – chiefly neologisms – that are not in the Diccionario de la lengua española («Dictionary of the Spanish language») and that have led Spanish speakers to convey their respective doubts to the Real Academia Española (Royal Spanish Academy, RAE). The Word Observatory's normative role, however, gives rise to confusion. Aiming to clarify it, this article takes some of these terms and analyses the evolution of their use and their diatopic variation at a pan-Hispanic level, using Twitter as a dialectological corpus. Neogicity criteria are met and a closer approach of the Academy to the speakers of the language is suggested. KEYWORDS: Twitter; neology; pan-Hispanic dialectology

TERMINÀLIA 27 (2023): 32-50 · DOI: 10.2436/20.2503.01.187
Data de recepció: 19/10/2022. Data d'acceptació: 05/04/2023
ISSN: 2013-6692 (impresa); 2013-6706 (electrònica) · <http://terminalia.iec.cat>

1 Introducción

El Observatorio de Palabras (OdP) fue presentado por la Real Academia Española (RAE) en su página web el 23 de octubre de 2020. Allí se describía de la siguiente manera (RAE, 2020a):

El «Observatorio de palabras» ofrece información sobre palabras (o acepciones de palabras) y expresiones que no aparecen en el diccionario, pero que han generado dudas: neologismos recientes, extranjerismos, tecnicismos, regionalismos, etc. Esta información es provisional, pues no está contemplada en las obras académicas, por lo que puede verse modificada en el futuro. La presencia de un término en este observatorio no implica que la RAE acepte su uso.

En tan solo unos pocos días, era noticia en todos los países de habla hispana: el pronombre *elle* se encontraba entre el conjunto de palabras descritas, y la confusión generalizada sobre su grado de aceptación y su estado de evaluación por parte de la RAE inundó la prensa de todas las ideologías. Tal fue la repercusión que la RAE pasó —de inmediato— a eliminar la entrada del controvertido pronombre. Alegó que se volvería a valorar su inclusión «cuando se difundiera ampliamente el funcionamiento y cometido de esta sección» (RAE, 2020b). Hoy día, casi dos años más tarde, la descripción permanece intacta (RAE, 2022). Sabemos, por respuestas que ha dado la RAE en Twitter a usuarios que consultaban, que las setenta y tres palabras que allí aparecen fueron incluidas en el 2020 (RAE, 2021a), y que su objetivo «no es recoger las palabras que están en fase de incorporación al diccionario, sino resolver dudas relacionadas con algunas de las palabras ausentes del DLE» (RAE, 2021b).

La introducción del OdP representa una innovación dentro de los procedimientos habituales de la RAE; sin embargo, el reclamo de una mejor especificación de su cometido y una adecuada distinción académica entre neologismos y extranjerismos, tecnicismos y localismos permanece vigente (García Platero, 2022, p. 29). En efecto, las entradas del OdP no poseen clasificación alguna. Más aún —favoreciendo la confusión general—, algunas de ellas pueden encontrarse en el *Diccionario de la lengua española* (DLE) (RAE y ASALE, 2022), en el *Diccionario panhispánico de dudas* (DPD) (RAE y ASALE, 2005) y en el *Diccionario de americanismos* (DA) (RAE y ASALE, 2010). En el OdP, la descripción de cada entrada detalla si son o no voces válidas, qué uso o usos son los sugeridos o recomendados, y cuáles se desaconsejan, sea por tratarse de extranjerismos, sea por ser considerados innecesarios, al existir otra variante en español. Esto último acarrea la inquietud respecto al prestigio de la variante española preferida: ¿se trata de la variante globalmente preponderante o es propia de ciertas regiones? ¿Cómo se condicen las preferencias y sugerencias que da la RAE con los usos y tendencias populares?

Lo cierto es que las palabras que conforman el OdP han generado dudas entre los hablantes del idioma y, por ello, se encuentran compiladas en este portal lingüístico. Estos hablantes pertenecen a una comunidad de casi seiscientos millones de personas (Instituto Cervantes, 2022, p. 23), geográficamente extensa y rica en variación dialectal. El español es concebido actualmente como una lengua pluricéntrica: existen muchos estándares de la lengua, con diferente alcance regional, pero en su mayoría carecen de apoyo institucional, de una codificación formal (Amorós-Negre, 2018, pp. 57-58). Estas variedades ejemplares constituyen gradaciones no solo desde el punto de vista diatópico, sino también diastrático y diafásico en el plano comunicativo. La norma española vigente, en tanto, describe y prescribe un estándar composicional (Senz et al., 2011, § 3.5.4), una única variedad formalmente codificada, entendida como una «lengua culta general panhispánica» (Amorós-Negre, 2018, p. 63). Como tal, tiene un carácter supradialectal que garantiza la comunicación entre la extensa comunidad de hispanohablantes (Amorós-Negre, 2018, pp. 48-49).

La actual norma española es el reflejo de una política lingüística reciente. Si bien la Asociación de Academias de la Lengua Española (ASALE) fue fundada en 1951 con el objeto de iniciar una política colaborativa entre todas las academias de la lengua (Paffey, 2021, p. 244), no fue hasta finales de siglo que se produjo una colaboración genuina (Amorós-Negre, 2018, p. 60) y hasta 2004 que se materializó en la «nueva política lingüística panhispánica». A la fecha, la ASALE está integrada por veintitrés academias de América, España, Filipinas y Guinea Ecuatorial (ASALE, 2022), pero es la RAE quien ha logrado mantener y naturalizar su papel público como autoridad de referencia de la lengua, representativa de todo el mundo hispanohablante (Paffey, 2021, p. 250; Valle, 2007, p. 88). Es a ella a quien el hablante común acerca sus dudas, y Twitter es uno de los medios de interacción, donde —desde 2012, como @RAEinforma— la RAE responde consultas y recibe todo tipo de opiniones y sugerencias.

En efecto, para el hablante común el buen lenguaje no es el reflejo de una variante prestigiosa, elegida y expresada para un estándar normativo, se trata de un ente más abstracto, que es bueno porque se presupone más lógico, claro y etimológicamente continuo (Niedzielski y Preston, 2000, p. 18). Así, acude a diccionarios y gramáticas como una incursión a tal abstracción del lenguaje y no como a meras autoridades sobre su uso. Las obras normativas trascienden su carácter descriptivo y se establecen como autoridad prescriptiva; por ello, la entrada de una palabra al diccionario le confiere un aval de garantía y de prestigio (Parakrama, 1995, pp. 10-12 y 28-29; Prieto de los Mozos, 1999, p. 249; Guerrero Ramos, 2022, pp. 191-193). Sin embargo, el OdP no es un diccionario, ni es parte de la norma; tampoco es necesariamente un estadio inter-

medio (aunque lo haya sido en algunos casos). Pero es natural que, siendo este portal lingüístico presentado por la autoridad prescriptiva, haya generado confusión en cuanto a la validación que concede a sus entradas.

Pero ¿por qué estos términos agrupados en el ODP han sido cuestionados? ¿Cómo ha sido su uso en los últimos años? ¿Se observan diferencias entre los distintos países de habla hispana? ¿Podrían algunos de estos términos ser incorporados al diccionario? ¿Qué características o comportamientos encontramos para los ya aceptados?

El objetivo de este trabajo es, tomando como base una selección de las entradas del ODP y utilizando Twitter como corpus dialectológico (véase § 3), establecer y analizar la evolución de sus frecuencias de uso y su distribución diatópica en los distintos países hispanohablantes; sean o no estos términos eventualmente aceptados, se busca compararlos con los que ya lo han sido (en la última edición electrónica del DLE) y observar cómo se ajustan las recomendaciones de la RAE a los usos de la gente corriente en su hablar cotidiano. Se pretende, en definitiva, esclarecer la función que cumple este observatorio.

2 Las nuevas voces del español

La variación y los cambios lingüísticos son inherentes a cualquier lengua viva; la reacción a estos cambios por parte de los hablantes también es natural y se encuentran en tensión permanente la conservación y la innovación del idioma (Jiménez Ríos, 2015, p. 46). Para que un cambio dado logre estabilidad y sea, posteriormente, incorporado a gramáticas y diccionarios, requiere que se alcance sobre él un consenso social, que se manifiesta tanto en la frecuencia de uso como en su perdurabilidad (Lorente Casafont, 2022, p. 35). Pero, tal como lo indica Guerrero Ramos (2022), aun si la neologicidad de una voz puede medirse con cierta objetividad según el criterio lexicográfico, su incorporación al diccionario depende tanto de la aptitud como de la ideología del equipo de redacción involucrado (p. 191). En efecto, las opiniones de los lexicógrafos respecto a la idoneidad o conveniencia de utilizar ciertos neologismos —en particular, la discriminación entre lo necesario y lo superfluo— difícilmente están alineadas (p. 192). Y el hablante —incluso a pesar de su naturaleza creativa— recurre al diccionario (o a la autoridad) en busca de la validación del uso de estas nuevas voces (p. 209). Al igual que los expertos, los hablantes difieren en sus opiniones, y tales discrepancias, sus expresiones de acuerdo y desacuerdo para con la RAE y entre ellos mismos, quedan manifestadas en los hilos públicos de Twitter de @RAEinforma, por lo que intervienen en la regulación de los usos, con prácticas prescriptivas populares o de base (Lukač, 2018, p. 5).

La variación del español —diatópica, diastrática y diafásica— se refleja, naturalmente, también en la

neología. Así, se encuentran voces nuevas comunes a las distintas variedades, otras propias a ciertas regiones y también voces nuevas distintas que responden a un mismo significado (Fernández-Silva, 2022, p. 259). Mientras la distancia geográfica y las diferencias culturales, sociopolíticas y económicas entre las comunidades influyen en la diferenciación entre las lenguas locales, factores externos como los medios masivos, la urbanización, la movilidad global y las redes sociales favorecen la interacción lingüística de los hablantes (Amorós-Negre, 2018, p. 65; Sánchez Lobato, 2019, pp. 68-69). Son ellos, en última instancia, los responsables de facilitar y preservar su comprensión mutua (Lara, 2015, pp. 27-28, citado en Amorós-Negre, 2018, p. 65), valorando sus similitudes y peculiaridades lingüísticas (Sánchez Lobato, 2019, p. 68). La innovación lingüística del español se encuentra, pues, en una tensión entre los procesos de variación, propiciados por los fenómenos locales, y los de homogeneización, fomentados por las dinámicas globales (Fernández-Silva, 2022, p. 262). Las interacciones en Twitter entre los hispanohablantes y la RAE reflejan estas tensiones, así como las variadas actitudes y opiniones respecto a las nuevas voces: «consulta» (Carla P., 2014), «sugerencia» (Arnandis, 2017), «aprobación» (Negrete-Portillo, 2018), «crítica» (exYge, 2020), «rechazo» (Daluna MX, 2018) y «sorpresa ante usos regionales sugeridos» (Felipe el Iguano, 2019).

Las obras normativas, por un lado, favorecen la homogeneización en la presentación de una norma común panhispánica que, simultáneamente, procura reconocer el pluricentrismo empírico de sus variedades. A pesar de las interpretaciones divergentes que genera el trasfondo ideológico de la política lingüística, el carácter compositivo del estándar actual se ha apartado de sus antiguas formas eurocéntricas para acercarse a las de sus hablantes (Greußlich, 2015, pp. 59-60; Amorós-Negre, 2018, p. 308). Sin embargo, aún es notorio el tratamiento diferenciado de las variantes no peninsulares como subsidiarias o marcadas, pues sigue vigente la distinción centralista entre el «diccionario general» (el DLE) y los «diccionarios de regionalismos», en particular, el DA (Lara, 2005, p. 182). Por un lado, estos diccionarios diferenciales o de -ismos solo incluyen lo disímil y periférico, no son autosuficientes y apelan a un público especializado y no al general; por el otro, hay aún una falta de representación de estos usos «marcados» en los diccionarios generales, que pretenden abarcar la lengua en su totalidad (Ávila, 2004, pp. 13-14; Fernández-Silva, 2022, pp. 259-260). Similarmente, los diccionarios de neologismos tampoco han mostrado un tratamiento prioritario hacia el eje diatópico. No obstante, tratándose el español de una lengua de notable expansión geográfica e histórica, con gran vitalidad léxica, la variación diatópica no puede ser desatendida, pues tiene inevitable influencia en las formas y las motivaciones de la producción neológica (Fernández-Silva, 2022, pp. 260-261).

Como observa Freixa (2022) en su estudio del proceso neológico, la institucionalización de una voz —entendida como la difusión y establecimiento social de su uso— no necesariamente culmina con su incorporación al diccionario (p. 57), ya que los criterios lexicográficos pueden primar sobre la evolución en su uso (p. 71). Por ello, Freixa (2022) sugiere que la solución más apropiada es analizar frecuencias de uso en grandes corpus de textos de la lengua general, por ejemplo, mediante histogramas (p. 58). Estos revelan distintos patrones en la evolución de neologismos; así por ejemplo, los candidatos a incorporarse al diccionario suelen compartir cinco características: no tender al desuso, estar en uso desde hace al menos cuatro años, tener más de diez ocurrencias, aparecer en variados tipos de textos y hacerlo en textos provenientes de variedades geolectales diferentes (pp. 64-65).

Cabe notar también la influencia de la variación diafásica en los neologismos. Como indica Sanmartín (2022), estos se originan y utilizan en distintos ámbitos comunicativos y discursivos, entre los que destacan, sobre todo, el léxico de especialidad y el léxico coloquial (pp. 234-236). La prensa, que es una fuente habitual para el estudio de neologismos, suele tener un registro más formal, aunque en ocasiones recurre a usos informales como recurso estilístico (pp. 241 y 257). Si bien, en el plano léxico, es más sencillo dar a una nueva voz una marca coloquial o neutra que una formal, debe verificarse como tal en un corpus coloquial o en redes sociales (p. 257).

3 Twitter como corpus dialectológico

El estudio computacional sobre cambio y variación lingüística mediante el uso de las redes sociales ha tenido un gran crecimiento en los últimos años (véase la revisión realizada por Nguyen *et al.* [2016] sobre sociolingüística computacional). Dado que el medio de comunicación no necesariamente define la concepción oral o escrita (Koch y Oesterreicher, 2007), el lenguaje de las redes sociales —a pesar de tratarse de un medio gráfico— puede presentar numerosas características de la oralidad, con lo que consigue gran heterogeneidad y es sumamente útil para el estudio variacional (Estrada Arráez y Benito Moreno, 2016, § 2). Twitter, en particular, por su extensión limitada, su naturaleza pública (aunque informal) no necesariamente conversacional y su exigua regulación —que concede absoluta libertad temática—, constituye su propia variedad, un híbrido entre las diferentes variedades de géneros en línea (Jehl, 2010, § 3.3). Adicionalmente, Twitter proporciona datos geolocalizados precisos, lo que convierte los tuits geoetiquetados en la fuente de numerosos estudios que han favorecido el mapeo de la variación lingüística (Grieve *et al.*, 2019, p. 1). Entre aquellos sobre el español global, Gonçalves y Sánchez (2014 y 2016) analizan la variación léxica y establecen

grupos dialectales, Ruiz Tinoco (2016) compara distintos usos léxicos y gramaticales en España y Latinoamérica, y Pardo (2022) analiza usos frecuentemente cuestionados por los hispanohablantes.

A pesar del alcance ilimitado en temática, proveniencia geográfica y número de informantes (Ruiz Tinoco, 2016, p. 599; Grieve *et al.*, 2019, p. 15), predominan las cohortes jóvenes, masculinas y urbanas (Longley *et al.*, 2015, § 3.1; Eisenstein, 2017, p. 370), por lo que la representatividad demográfica no es completamente equilibrada. Incluso así, varios estudios han ratificado a Twitter como una fuente válida para componer mapas dialectales (Cook *et al.*, 2014; Doyle, 2014; Huang *et al.*, 2016) y estudiar patrones generales de variación léxica regional (Grieve *et al.*, 2019). Lamentablemente, las restricciones legales de intercambio de datos dificultan la creación de corpus públicos, lo que afecta la reproducibilidad de los resultados, pues están sujetos a ser borrados (Ruiz Tinoco, 2016, § 2.1; Estrada Arráez y Benito Moreno, 2016, pp. 96 y 106). A pesar de la facilidad de recolectar muestras de gran tamaño (Huang *et al.*, 2016, p. 254; Grieve *et al.*, 2019, p. 15), la creación de corpus representativos requiere mucho tiempo de relevo, en especial cuando involucra a países de distinto producto interno bruto (PIB), pues varía su producción de tuits (Dunn y Adams, 2020, p. 4). En consecuencia, resulta imperante utilizar directamente el motor de búsqueda de la interfaz de programación de Twitter, incluso con la limitación de que —además de no reconocer diacríticos— no permite el uso de expresiones regulares (secuencias de símbolos y caracteres que expresan una cadena o patrón que pueda ser buscado dentro de un texto más extenso), lo que hace del proceso un trabajo más artesanal.

Si bien los tuits geoetiquetados representan solo el 1-2 % del total (Huang *et al.*, 2016, p. 253) y responden a un subgrupo sociodemográfico (Malik *et al.*, 2015), también es posible utilizar la ubicación en texto libre que provee el usuario en su perfil. A pesar de que depende de la voluntad del usuario de informar su ubicación real, Pavalanathan y Eisenstein (2015) encontraron que en la mayoría de los casos son precisas (§ 3.1) y que las ubicaciones geoetiquetadas y autoinformadas producen corpus diferentes (§ 4.2). Considerar ambas crea un conjunto de datos mayor que puede generar resultados más confiables (Huang *et al.*, 2016, p. 253).

4 Metodología

Con el objetivo de determinar las frecuencias de uso y la variación diatópica de las entradas del OdP, estas fueron seleccionadas con el criterio de poder establecer pares de variantes de alternancia (palabras alternativas para un mismo significado). El estudio de la frecuencia relativa de términos en el ámbito del país resulta fundamental para obtener resultados fiables en el caso de términos muy poco frecuentes, evalua-

dos en países cuya producción de tuits es muy dispar. En algunos casos, las entradas del Observatorio tienen un equivalente preexistente en el DLE, que se usó para constituir el par, y, en otros, son las propias entradas las que conforman el par de alternancia (con o sin equivalentes aceptados). Se desestimaron entradas polisémicas, a excepción de aquellas con un uso secundario muy infrecuente, en cuyo caso ese uso fue desestimado para el análisis (p. ej.: *dilogía*).

Así, de las 73 palabras del ODP, se seleccionaron 23 y se definieron un total de 13 pares de variantes de alternancia (algunas colectivas), presentados en la tabla 1. Allí se incluye una breve reseña de lo que indica la RAE en el ODP para cada palabra (RAE, 2022) y, en caso de ya encontrarse en alguno de los diccionarios DLE, DA y DPD, se recogen estos datos.

El proceso de extracción y análisis de datos se llevó a cabo usando Python 3.10.1 (Rossum y Drake, 2009) y sus bibliotecas de análisis de datos, así como la interfaz de programación de Twitter (<https://developer.twitter.com/>).

La frecuencia absoluta de uso de cada variante —siguiendo a Gonçalves y Sánchez (2014) y Ruiz Tinocho (2016)— se determinó por la cantidad de tuits que incluían dicha variante. Se incluyeron filtros para recolectar tuits en español y para desestimar retuits, con lo que se evitaron textos replicados. Según fuera posible, se contemplaron variación en género y número, conjugación en presente de indicativo, sustantivos, adjetivos, verbos y participios derivados. En algunos casos, se analizaron diferencias de uso entre verbos y sustantivos derivados de una misma entrada, bajo la hipótesis de que la conjugación verbal podría ser favorable a una versión ortográficamente más castellanizada de los extranjerismos.

La población de estudio son los usuarios hispanohablantes de Twitter, pero —siguiendo a Pardo (2022, § 8)— solo fueron considerados aquellos países cuyas academias pertenecen a la ASALE. Para determinar la ubicación de cada tuit, se tomó en cuenta su geolocalización y el campo de ubicación informado por el usuario en su perfil (rango de error de 0,4 a 1,1 % con un 95 % de confianza), en ese orden de prioridad y cuando estuvieran disponibles (Pardo, 2022, § 8.4). Esto permite extraer más tuits válidos (con una ubicación identificable), conveniente especialmente para palabras poco frecuentes y para lograr una mejor representatividad demográfica.

El análisis diatópico se realizó empleando frecuencias de uso relativas, utilizando cada país como unidad regional, esto es, la frecuencia de uso (cantidad de tuits) de cada variante con respecto a la de ambas variantes del par, por país (Pardo, 2022, § 8.6):

$$f. \text{ rel. } \text{var1} | \text{país}_i = \frac{\text{tuits. var1}}{\text{tuits. var1} + \text{tuits. var2}} \Big|_{\text{país}_i}$$

Por ello, los tuits fueron recolectados por par de alternancia: los tuits que contenían cualquiera de las variantes se buscaron y recopilamos simultáneamente durante el mismo período de tiempo, con lo que se aseguraba la proporción de uso apropiada de cada variante para el número total de tuits descargados. El periodo de recolección fue de enero de 2010 a julio de 2022, a excepción de usos muy frecuentes en los que el intervalo fue menor (siempre desde 2022 hacia atrás). Estos datos fueron luego procesados para determinar su proveniencia y utilizados para el mapeo dialectológico. Este método permite independizarse de la diferencia de producción de tuits entre países y ver claramente cómo se distribuyen los usos regionalmente. En pocos casos, en los que hay más de dos variantes alternantes, se creó una variante colectiva para poder establecer un mapeo adecuado, se agruparon las variantes poco frecuentes en un mismo conjunto y se contrastó su uso relativo respecto a la preponderante: se agrupó las variantes poco frecuentes en un mismo conjunto y se las contrastó en uso relativo respecto a la preponderante (p. ej.: *ranking vs. ranquing, ranquin, rankin*).

Los pares de alternancia se organizaron en dos grupos ordenados consecutivamente (en la tabla 1: 1 a 6 y 7 a 13), según las posibilidades de su análisis temporal, esto es, según el volumen de datos disponible para la recolección. En los primeros seis pares, ambas variantes del par (o una variante colectiva) permitieron la recolección de tuits desde 2010 hasta julio de 2022, lo que posibilitó un análisis temporal comparativo; en los siguientes siete pares, solo la variante minoritaria pudo recolectarse para ese mismo período, por lo que su análisis temporal fue individual (tras la recolección simultánea para el análisis diatópico, se prosiguió por separado en la recolección hasta 2010). En todos los casos, el número de tuits de cada año, para cada variante, fue ponderado por la cantidad de usuarios de Twitter del año respectivo a nivel global, según datos de la web Statista (Statista, 2022a y 2022b), de modo que se reflejasen las variaciones en el uso y no las variaciones en cantidad de usuarios (número significativo en los inicios de la red social). Si bien los factores de ponderación responden a números de usuarios globales y no hispanohablantes en particular, fue considerado un compromiso necesario.

Para ambos análisis se descartaron los tuits que incluyeran ambas variantes o que mencionaran a la RAE, ya que esto podría implicar una reflexión de los usuarios sobre el idioma, en lugar de una producción espontánea (la alternancia inconsciente se descartó en favor de lo anterior). Se exigió un mínimo de veinte tuits por país para incluir el par en el análisis; de esta forma, la probabilidad de una asignación incorrecta en la ubicación representaría únicamente un 5 % del total.

Tanto en el caso de *dilogía* como *destripe* (uso peninsular para *spoiler*), su polisemia fue desestimada para el análisis.

N.º par	Entrada del OdP	Descripción del OdP + Detalle en obras normativas	Par de alternancia
1	eficientizar	Verbo válido, de empleo habitual en varios países de América. + En el DA: Nicaragua y República Dominicana.	eficientizar vs. eficientar
	eficientar	Verbo válido, de empleo habitual en varios países de América. + En el DA: Nicaragua y Honduras.	
2	gamificación	Adaptación híbrida desaconsejada de <i>gamification</i> . Se recomienda «ludificación».	gamificar vs. ludificar
	ludificación	Equivalente recomendado a <i>gamification</i> . Propuesta para el DLE. + En el DLE desde diciembre del 2021.	
3	turistólogo	Voz reciente, en alternancia con «turismólogo».	turistólogo vs. turismólogo
	turismólogo	Voz reciente, en alternancia con «turistólogo».	
4	dilogía	Voz correcta, en alternancia con «bilogía». Se recomienda «obra de dos partes». + En el DLE, con otro significado.	dilogía vs. bilogía
	bilogía	Voz correcta, en alternancia con «dilogía». Se recomienda «obra de dos partes».	
5	ranquin	Adaptación recomendada para <i>ranking</i> . Propuesta para el DLE. + En el DPD.	ranking vs. ranquin, ranquing, rankin
6	webinar	Anglicismo. Se recomiendan: «webinario», «seminario web» o «ciberseminario». + «webinario» y «seminario web» en el DLE y «ciberseminario» se considera un compuesto a partir de «ciber-» (también en el DLE).	webinar vs. webinar, seminario web, ciberseminario
7	mutear	Derivado español, a partir de <i>mute</i> . Sustituir por «silenciar».	mutear vs. silenciar
8	legitimizar	Voz innecesaria y desaconsejada. Sustituir por «legitimar».	legitimizar vs. legitimar
9	spoiler	Anglicismo. Sustituir por «destripe» o «espóiler».	spoiler vs. espóiler
	espóiler	Adaptación documentada y válida.	
	destripe	Derivado de «destripar». Alternativa a <i>spoiler</i> .	
10	guasap	Adaptación válida.	guasap(ear) vs. wasap(ear)
	wasap	Adaptación válida.	
	guasapear	Graña adecuada y documentada.	
	wasapear	Graña adecuada y documentada.	
11	guglear	Graña adaptada y documentada.	guglear vs. googlear
	googlear	Graña híbrida y desaconsejable.	
12	emoyi	Posible adaptación de «emoji», graña preferida. + «emoji» en el DLE desde diciembre del 2021.	emoyi vs. emoji
13	paranoya	Graña aceptable de «paranoia», graña tradicional.	paranoya vs. paranoia

TABLA I. Entradas seleccionadas del OdP y par de alternancia asociado
 Todas las tablas son de elaboración propia. En este caso, las fuentes de información son:
 RAE, 2022; RAE y ASALE 2022, 2010, 2005 (DLE, DA y DPD, respectivamente)

5 Resultados

La cantidad de tuits recolectados fue muy variable en función del par de alternancia: se obtuvieron una media y una desviación estándar globales de (184 ± 120) mil tuits. El par más escaso fue *turistólogo/turismólogo*, con solo 8.361 tuits desde 2010. El requerimiento de un mínimo de veinte tuits por país limitó la inclusión de Guinea Ecuatorial en solo dos pares de alternancia y a la mitad en el caso de Filipinas. Sin embargo, en todos los análisis de alternancia se contó con la mayoría de los países hispanohablantes.

La tabla 2 resume los resultados de los análisis temporal y diatópico de cada par alternante. Para el análisis temporal, se clasifica el estado de evolución actual siguiendo a Freixa (2022, § 5); para el análisis diatópico, se presentan las distribuciones de uso encontradas para la variante minoritaria globalmente (no necesariamente localmente); se incluyen la media y la desviación estándar (SD) de su frecuencia relativa de uso a nivel global, así como el detalle de los tres países que mostraron un mayor uso (estos valores son complementarios a los de la variante mayoritaria; la SD expresa la dispersión de los datos respecto a la media). Tanto las gráficas temporales como los mapas dialectológicos se encuentran disponibles en el anexo.

El par *eficientar/eficientizar* presentó, para ambos casos, un uso estabilizado en el mundo hispanohablante, a pesar de diferencias en amplitud debido a la disparidad de uso de ambas variantes, con clara preponderancia a favor de *eficientar*; el análisis diatópico confirmó su preeminencia a nivel global (5,9 %) y local, aunque en República Dominicana alterna casi en igual medida con *eficientizar*. A pesar de que ambos términos figuran en el DA, sendos usos aparecen con la marca diatópica de Nicaragua; *eficientizar* con la de República Dominicana, y *eficientar* con la de Honduras. Sin embargo, en Nicaragua no se alcanzó el mínimo requerido de veinte tuits entre ambas variantes y, aunque se trata de un par poco frecuente (33.876 tuits), hubo otros pares alternantes de frecuencia similar para los que Nicaragua sí satisfizo el mínimo. Argentina es el otro país donde *eficientizar* tiene un uso relativamente significativo, un 35 %, seguido de Colombia, con tan solo el 4 %. España no quedó excluida del análisis (puesto diez de dieciséis, con un 0,2 % de uso), por lo que puede afirmarse que este uso no se encuentra restringido a América. Los resultados, para este par alternante, no se corresponden con lo que sugiere la norma académica vigente. Comparando este ejemplo con el par *legitimar/legitimizar*, donde podría esperarse algún paralelismo en el uso, se observa cierta estabilidad general en la variante minoritaria, que tiene una media de uso similar (6,3 %), pero una desviación estándar menor y una distribución distinta entre países, aunque los de mayor uso también son americanos. La RAE, en este caso, no recomienda su uso, lo que está en consonancia con la preferencia popular.

El par *gamificar/ludificar* muestra una clara preponderancia del término desaconsejado por la RAE. Mientras que *ludificar* alcanza, globalmente, apenas un 2,8 % de uso (SD 1,5 %), este fue incluido en el DLE en 2021 (seguramente, dada su etimología). Entre las críticas de los usuarios se encuentran el retraso de la RAE en su aporte y la falta de asociación del término propuesto con su uso (Lara, 2019; exYge, 2020).

El caso del par *turistólogo/turismólogo* evidencia un patrón neológico realista (ascenso, descenso y estabilización), aunque la preferencia a nivel global parece haberse invertido en favor de *turismólogo*. Sin embargo, el término minoritario a nivel global no lo es a nivel local, con una media de 37,8 % y una SD de 23,7 %, siendo el par que exhibió más variabilidad regional.

Bilogía ha primado sobre *dilogía*, mostrando la preferencia por el prefijo latino «bi-» sobre el griego «di-». Incluso habiendo desestimado el significado presente en el DLE, el uso de *dilogía* es notablemente minoritario (3,5 %); destaca Cuba con el 20,8 %.

Los casos de los anglicismos *ranking*, *webinar* y *spoiler* muestran que estos son ampliamente preponderantes sobre sus respectivas variantes en español. Dentro de las minoritarias, se observa el uso mayoritario de *rankin* sobre las variantes con «q» (*ranquin* está incluido en el DPD), y la preferencia de *webinario* y *seminario web* sobre un casi inexistente *ciberseminario*, todos ellos con un uso acentuado en la pandemia de 2020, pero respetando las tendencias previas. Notoriamente, estas versiones minoritarias en español (a excepción de *espóiler*, que ronda tan solo el 0,5 % de uso) tienen un uso más elevado en Estados Unidos, quizás motivado por un deseo de mayor distinción respecto al inglés. Es similar lo que ocurre con los pares *wasap/guasap* y *goglear/guglear*, donde, si bien se trata de versiones en español, la primera —más cercana a la inglesa— es la más usada; el pico temporal de *guasap* coincide con la popularización masiva de la aplicación Whatsapp, mientras que *guglear* ya se muestra estabilizada.

En todos los casos en los que se estudió la diferenciación entre sustantivos y verbos, se observó, para los últimos, un incremento notable en el uso de la variante de ortografía castellanizada (en especial, *ranquear* y *espoilear*, menor en *guasapear*).

En los pares *emoji/emoj* y *paranoya/paranoia* el uso con «y» es insignificante, prácticamente inexistente en conjunto en el caso de *emoji* y, para *paranoya*, se destaca únicamente España con un 6,4 % de uso. Mientras *paranoya* parece haber estabilizado su uso, restaría ver la evolución temporal de *emoji* en el futuro. Es llamativo que ambos usos, tan infrecuentes, sean sugeridos como válidos por la RAE.

Finalmente, cabe aludir al caso de *mutear*, del que, en competencia con *silenciar*, cabría esperar un uso muy limitado, pero que alcanzó un 8,1 % global y superó el 20 % de uso en Puerto Rico y República Dominicana. El pico de frecuencia se vio favorecido por la pandemia. De hecho, los usuarios se dirigieron a

N.º par	Var. minoritaria (var. mayoritaria)	media \pm SD	Análisis diatópico		Mapa	Análisis temporal
			Países principales			
1	eficientizar (eficientar)	5,9 % \pm 13,9 %	R. Dominicana 47,5 % Argentina 34,3 % Colombia 4,1 %			estabilización
2	ludificar (gamificar)	2,8 % \pm 1,5 %	Bolivia 7,4 % Cuba 5,5 % Uruguay 3,8 %			ascenso y descenso
3	turismólogo (turistólogo)	37,8 % \pm 23,7 %	Perú 79,7 % España 76,0 % Ecuador 61,5 %			ascenso, descenso y estabilización
4	dilogía (bilogía)	3,5 % \pm 4,7 %	Cuba 20,8 % Bolivia 8,3 % Perú 8,1 %			ascenso, descenso y estabilización
5	ranquin, rankin, ranqing, (ranking)	1,5 % \pm 1,4 %	Méjico 5,6 % EE. UU. 3,5 % Puerto Rico 3,2 %			descenso y estabilización
	ranquear (rankear)	15,7 % \pm 8,5 %	R. Dominicana 36,7 % Puerto Rico 31,2 % Colombia 23,4 %			descenso y estabilización
6	webinario, seminario web, ciberseminario (webinar)	9,5 % \pm 4,7 %	EE. UU. 19,7 % El Salvador 19,3 % Uruguay 13,6 %			ascenso y descenso
7	mutear (silenciar)	8,1 % \pm 5,5 %	Puerto Rico 22,9 % R. Dominicana 20,7 % Costa Rica 17,8 %			ascenso
8	legitimizar (legitimiar)	6,3 % \pm 4,3 %	Puerto Rico 20,2 % Panamá 13,0 % Bolivia 12,7 %			estabilización

N.º par	Var. minoritaria (var. mayoritaria)	media \pm SD	Análisis diatópico		Mapa	Análisis temporal
			Países principales			
9	espóiler (spoiler)	0,5 % \pm 0,2 %	Cuba Panamá España	1,0 % 0,8 % 0,7 %		 descenso y estabilización
	espoilear (spoilear)	3,4 % \pm 1,7 %	Puerto Rico Chile R. Dominicana	9,4 % 5,1 % 4,9 %		
10	guasap (wasap)	13,2 % \pm 9,1 %	Costa Rica Guatemala Honduras	39,6 % 30,3 % 23,9 %		 ascenso, descenso y estabilización
	guasapear (wasapear)	16,3 % \pm 7,1 %	Costa Rica Uruguay Honduras	30,7 % 30,3 % 24,5 %		
11	guglear (googlear)	3,0 % \pm 1,1 %	Chile Filipinas Cuba	4,9 % 4,6 % 4,5 %		 estabilización
12	emoyi (emoji)	0,04 % \pm 0,04 %	El Salvador Puerto Rico Venezuela	0,2 % 0,1 % 0,1 %		 ascenso y descenso
13	paranoya (paranoia)	0,6 % \pm 1,3 %	España El Salvador Bolivia	6,4 % 1,1 % 1,0 %		 ascenso, descenso y estabilización

TABLA 2. Análisis diatópico y temporal. Tabla de elaboración propia. Los datos surgen del análisis detallado en la metodología

la Academia, haciendo referencia a la influencia de la aplicación Zoom en la popularización de este término (Pérez, 2020).

6 Discusión

Se han estudiado ciertas voces neológicas del ODP con el objeto de clarificar su funcionamiento y su cometido. Los términos analizados han mostrado que, a pesar de una gran variabilidad en su frecuencia absoluta de uso, todos se extienden a lo largo del territorio hispanohablante. Aunque algunos países han quedado excluidos en unos pocos pares de variantes, en nin-

gún caso se observaron agrupaciones regionales, por lo que podríamos hablar de términos de alcance pan-hispánico. Asimismo, en general, el uso mayoritario desde el punto de vista global se reflejó también en el ámbito local.

Por otra parte, los análisis temporales han puesto de manifiesto no solo la existencia de distintos estadios evolutivos para cada voz, sino también la clara influencia de los fenómenos sociales en la variación lingüística. Incluso si ya estaban establecidos en uso, los picos de *webinario* y *mutear* reflejan su aumento de uso por las prácticas que impuso la pandemia en 2020, así como el de *guasap* indica la popularización de la aplicación. Con la excepción de *turismólogo/turistólogo*, en ningún

caso hubo —al menos a escala global— una inversión en las preferencias de uso entre variantes, lo que parece indicar cierta inflexibilidad en el cambio de voces establecidas. Quizás es justamente la baja frecuencia general del par referido al turismo lo que ha permitido su inversión en el uso.

De modo similar, tal como plantea Sánchez Lancis (2022), los extranjerismos tienden a estar más arraigados que las versiones castellanizadas propuestas; sea por prestigio, moda o especialización, su uso determina la preferencia y el diccionario puede acabar incluyéndolos y marcándolos en cursiva (pp. 144-145). Cuando el extranjerismo posee una grafía o fonética que facilita su adaptación sin grandes cambios, puede dar lugar a conservar grafías no patrimoniales a final de palabra (p. 145), como en *guasap/wasap* (cabe notar que la RAE las considera adaptaciones válidas). Se ha visto, empero, que las conjugaciones verbales sí favorecen las grafías tradicionales (*espoilear, ranquear*), por lo que quizás —aunque aún minoritarias— pueda haber una tendencia hacia ellas. El uso preponderante de la *k* en el caso *rankear/ranquear* no es suficiente para concluir que sea un mecanismo de uso actual, como sí podría manifestar la preferencia por el prefijo *bi-* sobre el *di-*. Sería interesante analizar más ejemplos similares.

Aunque posiblemente todas las palabras satisfagan las características a las que se refería Freixa (2002) para la incorporación de estos términos al diccionario (§ 2) y, aunque la mayoría fueran calificados por la RAE como aceptables, la autora notaba que las voces socialmente institucionalizadas no necesariamente son finalmente incorporadas al diccionario. De las palabras del OdP analizadas, solo una está en el DLE (*ludificación*), una en el DPD (*ranquín*) y dos en el DA (*eficientizar, eficiente*). Todas ellas respetan la ortografía española y, puesto que están recogidas en diferentes diccionarios, plantean dudas respecto a los criterios empleados para determinar su inclusión y a qué consecuencias pueden derivarse de este hecho en cuanto a la frecuencia o valoración sociolingüística del uso. Aunque el DPD y el DA son parte de la norma prescrita para el español y a la vez la configuran, no tienen la misma llegada al público ni el mismo prestigio que el *Diccionario de la lengua española*. De hecho, una de las dificultades que se le presenta al hablante al consultar una palabra es tener que hacerlo en todos los diccionarios por separado, pues no hay una búsqueda centralizada y, por defecto, se busca en el DLE. Lo más sencillo sigue siendo *guglear* el término en cuestión, agregando «RAE» a la búsqueda, o, alternativamente, enviar la consulta a la RAE (o a la FundéuRAE, fundación que también responde consultas).

Las palabras fueron incorporadas al OdP, según explica la RAE, por haber generado dudas entre los hablantes. Tratándose de términos de extensión panhispánica, cabría pensar que estas preguntas han tenido procedencias variadas. Lamentablemente, las consultas en Twitter sobre estos usos no son lo suficientemente

numerosas como para llevar a cabo un análisis estadístico fiable de cada consulta (como es el caso de las consultas generales [Pardo, 2022]) —cerca de la mitad de los casos estudiados no alcanza los cien tuits y retuits con ubicación válida—, pero un análisis cualitativo de los hilos de mensajes puede servir como introspección a las diferentes actitudes y opiniones de los hablantes.

Si bien la decisión en este trabajo de tomar pares de variantes de alternancia limitó la elección de los términos, permitió, al mismo tiempo, una comparación fidedigna de los usos locales a escala global. Un estudio de regionalismos puede ser comparativo si se limita a un solo país o puede realizarse a mayor escala si se desea corroborar cualitativamente la información provista en el observatorio.

A este respecto, cabe destacar que la creación del OdP responde a la actual política lingüística de la RAE, la presentación de una imagen, discurso y disposición renovada y tolerante de la institución, abierta al compromiso con el público (Paffey, 2021, p. 253). Aunque las entradas del OdP cumplen criterios de neologicidad y permiten interesantes análisis respecto a la incorporación de las voces, este no ha sufrido cambios desde 2020 y valdría evaluar qué voces fueron incorporadas al DLE sin pasar por el Observatorio. Ciertamente, los objetivos institucionales van más allá de los meramente lingüísticos, dada la compleja naturaleza sociopolítica de la gestión de la lengua y las disputas entre los diversos actores.

7 Conclusión

Este artículo ha estudiado —empleando Twitter como corpus dialectológico— la variación diatópica desde el punto de vista del país y la evolución temporal, durante la última década, de una selección de términos del OdP de la RAE, basada en la posibilidad de definir pares de variantes de alternancia. La metodología de mapeo dialectológico permitió llevar a cabo un análisis de extensión panhispánica y corroborar que todos los términos seleccionados tienen uso a escala global (incluso los preferidos muy acusadamente en ciertos países de la hispanofonía) y que, a excepción de un único caso, las variantes mayoritarias globalmente lo son también localmente. Las recomendaciones de la RAE —que en su mayoría son validaciones de las voces— se condicen con el uso popular, a excepción de los extranjerismos, que en el habla popular priman sobre las versiones españolas, incluso si, en los usos verbales, la ortografía muestra una tendencia hacia la castellanización de estos términos.

A pesar de que la RAE objete que el OdP sea un estadio intermedio en la aprobación de las voces y se desentienda de la aceptación de sus entradas, es claro que su presencia en este observatorio no es indiscriminada. Lejos de mostrar indiferencia, la Academia ha creado este novedoso espacio y ha clasificado sus

voces como adecuadas, aceptables, correctas y recomendadas, o como innecesarias, no recomendadas y desaconsejadas. Tal y como se señala explícitamente, esta información puede cambiar en el futuro, pero hoy en día son esos los juicios de valor establecidos por la autoridad de referencia de la lengua, a quien el hablan-

te común recurre mayoritariamente, no sin el esfuerzo continuado de la Academia. Los usos, sin embargo, progresarán y variarán en función de la negociación permanente entre la conservación y la innovación en el uso del idioma. ✿

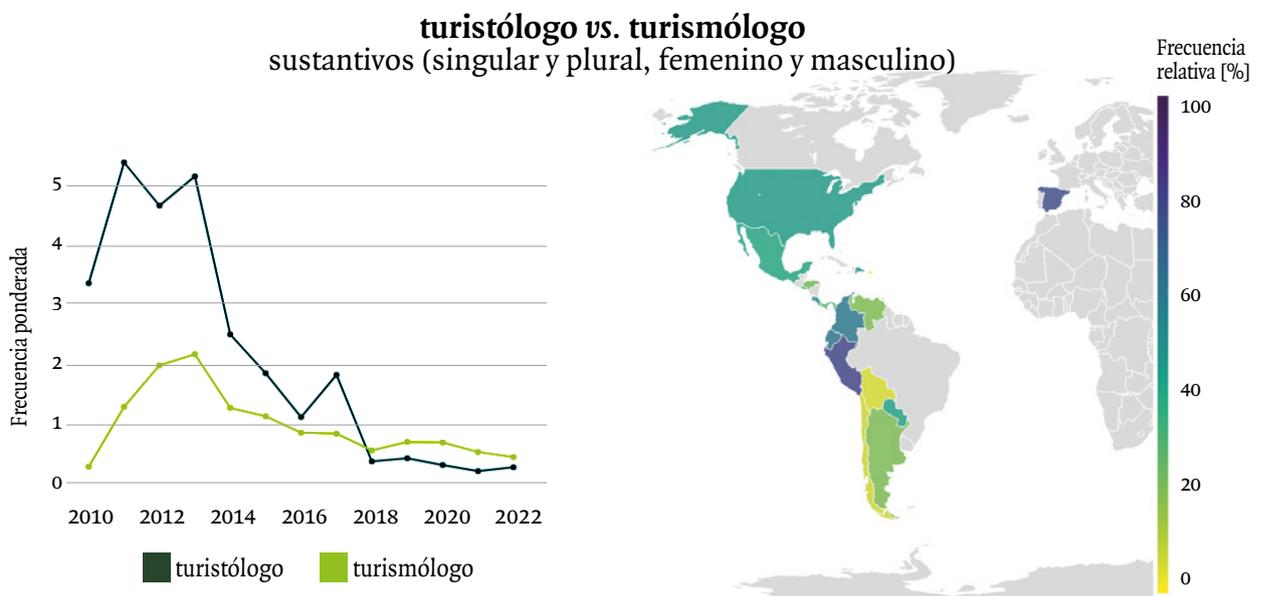
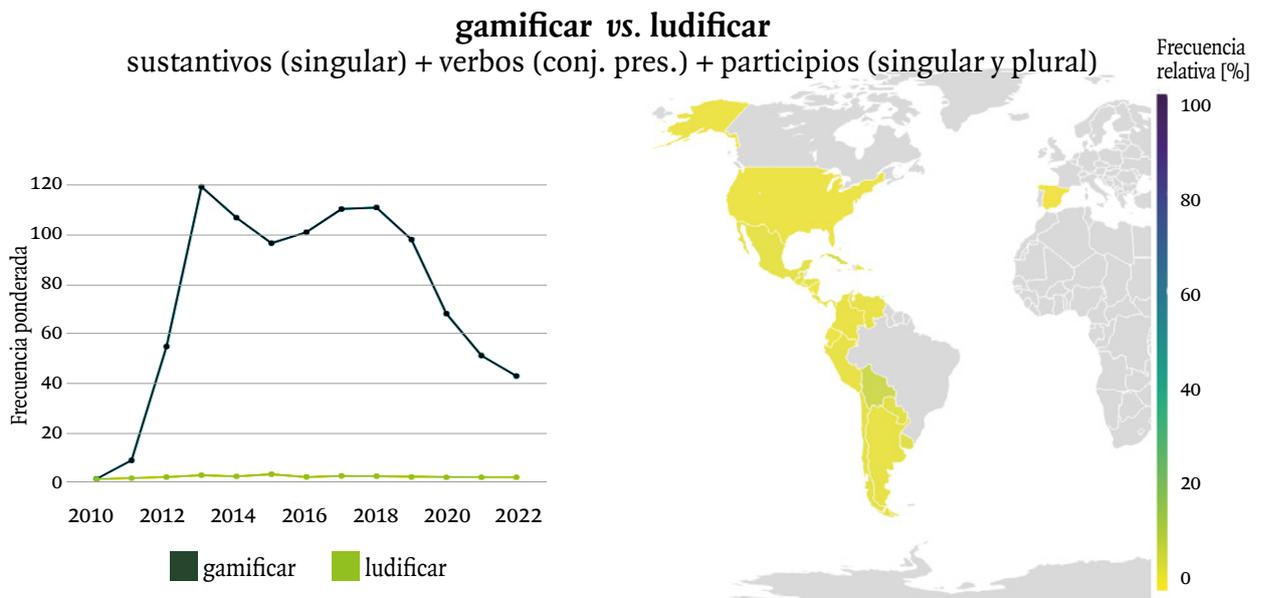
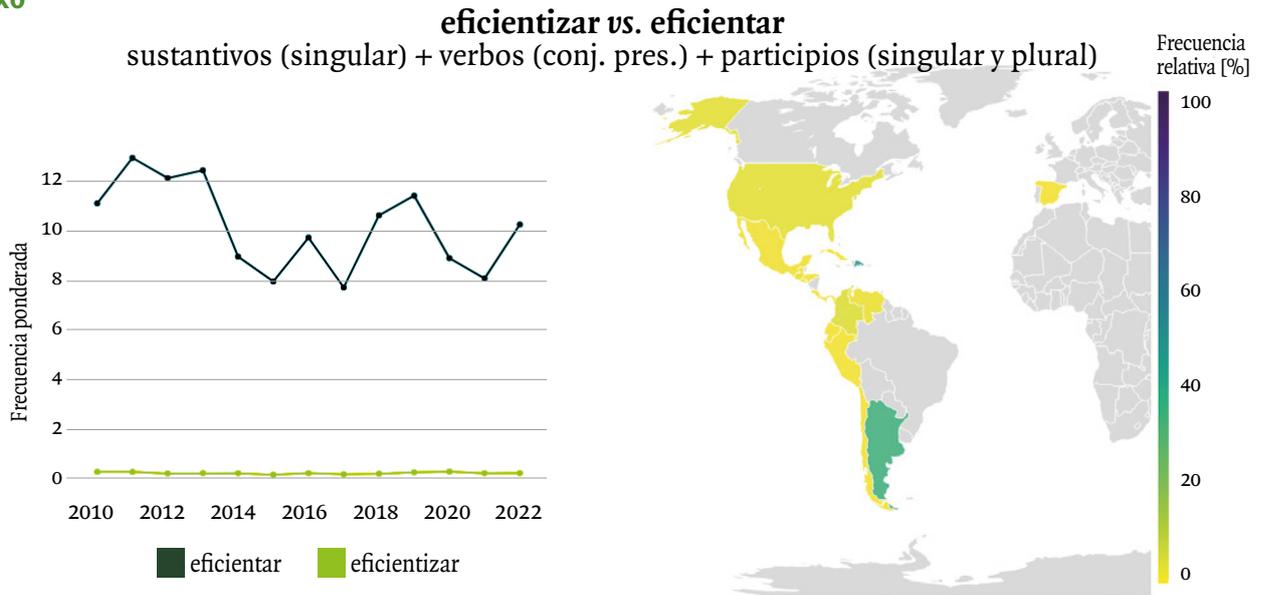
8 Bibliografía

- AMORÓS-NEGRE, Carla (2018). *La estandarización lingüística de los relativos en el mundo hispánico: Una aproximación empírica* [en línea]. Madrid: Iberoamericana; Frankfurt am Main: Vervuert. <<https://doi.org/10.31819/9783954877874>> [Consulta: 20 enero 2022].
- ARNANDIS, Rubén (@rubenarnandis) (2017). «Gracias @RAEinforma. Así lo haremos. Esperamos ver muy pronto “turismólogo/a” en las nuevas entradas del DLE!». Twitter, 25 mayo. <<https://twitter.com/rubenarnandis/status/867649499560374272>>.
- ASALE (2022). *Academias* [en línea]. <<https://www.asale.org/academias>> [Consulta: 30 enero 2022].
- ÁVILA, Raúl (2004). «¿El fin de los diccionarios diferenciales? ¿El principio de los diccionarios integrales?». *Revista de Lexicografía*, x, pp. 7-20.
- CARLA, Paola (@calapola). (2014). «@RAEinforma #RAEconsultas ¿Existe el verbo eficientar o eficientizar? Lo he escuchado mucho y me parece extraño». Twitter, 13 mayo. <<https://twitter.com/calapola/status/466288803792314369>>.
- COOK, Paul; HAN, Bo; BALDWIN, Timothy (2014). «Statistical methods for identifying local dialectal terms from GPS-tagged documents». *Dictionaries: Journal of the Dictionary Society of North America*, 35 (35), pp. 248-271.
- DALUNA MX (@dalunam) (2018). «¿“Guasap”? ¿Es en serio? ¡Quiero llorar! ¡Por favor díganme que es un sueño, una broma o que es una cuenta falsa de la RAE!». Twitter, 14 agosto. <<https://twitter.com/dalunam/status/1029349439729819648>>.
- DOYLE, Gabriel (2014). «Mapping dialectal variation by querying social media». *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics* (abril), pp. 98-106.
- DUNN, Jonathan; ADAMS, Benjamin (2020). «Mapping languages and demographics with georeferenced corpora» [en línea]. <<https://arxiv.org/abs/2004.00809>>.
- EISENSTEIN, Jacob (2017). «Identifying regional dialects in on-line social media». En: BOBERG, Charles; NERBONNE, John; WATT, Dominic (ed.). *The handbook of dialectology*. Hoboken, New Jersey: John Wiley & Sons, Inc., pp. 368-383.
- ESTRADA ARRAÉZ, Ana; Benito Moreno, Carlota de (2016). «Variación en las redes sociales: datos twilectales». *Revista Internacional de Lingüística Iberoamericana*, 28(2), 77-111.
- EXYGE (@exYge) (2020). «Si escribimos ludificación nadie va a saber de qué hablamos. Por otra parte no es nuevo el uso de terminología». Twitter, 21 abril. <<https://twitter.com/exYge/status/1252661135666536461>>.
- FELIPE EL IGUANO (@jelipe702) (2019). «@RAEinforma es la primera vez que escucho que se use “destripar” en este sentido. ¿Es propia de algún país o región?». Twitter, 25 marzo. <<https://twitter.com/jelipe702/status/1110185688723587074>>.
- FERNÁNDEZ-SILVA, Sabela (2022). «La variación geográfica de la neología en español y su representación lexicográfica». En: BERNAL, Elisenda; FREIXA, Judit; TORNER, Sergi (ed.). *La neología del español: Del uso al diccionario*. Madrid: Iberoamericana; Frankfurt am Main: Vervuert, pp. 259-282.
- FREIXA, Judit (2022). «El proceso neológico». En: BERNAL, Elisenda; FREIXA, Judit; TORNER, Sergi (ed.). *La neología del español: Del uso al diccionario*. Madrid: Iberoamericana; Frankfurt am Main: Vervuert, pp. 51-71.
- GARCÍA PLATERO, Juan Manuel (2022). «Panorama de la investigación de la neología en español». En: BERNAL, Elisenda; FREIXA, Judit; TORNER, Sergi (ed.). *La neología del español: Del uso al diccionario*. Madrid: Iberoamericana; Frankfurt am Main: Vervuert, pp. 17-34.
- GONÇALVES, Bruno; SÁNCHEZ, David (2014). «Crowdsourcing dialect characterization through Twitter». *PLoS ONE* [en línea], 9 (11), pp. e112074. <<https://doi.org/10.1371/journal.pone.0112074>> [Consulta: 20 enero 2022].

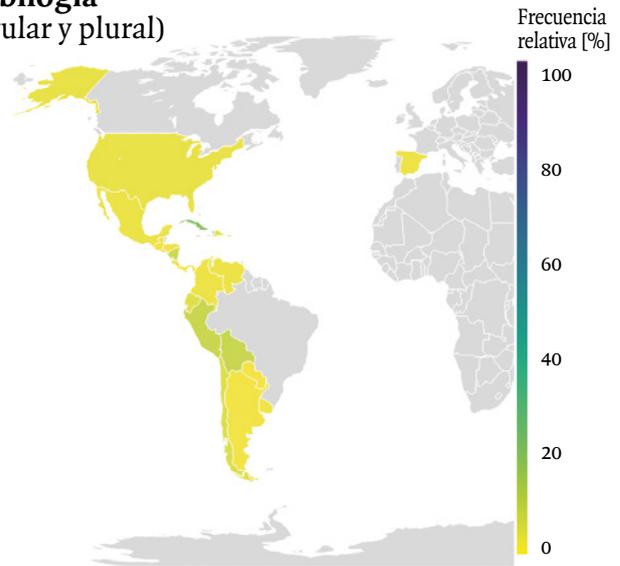
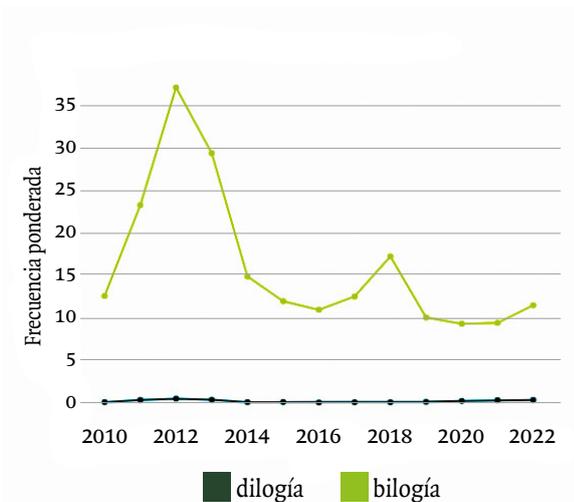
- (2016). «Learning about Spanish dialects through Twitter». *Revista Internacional de Lingüística Iberoamericana* [en línea], 28 (2), pp. 65-75. <<http://digital.casalini.it/4134616>> [Consulta: 20 enero 2022].
- GREUSSLICH, Sebastian (2015). «El pluricentrismo de la cultura lingüística hispánica: política lingüística, los estándares regionales y la cuestión de su codificación». *Lexis*, 39 (1), pp. 57-99.
- GRIEVE, Jack; MONTGOMERY, Chris; NINI, Andrea; MURAKAMI, Akira; GUO, Diansheng. (2019). «Mapping lexical dialect variation in British English using Twitter». *Frontiers in Artificial Intelligence* [en línea], 2 (11). <<https://doi.org/10.3389/frai.2019.00011>> [Consulta: 3 enero 2022].
- GUERRERO RAMOS, Gloria (2022). «La neología en los diccionarios del español». En: BERNAL, Elisenda; FREIXA, Judit; TORNER, Sergi (ed.). *La neología del español: Del uso al diccionario*. Madrid: Iberoamericana; Frankfurt am Main: Verbuert, pp. 191-209.
- HUANG, Yuan; GUO, Diansheng; KASAKOFF, Alice; GRIEVE, Jack. (2016). «Understanding US regional linguistic variation with Twitter data analysis». *Computers, Environment and Urban Systems*, 59, pp. 244-255. Birmingham: University of Birmingham.
- INSTITUTO CERVANTES (2022). *El español en el mundo 2022: Anuario del Instituto Cervantes* [en línea]. <https://cvc.cervantes.es/lengua/anuario/anuario_22/el_espanol_en_el_mundo_anuario_instituto_cervantes_2022.pdf> [Consulta: 5 junio 2022].
- JEHL, Laura Elisabeth (2010). «Machine translation for Twitter». En: *Edinburgh Research Archive* [en línea]. <<https://era.ed.ac.uk/handle/1842/5317>> [Consulta: 23 marzo 2022].
- JIMÉNEZ RÍOS, Enrique (2015). «Recorrido histórico por las razones para la admisión de voces nuevas en la lengua y en el diccionario». *Philologica Canariensis*, 21, pp. 45-80.
- KOCH, Peter; OESTERREICHER, Wulf (2007). *Lengua hablada en la Rumania: español, francés, italiano*. Traducción de Araceli López Serena. Madrid: Gredos. [Versión original: *Gesprochene Sprache in der Rumania: Französisch, Italienisch, Spanisch* (1990). Tübingen: Max Niemeyer Verlag]
- LARA, Luis Fernando (2005). «Para una reconstrucción de la idea de la lengua española: Más allá de las fronteras instituidas». *Revista Internacional de Lingüística Iberoamericana*, 6, pp. 171-187.
- LARA (@laraalonsosimon) (2019). «Todo eso llega tardísimo, me temo. “Gamificación” está tan arraigada en el ámbito de la enseñanza...». Twitter, 20 septiembre. <<https://twitter.com/laraalonsosimon/status/1175008999185571840>>.
- LONGLEY, Paul A.; ADNAN, Muhammad; LANSLEY, Guy. (2015). «The geotemporal demographics of Twitter usage». *Environment and Planning A*, 47 (2), pp. 465-484.
- LORENTE CASAFONT, Mercè (2022). «La vitalidad de las lenguas y el cambio léxico». En: BERNAL, Elisenda; FREIXA, Judit; TORNER, Sergi (ed.). *La neología del español: Del uso al diccionario*. Madrid: Iberoamericana; Frankfurt am Main: Verbuert, pp. 35-50.
- LUKAČ, Morana (2018). «Grassroots prescriptivism: An analysis of individual speakers’ efforts at maintaining the standard language ideology». *English Today*, 34 (4), 5-12.
- MALIK, Momin; LAMBA, Hemank; NAKOS, Constantine; PFEFFER, Jürgen. (2015). «Population bias in geotagged tweets». *Proceedings of the International AAAI Conference on Web and Social Media*, 9 (4), 18-27.
- NEGRETE-PORTILLO, Rafael (@_Rafael_Negrete) (2018). «Gracias. Pues empezaremos a llamarlo “ludificación” a ver [sic] conseguimos ganarle la batalla al calco». Twitter, 22 mayo. <https://twitter.com/_Rafael_Negrete/status/998895946149908480>.
- NGUYEN, Dong; DOĞRUÖZ, A. Seza; ROSÉ, Carolyn P.; JONG, Franciska de (2016). «Computational sociolinguistics: A survey». *Computational Linguistics*, 42 (3), pp. 537-593.
- NIEDZIELSKI, Nancy A.; PRESTON, Dennis R. (2000). *Folk linguistics*. Ed. a cura de W. Winter. Berlín: Nueva York: Mouton de Gruyter. (Trends in Linguistics: Studies and Monographs; 122)
- PARAKRAMA, Arjuna (1995). *De-hegemonizing language standards learning from (post)colonial Englishes about ‘English’*. Londres: Palgrave Macmillan.
- PAVALANATHAN, Umashanthi; EISENSTEIN, Jacob (2015). «Confounds and consequences in geotagged Twitter data» [en línea]. <<https://arxiv.org/abs/1506.02275>>.
- PRIETO DE LOS MOZOS, Emilio (1999). «Sobre las gramáticas normativas y las gramáticas no normativas». En: ARGENTER GIRALT, Joan Albert. *Homenatge a Jesús Tusón*. Barcelona: Empúries, pp. 245-258.
- RAE (2020^d). *Observatorio de Palabras* [en línea]. <<https://www.rae.es/portal-linguistico/observatorio-de-palabras>>. Internet Archive. <<https://web.archive.org/web/20201029041000/https://www.rae.es/portal-linguistico/observatorio-de-palabras>> [Consulta: 23 julio 2022].
- (@RAEinforma) (2020b). «Gracias por su interés. Debido a la confusión que ha generado la presencia de “elle” en el “Observatorio de palabras”...». Twitter, 30 octubre. <<https://twitter.com/RAEinforma/status/1322152976186806278>>.
- (@RAEinforma) (2021a). «Todas las palabras que aparecen actualmente en el “Observatorio” se incluyeron en él en 2020». Twitter, 15 febrero. <<https://twitter.com/RAEinforma/status/1361238500943802368>>.

- (@RAEinforma) (2021b). «El objetivo del observatorio no es recoger las palabras que están en fase de incorporación al diccionario, sino resolver dudas». Twitter, 14 mayo. <<https://twitter.com/RAEinforma/status/1393164970049875968>>.
- (2022). Observatorio de Palabras [en línea]. <<https://www.rae.es/portal-linguistico/observatorio-de-palabras>> [Consulta: 23 julio 2022].
- RAE; ASALE (2005). *Diccionario panhispánico de dudas* [en línea]. Madrid: Real Academia de la Lengua Española: Asociación de Academias de la Lengua Española: Santillana Ediciones Generales. <<https://www.rae.es/dpd/>> [Consulta: 25 julio 2022].
- (2010). *Diccionario de americanismos* [en línea]. Madrid: Real Academia de la Lengua Española. Asociación de Academias de la Lengua Española: Santillana Ediciones Generales. <<https://www.asale.org/damer/>> [Consulta: 25 julio 2022].
- (2022). *Diccionario de la lengua española* [en línea]. 23.^a ed. Madrid: Espasa [versión 23.5]. <<https://dle.rae.es>> [Consulta: 25 julio 2022].
- PARDO, Julia Elena (2022). #dudaRAE: A merging analysis of folk prescriptivism and dialectology [en línea]. Tesis de máster. Universitat Pompeu Fabra (UPF). Repositori Digital de la UPF. <<https://repositori.upf.edu/handle/10230/54016>> [Consulta: 1 julio 2022].
- PAFFEY, Darren J. (2021). «State-appointed institutions: Authority and legitimacy in the Spanish-Speaking world». En: AYRES-BENNETT, W.; BELLAMY, J. (ed.). *The Cambridge handbook of language standardization* [en línea]. Cambridge: Cambridge University Press, pp. 234-262. (Cambridge Handbooks in Language and Linguistics) <<https://doi.org/10.1017/9781108559249.009>> [Consulta: 10 marzo 2022].
- PÉREZ, Leopoldo (@leopoldo_perez7) (2020). «El uso de la herramienta del ZOOM es que, a puesto la palabra “mutear” de moda, yo la escuché hace un mes». Twitter, 26 mayo. <https://twitter.com/leopoldo_perez7/status/1265431929693900800>.
- RUIZ TINOCO, Antonio (2018). «Geocorpus del español de las redes sociales y cartografía automática». Monográficos SINOELE: IX Congreso Internacional de la Asociación de Hispanistas [en línea] (Bangkok, 2016), 17, pp. 598-608. <http://www.sinoele.org/images/Revista/17/monograficos/AAH_2016/AAH_2016_antonio_ruiz.pdf> [Consulta: 20 enero 2022].
- ROSSUM, Guido van; DRAKE, Fred L. (2009). *Python 3 Reference Manual* [en línea]. Scotts Valley, CA: CreateSpace. <<https://www.python.org/>> [Consulta: 03 enero 22].
- SÁNCHEZ LANCIS, Carlos (2022). «La neología, entre el uso y la prescripción». En: BERNAL, Elisenda; FREIXA, Judit; TORNER, Sergi (ed.). *La neología del español: Del uso al diccionario*. Madrid: Iberoamericana; Frankfurt am Main: Verbuert, pp. 129-150.
- SÁNCHEZ LOBATO, Jesús (2019). «La lengua española en busca del sobresaliente». En: MERINO, J. M.; GRIJELMO, A. (ed.). *Más de 555 millones podemos leer este libro sin traducción: La fuerza del español y cómo defenderla*. Taurus, pp. 65-79.
- SANMARTÍN, Julia (2022). «Neología y registro: sobre lo coloquial, lo formal y lo especializado en medios de comunicación escritos». En: BERNAL, Elisenda; FREIXA, Judit; TORNER, Sergi (ed.). *La neología del español: Del uso al diccionario*. Madrid: Iberoamericana; Frankfurt am Main: Verbuert, pp. 233-258.
- SENZ, Silvia; MINGUELL, Jordi; ALBERTE, Mon (2011). «Las academias de la lengua española, organismos de planificación lingüística». En: *El dardo en la Academia: Esencia y vigencia de las academias de la lengua española*, 1, pp. 371-550.
- STATISTA (2022a). *Number of monthly active Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019* [en línea]. <<https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/>> [Consulta: 12 marzo 2022].
- (2022b). *Annual Twitter user growth rate worldwide from 2015 to 2024* [en línea]. <<https://www.statista.com/statistics/303723/twitters-annual-growth-rate-worldwide/>> [Consulta: 12 marzo 2022].
- VALLE, José del del (2007). «La RAE y el español total: ¿esfera pública o comunidad discursiva?». En: VALLE, José del (ed.). *La lengua, ¿patria común?: Ideas e ideologías del español*. Madrid: Iberoamericana; Frankfurt am Main: Verbuert, pp. 81-96.

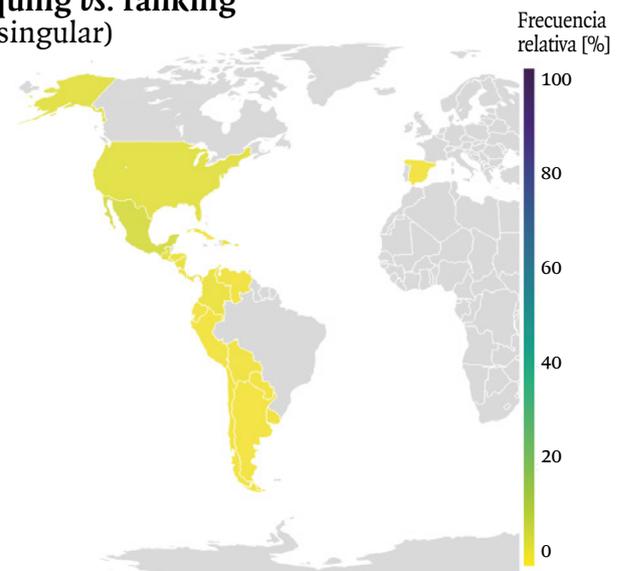
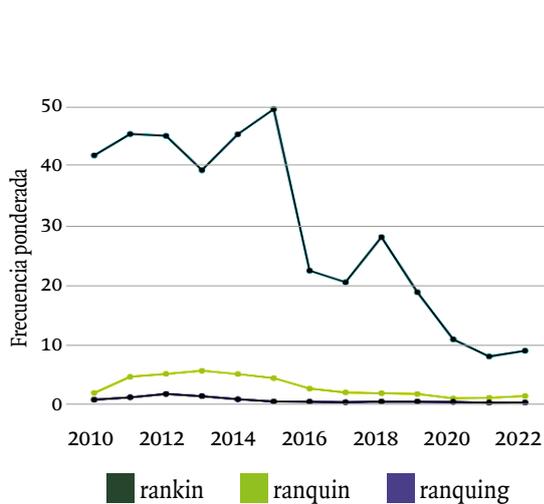
Anexo



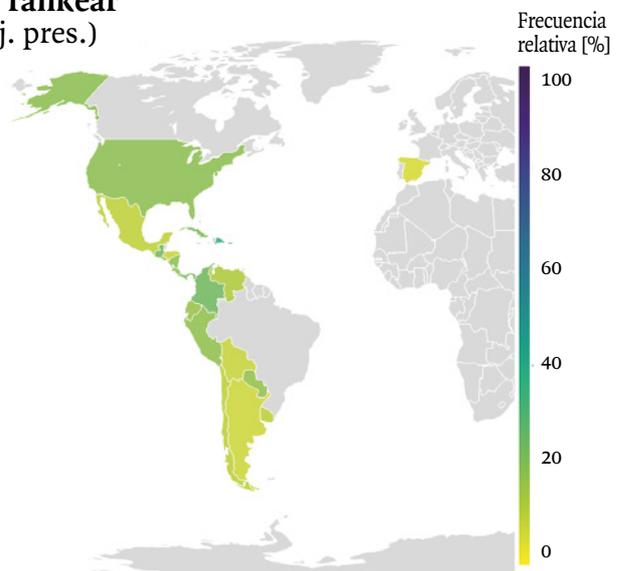
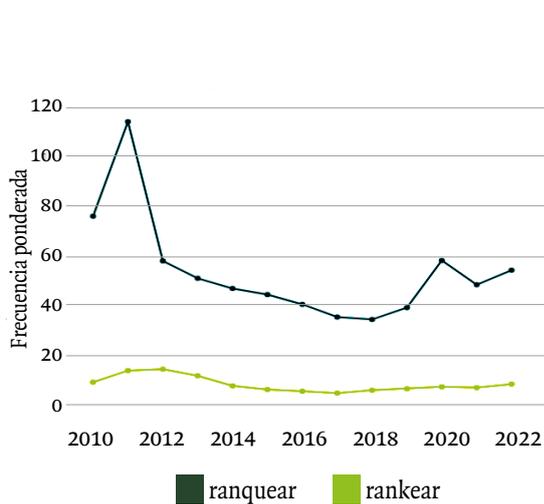
dilogía vs. bilogía
 sustantivos (singular y plural)



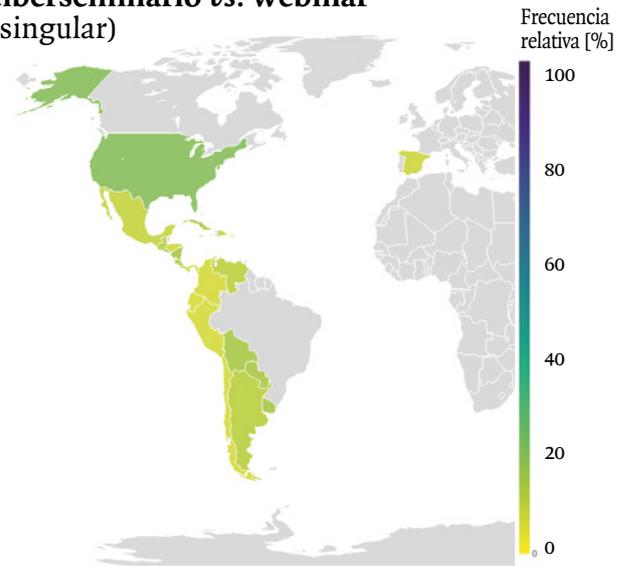
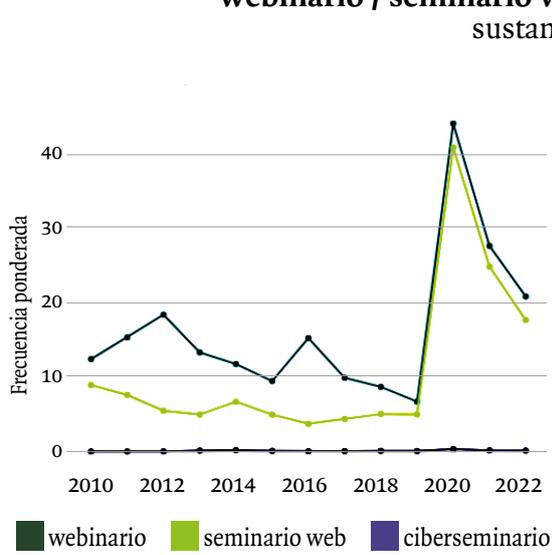
rankin/ranquin/ranquing vs. ranking
 sustantivos (singular)



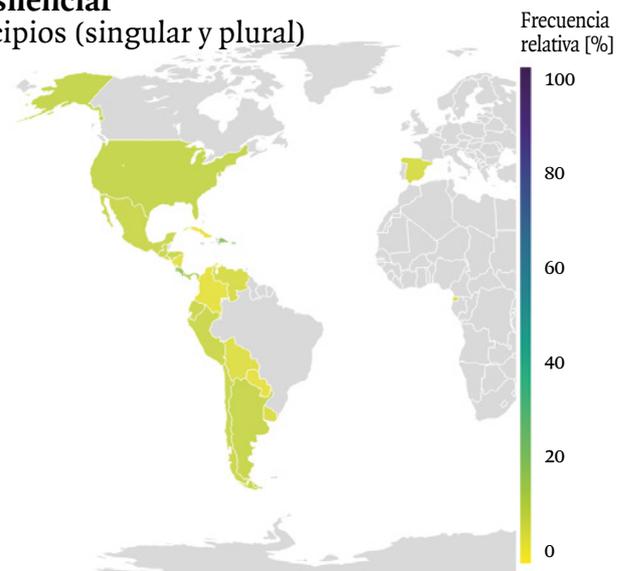
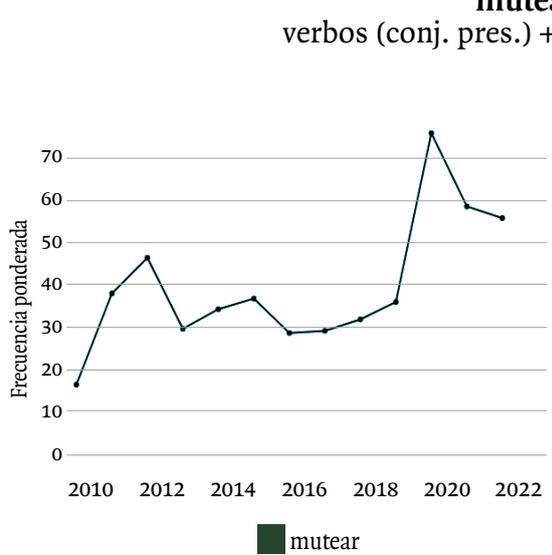
ranquear vs. rankear
 verbos (conj. pres.)



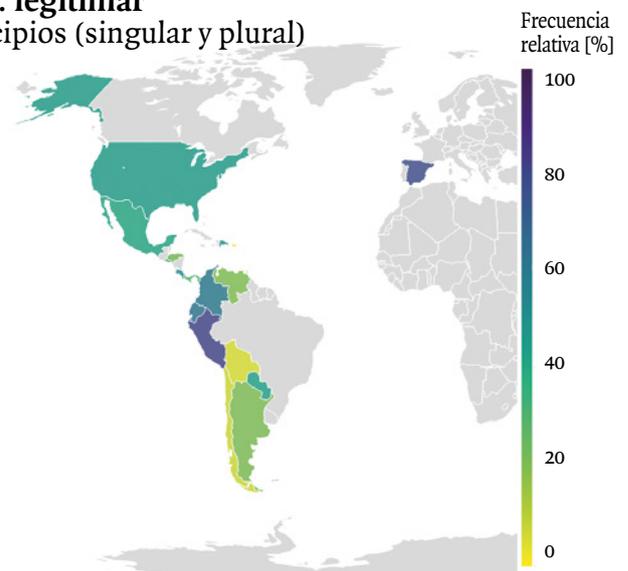
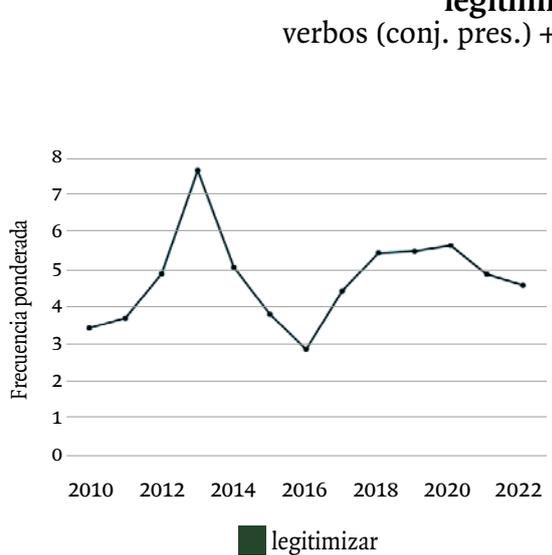
webinario / seminario web / ciberseminario vs. webinar
 sustantivos (singular)



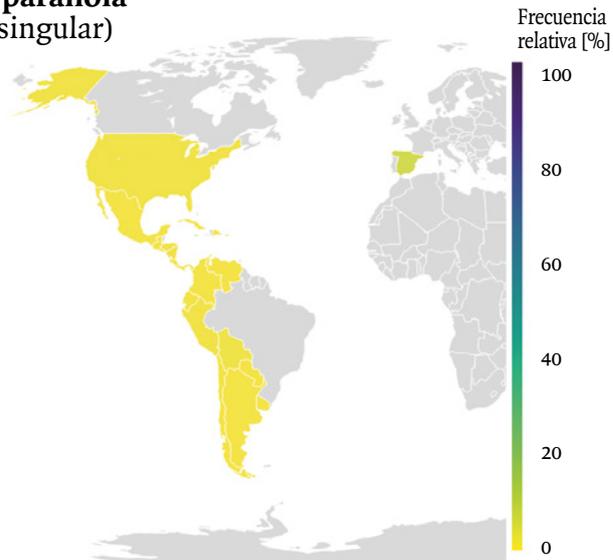
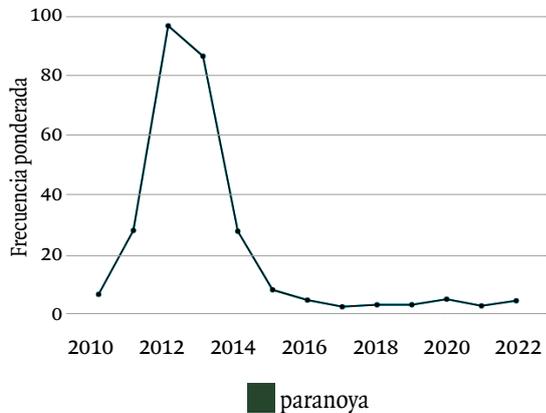
mutear vs. silenciar
 verbos (conj. pres.) + participios (singular y plural)



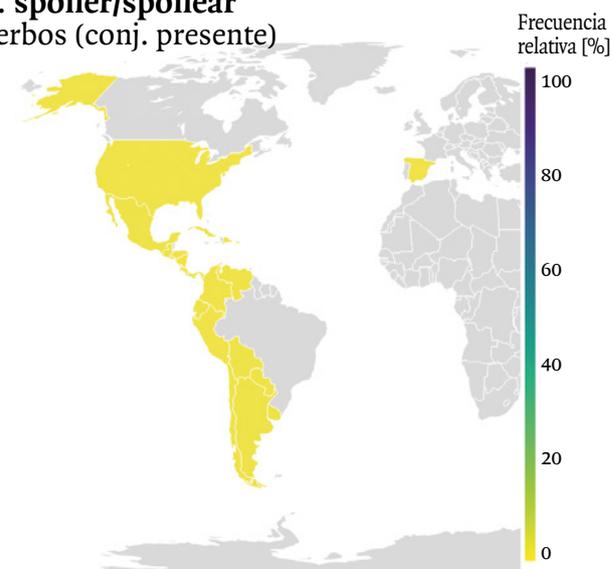
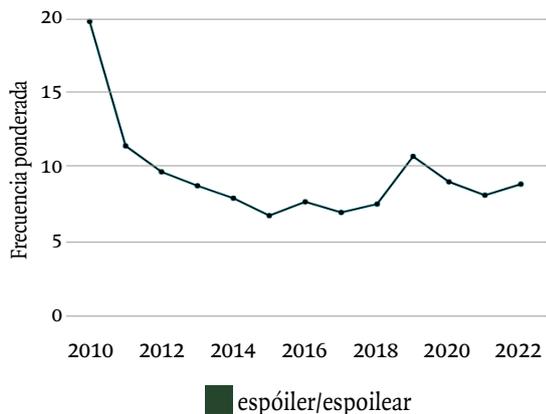
legitimizar vs. legitimar
 verbos (conj. pres.) + participios (singular y plural)



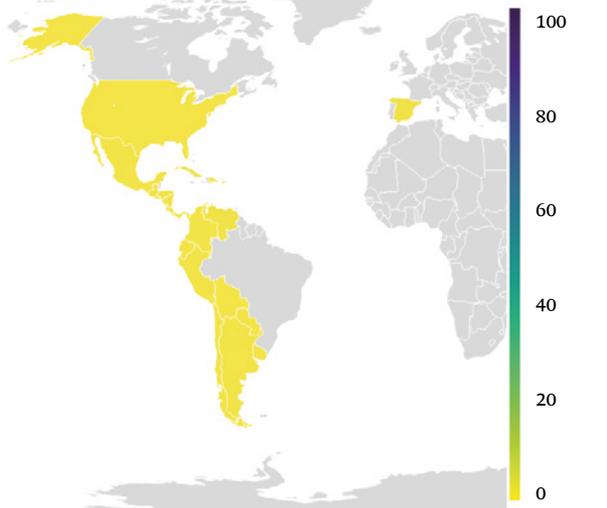
paranoya vs. paranoia
 sustantivos (singular)



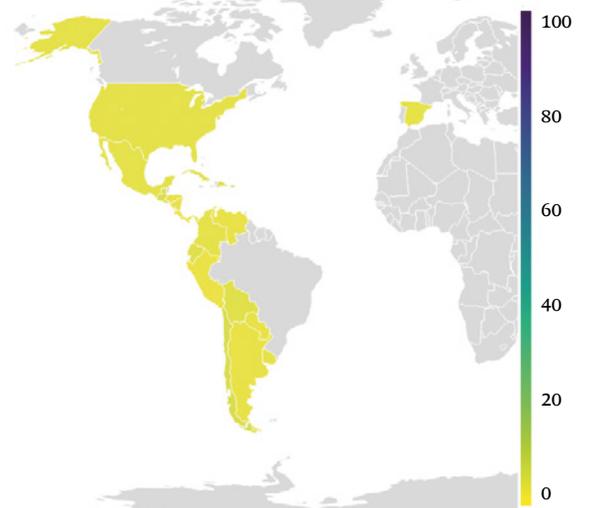
espóiler/espoilear vs. spoiler/spoilear
 sustantivos (singular) + verbos (conj. presente)



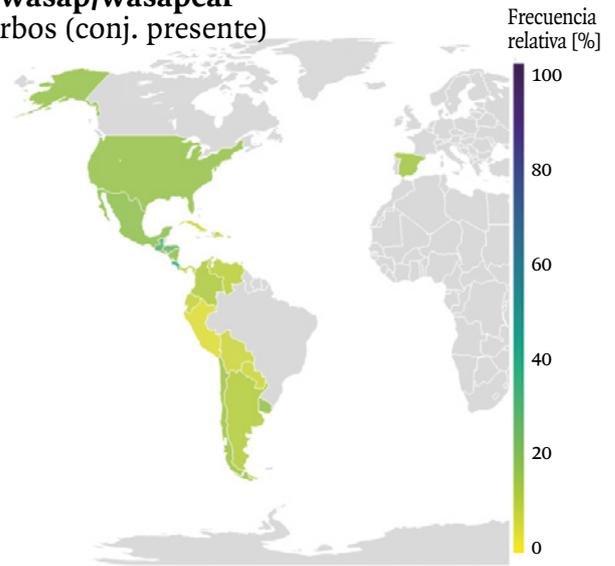
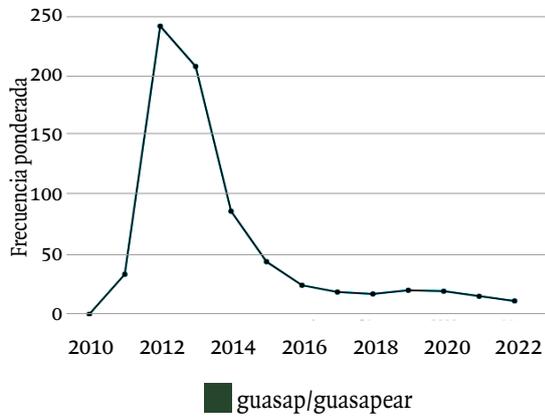
espóiler vs. spoiler
 sustantivos (singular)



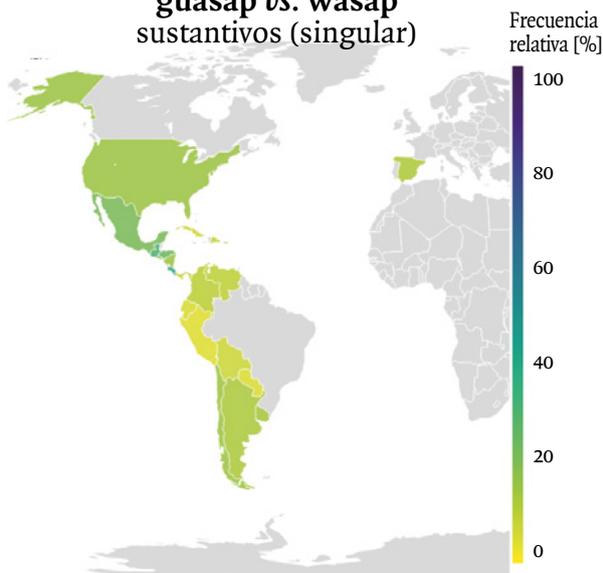
espoilear vs. spoilear
 verbos (conj. presente)



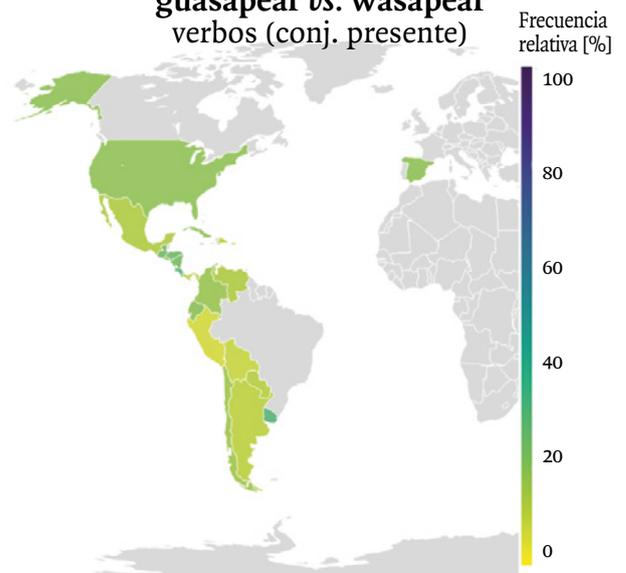
guasap/guasapear vs. wasap/wasapear
 sustantivos (singular) + verbos (conj. presente)



guasap vs. wasap
 sustantivos (singular)



guasapear vs. wasapear
 verbos (conj. presente)



guglear vs. googlear
 verbos (conj. presente)

