



RILEX
REVISTA SOBRE INVESTIGACIONES LÉXICAS

VOLUMEN IX - NÚMERO 2
JULIO, 2026

Marina Albers
Santiago Rodríguez-Rubio
Blanca Fernández Soriano
Francisco Escudero Paniagua
Natalia Alarcón Mosquera

REVISTAS CIENTÍFICAS DE LA UNIVERSIDAD DE JAÉN

<https://doi.org/10.17561/rilex.9.12>

Los estudios e investigaciones que se recogen en esta revista están sujetos a una licencia de reconocimiento de *Creative Commons*. Esta licencia permite **compartir** (copiar y redistribuir el material en cualquier medio o formato) y **adaptar** (remezclar, transformar y construir a partir del material para cualquier propósito, incluso comercialmente) el material siempre que se indique adecuadamente el origen y los cambios.

CONSEJO EDITORIAL

EDITORA

Dr.^a M.^a Águeda Moreno Moreno (Universidad de Jaén)

DIRECTOR EDITORIAL

Dr. Jesús Camacho Niño (Universidad de Jaén)

SECRETARÍA

Dr.^a Marta Torres Martínez (Universidad de Jaén)

CONSEJO DE REDACCIÓN

DIRECCIÓN

Dr.^a M.^a Águeda Moreno Moreno (Universidad de Jaén)

SUBDIRECCIÓN/SECRETARÍA

Dr.^a Marta Torres Martínez (Universidad de Jaén)

VOCALES

Dr.^a Eleni Leontaridi (Aristotle University of Thessaloniki)

Dr.^a Elisabeth Fernández Martín (Universidad de Almería)

Dr. Francisco Pedro Pla Colomer (Universidad de Jaén)

Dr. Jesús Camacho Niño (Universidad de Jaén)

Dr. Matías Hidalgo Gallardo (Università degli Studi di Bergamo)

Dr. Narciso Contreras Izquierdo (Universidad de Jaén)

Dr. Tibor Berta (Universidad de Szeged)

Dr.^a Victoria Rodrigo (Georgia State University)

EQUIPO TÉCNICO

EDITOR TÉCNICO

Dr. Jesús Camacho Niño

ASISTENCIA TÉCNICA

Alicia Arjonilla Sampedro (Universidad de Jaén)

Inmaculada Ruiz Sánchez (Universidad de Jaén)

COMITÉ CIENTÍFICO

Ángel López García-Molins, Universidad de Valencia, España

Cecilio Garriga Escribano, Universidad Autónoma de Barcelona, España

Concepción Maldonado González, Universidad Complutense de Madrid, España

Dolores Azorín Fernández, Universidad de Alicante, España

Giuseppe Trovato, Universidad de Venecia, Italia

Gloria Clavería Nadal, Universidad Autónoma de Barcelona, España

Humberto Hernández Hernández, Universidad de La Laguna, España

Josefina Prado Aragonés, Universidad de Huelva, España

José Ignacio Pérez Pascual, Universidad de A Coruña, España

José Ramón Carriazo Ruiz, Universidad Nacional del Educación a Distancia, España

Mar Campos Souto, Universidad de Santiago de Compostela, España

Mar Cruz Piñol, Universidad de Barcelona, España

M.^a Luisa Calero Vaquera, Universidad de Córdoba, España

Marta Higuera García, Instituto Cervantes, España

Matteo de Beni, Universidad de Verona, Italia

Pedro Fuertes-Olivera, Universidad de Valladolid, España

Stefan Ruhstaller, Universidad Pablo de Olavide, España

Sven Tarp, Universidad de Aarhus, Dinamarca

ÍNDICE

Marina Albers

Tratamiento de los *-ismos* en la lexicografía americana: los diccionarios
del Río de la Plata de los siglos XIX y XX.....7

Santiago Rodríguez-Rubio

De la “a” a la “az”. Estudio de las anomalías formales del texto en español
del Cambridge Dictionary Online (English-Spanish/Spanish-English)37

Blanca Fernández Soriano

Lexico-semantic construction of fear in Maya Angelou’s autobiography of girlhood..... 71

Francisco Escudero Paniagua

¿Carcamusas para los carcas y las musas? La historia del nombre del plato toledano..... 103

Natalia Alarcón Mosquera

Orientaciones para instruir al profesorado de E/L2 en la enseñanza de verbos frecuentes:
el caso de ‘tenera’..... 125



TRATAMIENTO DE LOS *-ISMOS* EN LA LEXICOGRAFÍA AMERICANA: LOS DICCIONARIOS DEL RÍO DE LA PLATA DE LOS SIGLOS XIX Y XX

TREATMENT OF *-ISMS* IN AMERICAN LEXICOGRAPHY: THE DICTIONARIES OF THE RÍO DE LA PLATA IN THE 19TH AND 20TH CENTURIES

Marina Albers

Universidad de Salzburgo

marina.albers@plus.ac.at

RESUMEN

En la presente contribución, nos proponemos estudiar una serie de obras lexicográficas del área del Río de la Plata del siglo XIX —como época inicial de la lexicografía americana propia tras las independencias de las nuevas naciones— que presentan los indigenismos y ruralismos como expresión de la identidad rioplatense, y obras del siglo XX, enfocadas mayoritariamente en el concepto de americanismo. La contribución se centrará en la pregunta de cómo los diferentes *-ismos* —americanismos, argentinismos, uruguayismos, indigenismos, ruralismos, regionalismos y provincialismos— son clasificados, definidos y valorados en los diccionarios rioplatenses, con vistas a su comparación. Además, en un pequeño estudio metalexicográfico cualitativo basado en un corpus epistolar inédito del siglo XVIII, trataremos de analizar una selección representativa de siete lemas de origen patrimonial o indígena con el fin de contrastar y comparar el tratamiento de estos *-ismos* en las obras lexicográficas presentadas. Si bien los términos son conocidos y usados o bien en toda América o bien principalmente en la región del Río de la Plata —denominando siempre realidades y referentes propios del continente americano o de la región—, no hay consenso en cuanto a la aplicación de los *-ismos*.

Palabras clave: americanismo, argentinismo, diccionarios, indigenismo, lexicografía rioplatense, uruguayismo.

ABSTRACT

In this contribution, we propose to study a series of lexicographical works from the Río de la Plata area in the 19th century—as the initial period of American lexicography following the independence of the new nation—which present indigenous and rural terms as expressions of Río de la Plata identity, as well as works from the 20th century, which focus for the most part on the concept of americanism. The contribution will focus on the question of how these different *-isms*—americanisms, argentinisms, uruguayisms, indigenisms, ruralisms, regionalisms and provincialisms—are classified, defined and evaluated in Río de la Plata dictionaries, with a view to comparing them. In addition, in a small qualitative metalexicographical study based on an unpublished 18th-century epistolary corpus, we will attempt to analyse a representative selection of seven entries of patrimonial or indigenous origin in order to contrast and compare the treatment of these *-isms* in the lexicographical works presented. Although the terms are known and used either throughout the Americas or mainly in the Río de la Plata region—always referring to realities and references specific to the American continent or the region—there is no consensus on the application of the *-isms*.

Keywords: Americanism, Argentinism, dictionaries, indigenism, Río de la Plata Lexicography, Uruguayism.

Recibido: 01-10-2025
Aceptado: 31-03-2026

DOI: <https://doi.org/10.17561/rilex.9.2.9979>



1. INTRODUCCIÓN

En la época colonial, la lexicografía o, mejor dicho, la proto-lexicografía de América, empezó paralelamente a la expansión española (Pöll, 2018, p. 120), ya que tanto en crónicas como en glosas de obras historiográficas, por ejemplo, pueden encontrarse referencias y menciones de voces nuevas o adaptadas para referirse a las nuevas realidades con las que los conquistadores entraron en contacto. Del mismo modo, en el ámbito de la lingüística misionera, las muy diversas lenguas y culturas indígenas fueron descritas en gramáticas, vocabularios, catecismos, etc. si bien no con un fin meramente lingüístico, sino principalmente con un objetivo práctico en el marco de la tarea evangelizadora (Zimmermann, 2016, p. 169).

En esta primera fase de contacto, los españoles no poseían términos adecuados en castellano para referirse a “la fauna, la flora, los accidentes geográficos, la toponimia y aún [...] las relaciones sociales, la vida cultural y la organización institucional” (Fontanella de Weinberg, 1992, p. 97) de América. Por lo tanto, se vieron en la necesidad de elaborar estrategias para lidiar con este desafío lingüístico.

Por un lado, adaptaron las voces patrimoniales existentes para referirse a las nuevas realidades, lo cual generó, en consecuencia, una ampliación semántica. El empleo de la palabra española *león* para referirse a ‘puma’ o del adjetivo *perezoso* para nombrar al mamífero que vive en zonas tropicales de América, sirven como ejemplos de la adaptación metafórica de un término existente a un nuevo referente sin cambios morfológicos, solo a base de semejanza y comparabilidad del concepto conocido con el desconocido (Bravo-García & Cáceres-Lorenzo, 2011, pp. 59-60; Ramírez Luengo, 2007, p. 73). De estas voces patrimoniales que ampliaron su dimensión semántica se diferencian aquellos términos españoles adaptados que sufrieron además cambios morfológicos en la forma o que fueron acompañados de complementos para la diferenciación del concepto conocido del nuevo. Cabe mencionar, por ejemplo, las derivaciones a través de las cuales fueron formadas

nuevas palabras para referirse a conceptos desconocidos, como la sufijación mediante *-ada* en *caballada* o *potrada*, para el primer caso, o especificaciones a través de ‘de allá’ o ‘de la tierra’ como aclaraciones complementarias en *jazmín de la tierra* o *conejillo de Indias*, en el segundo (Lapesa, 2008[1981], p. 499; Ramírez Luengo, 2007, p. 73). Esta primera estrategia de adaptación y ampliación del léxico patrimonial a las nuevas circunstancias y realidades americanas suele llevar a la creación de *americanismos*¹. Estos o bien se difundieron por toda América o gran parte del continente o llegaron incluso a España, o bien quedaron restringidos a ciertos límites geográficos, constituyendo de este modo *regionalismos* o *provincialismos*.

Por otro lado, en casos en los que los conceptos o realidades americanos carecían de referentes parecidos en el imaginario europeo conocido, los españoles introdujeron términos prestados de las múltiples lenguas indígenas. De esta manera, los primeros *indigenismos* entraron al español en fechas muy tempranas en la primera fase de la Conquista y provenían de las lenguas autóctonas de la región antillana: del taíno y del caribe vienen palabras como *canoa*, *cacique*, *maíz* y *tiburón* (Bravo-García & Cáceres-Lorenzo, 2011, p. 65; Lapesa, 2008[1981], p. 465; Ramírez Luengo, 2007, pp. 75-76). En cuanto a la incorporación de las voces de origen indígena, cabe diferenciar, por una parte, aquellos préstamos rápidamente conocidos, difundidos e integrados al español, y por tanto capaces de aparecer en los textos coloniales sin glosas explicativas, y por otra parte, aquellos que requerían comentarios metalingüísticos —tales como descripciones, sinónimos, definiciones o traducciones— como modo de proceder necesario ante la nula o escasa integración de los términos en el español (Enguita Utrilla, 2010, pp. 202-215). Mientras que los préstamos tempranos se expandieron generalmente en todo el mundo

¹ Véanse las explicaciones acerca de los americanismos en el apartado 2. del presente trabajo.

hispanico², los indigenismos provenientes de lenguas con las que los españoles entraron en contacto posteriormente o bien quedaron restringidos a sus regiones de origen —tal como es el caso de los indigenismos de lenguas como el maya³ de Mesoamérica, del aimara o del mapudungun de América del Sur—, o bien se extendieron más allá de su geografía y temporalidad originales en el caso de las ‘lenguas generales’⁴ náhuatl, quechua y guaraní (Ramírez Luengo, 2007, pp. 76-79):

La situación de las lenguas indígenas que proporcionan préstamos en etapas posteriores a este primer momento antillano es muy diferente. Algunas fueron consideradas lenguas generales por la Corona española y alcanzaron una amplia difusión que refuerza su vitalidad actual; las palabras que aportan perviven, en este caso, junto a la lengua de las que proceden (Bravo-García & Cáceres-Lorenzo, 2011, p. 138).

Terminada la época colonial, la creación de estados independientes desembocó en América —junto al surgimiento de identidades americanas propias— también en una nueva conciencia acerca de las características americanas de la lengua española y de las variedades lingüísticas nacionales desde un nuevo punto de vista propiamente americano y emancipado (Billaudelle, 1999, p. 87; Pöll, 2018, p. 120), como ilustran, por ejemplo, los discursos de ciertos próceres y líderes de los nuevos estados americanos, quienes recurrían a menudo a ciertos rasgos lingüísticos hoy considerados características

²La rápida difusión de los primeros indigenismos se debe, entre otros aspectos, a las relaciones comerciales entre Europa y América y al intercambio de productos exóticos, así como a la literatura cronística y a la burocracia colonial que preveía la documentación minuciosa de los sucesos en América (Ramírez Luengo, 2007, p. 75).

³Ramírez Luengo y Quirós García (2015, pp. 199-201), por ejemplo, confirman la escasa presencia de mayismos —en comparación con indigenismos provenientes del náhuatl o de las lenguas antillanas— incluso en las zonas en las que el maya era la lengua habitual, como lo constituye el español de Yucatán de los siglos XVII y XVIII estudiados.

⁴El término de ‘lengua general’ describe una lengua indígena que servía como *lingua franca* o *koiné*, o sea, como vehículo de comunicación suprarregional, entre otros para el objetivo de la evangelización, entre los españoles y la población indígena —incluso la que poseía una lengua materna de menor relevancia— más allá de su territorio de uso original. Véase Ramírez Luengo (2007, pp. 77-79) para explicaciones acerca de la expansión y los aportes léxicos de las lenguas generales.

americanas (Frago Gracia, 2011)⁵. De este modo, las consecuencias normativas a nivel metalingüístico fueron, por una parte, una gramaticografía americana, de la que surgió la obra de Andrés Bello como máximo reflejo de la identidad americana en una propia norma lingüística hispánica⁶, y por otra parte, el comienzo de una lexicografía americana estandarizadora en el siglo XIX, que posibilitó el desarrollo de vocabularios y diccionarios regionales y nacionales, incluyendo asimismo elementos indígenas (Sánchez Méndez, 2016, pp. 21-37). No obstante, la nueva conciencia identitaria y lingüística de los Estados nacientes mantuvo asimismo viva la idea de la posesión de una lengua vehicular en común, el español, y de la unidad idiomática (Chávez Fajardo, 2022, pp. 129-130; pp. 139-140), de ahí que tanto el legado español como lo propio convivieran —aunque con diferentes ideas lingüísticas e ideologías— en el imaginario americano de la época⁷.

La reflexión acerca de los rasgos identitarios reflejados en la lengua empieza a desarrollarse también en la lexicografía del siglo XIX, diferenciando siempre lo propio de la norma peninsular⁸ y constituyendo el diccionario “un

⁵ Frago Gracia (2011) analiza en una amplia serie de documentos de personajes ilustres como Rosas, Sarmiento, Belgrano, Bolívar, San Martín, etc., la sustitución de *vosotros* por *ustedes* y las estructuras posesivas de tercera persona.

⁶ El criollo venezolano Bello publicaría, entre otras muchas obras, en 1823 las *Indicaciones sobre la conveniencia de simplificar i uniformar la ortografía en América* como esbozo de una ortografía americana propia, y finalmente, en 1847, la *Gramática de la lengua castellana destinada al uso de los americanos* como máxima expresión de la identidad americana en una norma lingüística propia, pero no con el objetivo de alejarse de las propuestas de la Real Academia Española en cuanto al español peninsular, sino con el propósito de encontrar igualdad y de conservar el legado español.

⁷ Cabe señalar la relevancia de los paratextos como prólogos, introducciones o advertencias de los diccionarios, ya que estos reflejan siempre la ideología —sea purista, nacionalista, romántica, racionalista, etc.— con la que el respectivo diccionario está confeccionado. Por tanto, puede aplicarse el concepto de *ideologema* a la hora de estudiar la lexicografía americana (Chávez Fajardo, 2022, pp. 119-120). Véase para la discusión de las diferentes posturas de la época el capítulo correspondiente en Chávez Fajardo (2022, pp. 119-142).

⁸ Las razones por las cuales el español peninsular constituye la variedad de referencia de contraste por antonomasia en la lexicografía diferencial son varias. Cabe mencionar, por ejemplo, la importante tradición lexicográfica, la irradiación cultural y lingüística de España, el número de grandes editoriales o simplemente el hecho de que la variedad peninsular seguía siendo la más conocida y la mejor descrita (Chávez Fajardo, 2022, pp. 48-49; Haensch & Werner, 1993a, p. XX).

espacio de memoria” que refleja el ingenio, las costumbres, etc. de su pueblo (Chávez Fajardo, 2022, pp. 120-121). Por lo tanto, los primeros diccionarios solían ser de naturaleza contrastiva y diferencial, es decir, que comparaban la respectiva variedad regional o nacional de América con la variedad peninsular (Billaudelle, 1999, p. 87; Barcia, 2006, pp. 18-19; Lauria, 2011, p. 112; Pöll, 2018, p. 120; Coll, 2024, p. 482).

Si bien son archiconocidas y frecuentemente usadas las designaciones como *argentinismo*, *mexicanismo*, *peruanismo*, etc., cabe tener en cuenta que carecen de una definición inequívoca en la lexicografía histórica americana. Estos *-ismos* siempre parten de la idea de una norma monocéntrica, la peninsular, de la que las variedades nacionales americanas se distancian debido a sus peculiaridades léxicas nacionales (Huisa Téllez, 2024, p. 474).

En el caso concreto del Río de la Plata, la producción de instrumentos lexicográficos propios comenzó en los años 1870 como consecuencia del proceso emancipador de las naciones de Argentina y de Uruguay. Los diccionarios confeccionados no sirvieron solamente como instrumentos lingüísticos, sino también como testimonios culturales, históricos, discursivos e identitarios (Lauria, 2011, p. 107, pp. 116-117).

2. LOS -ISMOS

Los *-ismos* más frecuentes que se nos presentan en los diferentes diccionarios del Río de la Plata son los propios *argentinismos* y *uruguayismos*, aparte de los *americanismos* e *indigenismos*. También suelen aparecer los términos *ruralismo*, *regionalismo* y *provincialismo*.

Como ya hemos dicho, los conceptos y las perspectivas que abordan estos *-ismos* no son homogéneos en la lexicografía americana, ya que pueden referirse a conceptos y realidades muy diferentes desde distintas perspectivas geográficas, temporales e ideológicas.

No obstante, por una parte, los *regionalismos* y *provincialismos*, así como los *ruralismos* poseen caracterizaciones bastante homogéneas. *Ruralismos*

son las voces que designan conceptos de la vida campestre, tales como flora, fauna, geografía, costumbres, hábitos, etc. Los *regionalismos* tienen un uso y una difusión exclusiva en una región, más allá de las fronteras políticas, como lo es la rioplatense, mientras que los *provincialismos* se limitan a una determinada provincia, independientemente de las fronteras nacionales.

En cambio, los *indigenismos*, si bien coinciden en el aspecto de que proceden de una lengua indígena de América, pueden incluir o bien las voces de origen indígena que forman parte del español general y son conocidas en toda América y en España, o bien únicamente las que se limitan a la zona de su influencia histórica. Además, algunos indigenismos se adaptaron al sistema fónico y morfológico del español, mientras que otros pervivieron en su forma original.

Por otra parte, los *americanismos*, al igual que los *argentinismos* y *uruguayismos*, no poseen delimitaciones tan homogéneas en las obras lexicográficas americanas de los siglos XIX y XX. A este respecto, resulta convincente la definición de *mexicanismo* propuesta por Company Company (2007) desde una perspectiva meramente sincrónica, que puede aplicarse tanto a los americanismos⁹ en general como a los argentinismos y uruguayismos o, más generalmente, a *x-ismos*:

Entenderemos por [x-ismo] [...] el conjunto de voces, formas o construcciones que son caracterizadoras del habla urbana, popular o culta, o ambas, de este país [x, o de América en general] [...] y cuyo uso muy frecuente y cotidiano distancia la variante [x] respecto del español peninsular (Company Company, 2007, pp. 28-29; los cambios marcados a favor de una aplicación general son nuestros).

La misma autora diferencia tres tipos de mexicanismos (Company Company, 2010), que permiten una paralela distinción de tres tipos de *americanismos*¹⁰: (1) los *americanismos puros* se definen como aquellas voces que se emplean en el español general de América y que son inexistentes en el

⁹ Ramírez Luengo (2014, p. 2) propone la aplicación de *mexicanismo* y de los tres tipos según Company Company (2007; 2010) al *americanismo*.

¹⁰ Trataremos clasificar los lemas analizados en nuestro estudio metalexiconográfico en el apartado 4. según estos tres tipos de americanismos.

español peninsular general; (2) los *americanismos de frecuencia* son voces o formas compartidas tanto en forma como en significado con el español peninsular, pero que se destacan en América por una frecuencia de empleo y una generalización mucho mayor que en España; y (3) los *americanismos semánticos*, entendidos como voces y formas que, si bien son compartidas con el español peninsular, presentan en América valores semánticos propios (Company Company, 2010, pp. XVII-XVIII; Ramírez Luengo, 2014, pp. 2-3). No obstante, las obras lexicográficas del Río de la Plata aquí consultadas carecen de estas perspectivas y definiciones y ofrecen conceptos propios.

De este modo, nos proponemos en esta contribución ocuparnos de las siguientes preguntas desde un punto de vista metalexicográfico: En primer lugar, quisiéramos comparar las definiciones de los mencionados *-ismos* a lo largo de los siglos XIX y XX en la lexicografía en Argentina y Uruguay. Nos interesa, además, centrarnos en las diferencias entre unos *-ismos* y otros. En segundo lugar, nos enfocaremos en las valoraciones que los diferentes *-ismos* experimentan en las obras lexicográficas consultadas, así como en la relevancia que tienen para la cultura, la identidad y la lengua nacionales. En tercer lugar, trataremos de comparar en un pequeño estudio metalexicográfico cualitativo una serie de lemas ejemplares de estos *-ismos* en varios diccionarios rioplatenses de los siglos XIX y XX con el fin de descubrir y comentar el empleo de las diferentes voces de origen patrimonial o indígena en datos históricos reales y su tratamiento en la lexicografía americana regional.

3. LOS -ISMOS EN ALGUNOS DICCIONARIOS RIOPLATENSES DE LOS SIGLOS XIX Y XX

A continuación, iremos presentando una serie no exhaustiva de diccionarios y vocabularios ejemplares, unos muy conocidos y otros menos, confeccionados en territorio argentino o uruguayo, respectivamente, y las posturas y definiciones de los mencionados *-ismos*, comenzando en orden cronológico con algunas obras decimonónicas del Río de la Plata.

3.1. LOS DICCIONARIOS DEL SIGLO XIX

Los diccionarios y vocabularios que comenzaron a surgir en todas las recién creadas naciones americanas a partir de la segunda mitad del siglo XIX trataron de presentar los rasgos del habla regional o nacional como propios del país, de sus provincias o bien de regiones enteras, más allá de las fronteras políticas recientemente formadas, pero manteniendo siempre el contraste con el español peninsular como variedad de referencia.

3.1.1. *Un Inédito Diccionario de argentinismos del siglo XIX*

Un Inédito Diccionario de argentinismos *del siglo XIX* (IDA) se basa en el *Diccionario de Argentinismos*, originalmente elaborado entre 1875 a 1879 por la Academia Argentina de Ciencias, Letras y Artes, pero jamás publicado. Finalmente, el equipo editorial de la Academia Argentina de Letras digitalizó y editó el primer diccionario sobre el léxico argentino en 2006 (Barcia, 2006, pp. 11-24).

En este diccionario, aunque lleva en el título el término de *argentinismo*¹¹, entendido como voz de uso exclusivo en Argentina, la Academia trata de evitar el término y de sustituirlo por el concepto menos restringido *habla de los argentinos*. De este modo, son incluidos los vocablos de uso cotidiano del pueblo argentino que son patrimonio no solamente de la región del Río de la Plata —refiriéndose a Argentina, Uruguay y Paraguay— y de países vecinos como Chile y Bolivia, sino también conocidos en otras naciones americanas, unidas a través de una lengua (Barcia, 2006, pp. 23-24). Por lo tanto, se trata de un diccionario confeccionado según un criterio de uso y no de origen, es decir, que incluye las voces usadas en Argentina y las zonas vecinas —además de contrastar los significados rioplatenses con los del diccionario de la

¹¹ Barcia (2006, p. 24, p. 36) hace hincapié en el hecho de que no puede hablarse de *argentinismo* ni de un léxico nacional hasta que no se conozca con certeza qué vocablos son realmente propios y de uso exclusivo de Argentina. Lo mismo valdría, por consiguiente, para cada *-ismo* que insinúa ser de empleo limitado exclusivamente a una nación.

RAE¹²— y no solamente las que tienen su origen ahí o en una de las lenguas indígenas de la región.

3.1.2. *Lenguaje del Río de la Plata*

Washington Bermúdez empezó a elaborar a partir de 1880 el *Lenguaje del Río de la Plata* (LRP), cuya publicación estaba planeada en tres volúmenes temáticos, a saber, verbos, voces y refranes. No obstante, nunca llegó a publicarse sino la parte que iba desde “a asunto de que” hasta “acomodar”. Tras el fallecimiento del autor, la labor fue retomada por su hijo, Sergio Bermúdez.

Los originales de este diccionario, que incluyen un total de 24 volúmenes con más de 9000 folios manuscritos o mecanografiados, se encuentran actualmente en el archivo de la Academia Nacional de Letras de Uruguay, en su sede en Montevideo (Da Rosa & Lucián, 2016)¹³.

El diccionario de Bermúdez reúne las voces usadas en Argentina, Uruguay y en Paraguay, de ahí que constituyera el primer diccionario que comprende la región rioplatense entera, incluyendo la paraguaya. Por lo tanto, el LRP es considerado un diccionario de regionalismos. La crítica de Bermúdez a la lexicografía americana de aquel entonces se centraba en el hecho de que solían ridiculizarse las voces corrientes en la región, ya que fueron designadas *barbarismos* o *provincialismos*, dos designaciones que no dan cuenta del valor del habla rioplatense y de la dimensión cultural e identitaria de la región que está detrás de la lengua (Billaudelle, 1999, p. 238).

¹² A modo de ejemplo, el IDA menciona aparte de las acepciones rioplatenses de los lemas *estancia* o *rancho* también los significados según la RAE (*estancia*: ‘mansión, habitación y asiento; cualquier lugar, casa o paraje’; *rancho*: ‘la reunión de los soldados que, en forma de rueda, comen juntos’), siempre con la indicación diatópica “Esp.” (Barcia, 2006, s. v. *estancia*; s. v. *rancho*).

¹³ Agradezco a la Academia Nacional de Letras de Uruguay por facilitarme el acceso a los originales de los tomos en abril de 2023, especialmente a Virginia Bertolotti por la invitación y la gestión y a Juan Justino da Rosa por su cordial recepción en la sede en Montevideo, así como por darme a conocer la historia detrás del diccionario.

3.1.3. *Vocabulario rioplatense razonado*

Daniel Granada, oriundo de Galicia, pero residente en Uruguay desde su infancia, confeccionó el *Vocabulario rioplatense razonado* (DRR), que cuenta con dos tempranas ediciones, de 1889 y 1890, y una tercera de 1957, así como con una reedición de la segunda, elaborada por Kühl de Mones en 1998. El vocabulario contiene un inventario de los americanismos de la región del Plata —área correspondiente a la zona conquistada por Juan Díaz de Solís y al virreinato del Río de la Plata— de los que no hace mención el diccionario de la RAE.

Las unidades léxicas del vocabulario son denominadas *americanismos del Plata* y siguen un criterio de uso y no de origen. Como ilustra la Tabla 1, son incluidos tanto los americanismos de uso exclusivo de la región del Plata —los *regionalismos* y *provincialismos*— como los americanismos generales. En cuanto al origen, Granada toma americanismos de procedencia indígena y americanismos castellanos. En total, el DRR clasifica como *americanismos del Plata* las siguientes unidades léxicas (Kühl de Mones, 1998, pp. XX-XXVI).

Americanismos del Plata
Voces procedentes de lenguas indígenas de la región: del quichua en el noroeste, del guaraní en el noreste y del araucano en el sur.
Voces de otro origen.
Voces castellanas obsoletas en España, pero usuales en el Río de la Plata.
Voces castellanas con cambios en el significado.
Voces de raíz castellana, pero con cambios morfológicos a través de la derivación.

TABLA 1: *Americanismos del Plata en el DRR (1889-1890)*

Un estudio de Lauria (2011, p. 129) demuestra que el vocabulario de Granada registra los indigenismos sobre todo para los campos semánticos como instrumentos, armas, vestimenta, comidas y bebidas, así como lo que llama *ruralismos* para los conceptos de la vida campestre, o sea, los vinculados con flora, fauna, geografía local, hábitos y costumbres, etc.

La razón por la cual el enfoque del vocabulario de Granada está centrado en la incorporación de *indigenismos* y de *ruralismos* es la Conquista del desierto, es decir, la paulatina expansión del territorio de la República Argentina hacia el sur, tierra hasta entonces habitada únicamente por pueblos indígenas, tales como los mapuches, los tehuelches o los ranqueles. A través de este contacto con hablantes del mapudungun o de las lenguas chon —a las que pertenece la lengua de los tehuelches—, entraron nuevos vocablos al español de la región, y posteriormente también a las obras lexicográficas, sirviendo como preservación de los elementos culturales indígenas como patrimonio histórico y etnográfico cultural de la región entera (Lauria, 2010, pp. 176-179):

En el marco, entonces, de la Conquista del Desierto [...] vemos un movimiento tendiente a atesorar con la idea de preservar ciertos elementos culturales de los pueblos originarios a modo de elementos constitutivos del patrimonio histórico y etnográfico nacional. Entre ellos, se comenzó a estudiar el material léxico en la medida en que dejó profundas marcas en las diferentes variedades del español americano. En este sentido, los *indigenismos* fueron fundamentalmente considerados como objetos de gran valor, de una valía preciosa y preciada debido a su condición de ser rastros de lenguas o bien muertas o bien en vías de extinción y poco estudiadas (Lauria, 2011, pp. 128-129).

Asimismo, el registro de *ruralismos* revaloriza e idealiza la figura del gaucho y la vida campestre como íconos de la tradición argentina (Lauria, 2011, p. 120). Por lo tanto, los *indigenismos* y los *ruralismos* son considerados los elementos que reflejan la identidad regional y nacional argentina en el léxico (Lauria, 2010, p. 187). En cambio, en otros niveles de la lengua, tanto el yeísmo rehilado (Canale & Coll, 2016) como el voseo (Bertolotti, 2015; Carricaburo, 2015), por ejemplo, constituyen rasgos lingüísticos identitarios del español no solo argentino, sino rioplatense¹⁴.

¹⁴ Tanto Carricaburo (2015, p. 27) como Bertolotti (2015, pp. 125-127, p. 142) datan la preponderancia del voseo, especialmente difundido entre indígenas, mestizos y gauchos de las zonas rurales, en la época de las Independencias argentina y uruguaya, respectivamente.

3.2. LOS DICCIONARIOS DEL SIGLO XX

Mientras que los diccionarios del siglo XIX que hemos consultado se centran mayoritariamente en la incorporación de *provincialismos*, *regionalismos*, *indigenismos* y *americanismos del Plata* —y por tanto, también regionalismos— como expresión de la identidad propia en contraste con el español peninsular, los del siglo XX se ocupan generalmente del concepto más bien controvertido de *americanismo* (Coll, 2024, p. 482).

3.2.1. Diccionario de argentinismos, neologismos y barbarismos

En el *Diccionario de argentinismos, neologismos y barbarismos* (DANB) de 1911, el jurista y ensayista Segovia elogia, por una parte, el *Diccionario de la lengua española*, pero critica a la vez la falta de voces, modismos y refranes de Argentina, falta que se debe —según el autor— al vago conocimiento de las realidades americanas por parte de los integrantes de la RAE¹⁵, cuya obra general, sin embargo, valora (Segovia, 1911, pp. 5-7). En cuanto a los conceptos de *americanismo*, *argentinismo* etc., Segovia (1911) propone la clasificación expuesta en la Tabla 2.

-ismos	Contenido
Castellanismos y neologismos	Vocablos conocidos y usados en América y en España, incluso si tienen acepciones distintas.
Americanismos	Palabras usadas en Hispanoamérica y que no se emplean en España. Son voces de origen americano.
Argentinismos (<i>sensu strictu</i>)	Vocablos de uso corriente en el Río de la Plata, o también en Chile, o bien solamente en Argentina. Palabras que no se usan en otros países americanos.

TABLA 2: Clasificación propuesta por Segovia en el DANB (1911)

¹⁵ En cambio, Chávez Fajardo (2022, p. 128-129), por ejemplo, hace referencia a los prólogos de los diccionarios del académico chileno Manuel Antonio Román, quien postula que la integración de variantes diatópicas —como reflejo de nuevas realidades— en los diccionarios americanos, así como la posterior adaptación y difusión de las mismas por la RAE llevarían al enriquecimiento de la lengua, dado que la diversidad léxica forma parte de la homogeneidad y unidad de la lengua española.

En su diccionario, Segovia (1911) distingue, además, las voces extranjeras, las lenguas indígenas, así como los *barbarismos*. En este sentido, clasifica como *voces extranjeras*, que abundan en Buenos Aires por ser en la época “una de las ciudades más adelantadas y cosmopolitas del mundo” (Segovia, 1911, p. 8), las que mantienen su significado, ortografía y pronunciación originales. A las lenguas indígenas de las que los argentinos son herederos, el guaraní, el quichua y el pampa, adscribe un valor cultural importante. Por último, son denominadas *barbarismos* las faltas cometidas por la plebe ignorante, tales como la elisión de la /d/ intervocálica o el uso de imperativos voseantes como *tomá*, *bebé*, etc. (Segovia, 1911, p. 8).

3.2.2. *Nuevo vocabulario campesino rioplatense*

Guarnieri publicó una serie de diccionarios sobre el lenguaje popular rioplatense. En el *Nuevo vocabulario campesino rioplatense* (NVCR), por ejemplo, el autor reúne los términos populares del habla rioplatense, que está compuesta tanto por el habla gauchesca o campesina como por el lunfardo.

En su prólogo, Guarnieri hace referencia a la pervivencia de *arcaísmos* debido al aislamiento de la sociedad rural, a la escasa presencia de *indigenismos*, así como a la introducción de *ruralismos* que se refieren a actividades, utensilios, trabajos campestres, etc. (Guarnieri, 1957, pp. 9-16).

3.2.3. *Nuevo Diccionario de Americanismos*

Los tomos II y III del *Nuevo Diccionario de Americanismos* forman parte de un vasto proyecto de colaboración dirigido a partir de los años 1970 desde Augsburgo por Haensch y Werner, que persigue el objetivo de confeccionar diccionarios de cada uno de los actuales países hispanoamericanos. “[E]l proyecto de Augsburgo tenía como objetivo central desde el primer momento renovar aquella lexicografía diferencial del español de América originada en los nuevos países independientes ya bien entrada la segunda mitad del siglo

XIX” (Huisa Téllez, 2024, p. 474). Hasta el momento, contamos tan solo con los tomos sobre *colombianismos*, *cubanismo*s, *argentinismos* (NDA) y *uruguayismos* (NDU), que fueron elaborados junto con equipos nacionales y publicados finalmente en 1993.

Estos *Nuevos Diccionarios de Americanismos* se destacan por su carácter sincrónico, descriptivo y diferencial, esto es, que abarcan las voces consideradas ‘normales’ en la respectiva variedad nacional en la segunda mitad del siglo XX, tomando como referencia el español peninsular (Haensch & Werner, 1993a, pp. XIX-XX; Haensch & Werner, 1993b, pp. XVII-XVIII; Pöll, 2018, pp. 121-122; Chávez Fajardo, 2022, pp. 45-48).

En este sentido, son clasificados como *argentinismos* o *uruguayismos*, respectivamente, ni solo conceptos o referentes americanos ni tampoco solo unidades léxicas creadas en territorio uruguayo o argentino, sino que presentan cómo se denomina un concepto en Uruguay o Argentina, dentro de los límites del país. De este modo, registran solo unidades léxicas que se emplean en Uruguay o Argentina, respectivamente, pero no en España.

3.2.4. *Nuevo Diccionario de Americanismos e Indigenismos*

El *Nuevo Diccionario de Americanismos e Indigenismos* (NDAI) fue elaborado por Morínigo, oriundo de Paraguay, y publicado en 1998. Tomando como fuente principal la lengua escrita, el autor reúne en el diccionario una amplia serie de términos, difundidos y familiares en todo el continente, en amplias áreas del mismo o bien en zonas restringidas (Morínigo, 1998, pp. 9-11).

Si bien enumera en el título de su obra los *americanismos* y los *indigenismos*, aclara que los segundos forman parte de la primera categoría, ya que el concepto de *americanismo*, como puede verse de manera resumida en la Tabla 3, resulta ser muy amplio e incluye las siguientes categorías.

Americanismo
Voces provenientes de las lenguas indígenas (indigenismos) incorporadas al español general o regional, en su forma etimológica o adaptada a la fonología y morfología del español.
Palabras creadas o inventadas en América o derivadas de voces patrimoniales según las normas de la lengua española, para designar referentes americanos (naturaleza, cultura, vida americana, etc.).
Voces españolas que tienen en América acepciones diferentes de las tradicionales peninsulares. La diferencia viene de la aplicación de un nombre para un referente europeo a uno americano parecido o bien resulta de la variación semántica 'normal'.
Arcaísmos, marinerismos, regionalismos españoles desconocidos hoy en España, pero que se mantienen vivos en América con sus acepciones antiguas.
Préstamos como latinismos, helenismos, africanismos, anglicismos, galicismos que forman parte del español americano.

TABLA 3: *Clasificación de americanismo propuesta por Morínigo en el NDAI (1998)*

Como acabamos de ver, reina un cierto desacuerdo en cuanto a las definiciones de los diferentes *-ismos* en los diccionarios presentados, sobre todo con respecto a *americanismos*, *argentinismos* y *uruguayismos*, ya que no nacen de las mismas perspectivas, ideas lingüísticas e ideologías ni tampoco concuerdan en cuanto a la restricción de los conceptos. No obstante, sí ha quedado manifiesto que solo aparecen en los diccionarios de los siglos XIX y XX con esta etiqueta desde un punto de vista diferencial, es decir, comparándose y contrastándose con el léxico peninsular, por lo que en la lexicografía rioplatense misma nunca se hablaría ni de *americanismo* ni de *argentinismo* o *uruguayismo* por constituir la única variedad de objeto y por no haber comparación con otra.

4. ESTUDIO METALEXICOGRÁFICO

A continuación, procederemos a la comparación de las entradas lexicográficas de algunas voces de uso histórico reales, elegidas ejemplarmente de un corpus epistolar rioplatense inédito¹⁶ en los ocho diccionarios seleccionados.

¹⁶ Este constituye el corpus de trabajo de mi tesis doctoral dedicada al estudio del español del siglo XVIII en la antigua Provincia jesuítica del Paraguay (Albers, 2025), la cual corresponde en la diacronía a la región del Río de la Plata y del Paraguay. En el apéndice del mismo trabajo se encuentran las transcripciones paleográficas de los documentos, principalmente cartas, redactadas por autores jesuitas criollos oriundos de la misma provincia de la orden de la Compañía de Jesús.

El corpus del que proviene la selección de las voces de origen patrimonial e indígena es constituido por cien documentos principalmente epistolares del Archivo General de la Nación en Buenos Aires (AGN). Los documentos fueron redactados por jesuitas durante el siglo XVIII en las áreas correspondientes a los actuales países de Argentina, Paraguay y Uruguay.

Los términos que hemos elegido con el criterio de que sirvan como ejemplos representativos para nuestro pequeño estudio metalexicográfico son, por una parte, voces indígenas de difusión panhispánica (*cacique*) o bien supra-regionales o regionales de las lenguas generales del Río de la Plata (*chacra* del quechua y *pacú* del guaraní). Por otra parte, se trata de términos de origen castellano que vivieron ampliaciones semánticas (*estancia*) y/o morfológicas (*caballada*, *potrero*) en algunas partes de América, así como de un préstamo del francés, (*flete*, del fr. *fret*). Todos pertenecen a la vida americana campestre y se refieren, por tanto, a referentes y realidades propiamente americanos.

4.1. CACIQUE

El término caribe *cacique* está registrado solo en el LRP, el NVCR y en el NDAI, y sus acepciones abarcan el significado original del término (1) y la ampliación al área rioplatense propia (2), como puede observarse en las entradas del NDAI y del NVCR:

- (1) Nombre con que los indios araucos de Santo Domingo designaban a sus señores. Los españoles lo difundieron por toda América para designar a los indios señores de vasallos o jefes guerreros. (Morínigo, 1998, s. v. *cacique*)
- (2) Jefe indígena. El cacique de nuestras parcialidades indígenas era solo obedecido en tiempos de guerra. En los días de paz nuestros indios no conocían ninguna forma de jerarquía social ni política. (Guarnieri, 1957, s. v. *cacique*)

El origen indígena de la voz *cacique* es indicado tan solo en dos de los tres diccionarios, esto es, como “voz caribe” (Bermúdez & Bermúdez, 1880-1947, s. v. *cacique*) en el LRP, y como “del taíno de Santo Domingo” (Morínigo, 1998, s. v. *cacique*) en el NDAI.

En cuanto a la difusión del término, el NDAI señala —pese a que tan solo tres de los diccionarios rioplatenses consultados tengan registrado el

término— su expansión por todo el continente americano. Además, es archiconocido el temprano uso de la voz *cacique* al otro lado del océano en el español peninsular —donde en la actualidad suele referirse a una persona que ejerce poder sobre una comunidad de manera abusiva—, por lo que no resulta improbable la hipótesis de que algunos de los préstamos de las lenguas antillanas se difundieran desde España a las otras zonas de América (Quirós García & Ramírez Luengo, 2015, p. 198; Ramírez Luengo, 2007, p. 77). Dadas las diferencias semánticas actuales, el indigenismo *cacique* constituye un americanismo semántico que, a diferencia de la mayoría de los americanismos semánticos, no recibió en América un significado distinto al del español peninsular original, sino al revés, que tomó en el español peninsular actual uno propio, distinto al americano original.

4.2. CHACRA

La totalidad de los diccionarios rioplatenses consultados disponen de la entrada de la voz quechua *chacra*, lo cual puede deberse al hecho de que se trata de un préstamo de una lengua indígena presente en la región. El término designa un ‘establecimiento o una finca rural’ como definen ejemplarmente el LRP (3), el IDA (4) y el NDU (5):

- (3) Extensión de tierra más o menos grande, dedicada a la agricultura, a la explotación en pequeño de ciertas industrias rurales y a la cría de animales de corral en escala reducida. (Bermúdez & Bermúdez, 1880-1947, s. v. *chacra*)
- (4) Casa de campo destinada a siembras, distante de la ciudad. Hay generalmente en ellas una quinta o plantío de árboles frutales y un jardín. (Barcia, 2006, s. v. *chacra*)
- (5) Propiedad rural que consta de una vivienda y una determinada extensión de terreno en la que se cultivan esencialmente productos hortícolas y se crían aves de corral, cerdos, etc. (Kühl de Mones, 1993, s. v. *chacra*)

Cuatro de los diccionarios proporcionan información acerca del origen quechua del término. De este modo, el LRP señala “del quich”. (Bermúdez & Bermúdez, 1880-1947, s. v. *chacra*), el NVCR aclara que “es una voz quichua” (Guarnieri, 1957, s. v. *chacra*) y el DANB informa acerca de su forma original (“del quichua *chajra*” (Segovia, 1911, s. v. *chacra*)), mientras que el NDAI

explica que el término proviene “del quichua chac-ra. Sementera” (Morínigo, 1998, s. v. *chacra*), agregando así el significado quechua original ‘sementera’.

En lo que se refiere a las áreas de uso del término *chacra*, Segovia (1911, s. v. *chacra*) lo sitúa en su DANB en la sección de americanismos, de ahí que quepa suponer que el término es conocido en toda América, pero no en España. A la misma difusión hace referencia el LRP, indicando que posee el mismo significado en casi toda Hispanoamérica (Bermúdez & Bermúdez, 1880-1947, s. v. *chacra*), si bien el área original de uso de la lengua general quechua se restringía a la zona andina de América. Desde un punto de vista sincrónico, sin embargo, el *Diccionario de Americanismos* (DAMER) limita la expansión de este americanismo puro —por no ser conocido en el español peninsular— a las variedades de Sudamérica, ya que no está difundido ni en México y Centroamérica ni en las islas del Caribe (Asociación de Academias de la Lengua Española, 2010, s. v. *chacra*).

4.3. PACÚ

La voz indígena seguramente menos conocida y difundida, *pacú*, proviene del guaraní. Tanto el IDA como el DANB carecen de la entrada del lema. El *pacú* designa un ‘pez de agua dulce’ y posee las características descritas en el LRP (6), el DRR (7) y el NDA (8):

- (6) Pez de agua dulce & de blanca y buena carne aunque inferior en calidad a la del pejerrey o la corbina; de color pardo, cuerpo oval y achatado, escamoso, con cabeza pequeña y de fuertes dientes, perteneciente a la familia de los charanícidos, subgénero de los serralmoninos. Son comunes los ejemplares de más de diez quilos de peso. (Bermúdez & Bermúdez, 1880-1947, s. v. *pacú*)
- (7) Pez grande de los ríos, escamoso, achatado, pardo y de carne exquisita. (Granada, 1998[1890], s. v. *pacú*)
- (8) Pez de agua dulce, de gran tamaño, que puede alcanzar 80cm de longitud y 18kg de peso. Su cuerpo es oval, algo comprimido y de color dorado lustroso, con las aletas de un color que oscila entre anaranjado y rojizo. Vive en los grandes ríos. Se alimenta de frutos, vegetales y crustáceos. (Chuchuy & Hlavacka de Bouzo, 1993, s. v. *pacú*)

Por una parte, tanto el DRR como el NDAI indican el origen guaraní de *pacú* (Granada, 1998[1890], s. v. *pacú*; Morínigo, 1998, s. v. *pacú*), y por otra, el

NVCR lo trata como un referente propio de la región, señalando que se trata de un “Pez de nuestros ríos y arroyos” (Guarnieri, 1957, s. v. *pacú*), o sea, como un referente y término propio y característico de la región del Río de la Plata, por lo que resulta comprensible que el NDAI limite su difusión a Argentina, Bolivia, Paraguay y Uruguay (Morínigo, 1998, s. v. *pacú*). De todos modos, se trata de un caso ejemplar de indigenismo y a la vez americanismo puro, que incluso queda restringido a la zona de influencia de la lengua indígena de la que proviene, el guaraní, y que no vivió una difusión a otras regiones americanas ni mucho menos a España.

4.4. ESTANCIA

Uno de los más conocidos términos de origen patrimonial con la acepción de ‘mansión, habitación, dependencia, alcoba, aposento, cuarto, pieza, etc.’ que vivió un cambio semántico en América es *estancia*, que en amplias partes del nuevo continente tiene el significado de ‘hacienda o establecimiento de campo de grandes dimensiones, destinado a la ganadería o la agricultura’, por lo que cabe clasificarlo como americanismo semántico. Dado que el término es considerado un americanismo de amplia difusión sobre todo en Sudamérica¹⁷, todos los diccionarios rioplatenses consultados lo registran. Tal como señalan el IDA (9), el DRR (10) y el DANB (11), *estancia* puede referirse tanto a la ‘hacienda de campo’ como a la ‘casa perteneciente a la hacienda’:

- (9) 1. Hacienda de campo cuya área de terreno no baja generalmente de media legua de frente por una y media de fondo (construcción que constituye lo que se llama una suerte de estancia) destinada a la cría de ganado, particularmente de vacas, yeguas y ovejas. 2. La casa donde reside el patrón o mayordomo de una hacienda de campo. (Barcia, 2006, s. v. *estancia*)
- (10) 1. Establecimiento de ganadería. 2. Conjunto de edificios y construcciones a él pertenecientes, por lo regular en el punto más eminente del campo [...]. (Granada, 1998[1890], s. v. *estancia*)

¹⁷ Según el DAMER, el término *estancia* con esta acepción propiamente americana está expandido en Estados Unidos, Nicaragua, Panamá, Venezuela, Ecuador, Perú, Bolivia, Chile, Argentina, Paraguay y Uruguay, faltando entonces su uso en México, partes de Centroamérica, las islas del Caribe y Colombia (Asociación de Academias de la Lengua Española, 2010, s. v. *estancia*).

- (11) 1. O establecimiento de campo o simplemente establecimiento, es la hacienda o finca rural destinada a la ganadería y ordinariamente a la cría del ganado vacuno, como fin principal. 2. La población o poblaciones, es decir, el conjunto de edificios y construcciones pertenecientes a dicho establecimiento. (Segovia, 1911, s. v. *estancia*)

El último de los diccionarios citados, el DANB, clasifica *estancia* como castellanismo, de ahí que queden manifiestos su origen peninsular y su expansión tanto en España como en América, si bien con acepciones diferentes (Segovia, 1911, s. v. *estancia*). Además, el NVCR ofrece asimismo información acerca de la historia del término en América, explicándola de la siguiente manera:

- (12) Es una voz antigua, que se aplicó a establecimientos similares de América desde los primeros años de su conquista. (Guarnieri, 1957, s. v. *estancia*)

Con respecto a la difusión del término, el LRP informa acerca de que es “lo mismo en Chile, Río Grande del Sur, y Puerto Rico” (Bermúdez & Bermúdez, 1880-1947, s. v. *estancia*), mientras que el DRR solo lo sitúa en el Río de la Plata. El NDAI, en cambio, enumera “Arg., Chi., Par. y Uru”. (Morínigo 1998, s. v. *estancia*).

4.5. CABALLADA

El término *caballada* —derivado del sustantivo *caballo* y, por tanto, una voz de raíz patrimonial— se encuentra en seis de los diccionarios rioplatenses, faltando solo en ambos diccionarios de argentinismos, el IDA y el NDA.

Tal como coinciden los diccionarios, la voz se emplea para referirse a un ‘conjunto de caballos’ o una ‘manada de caballos’ y puede poseer, según el DRR (13), el NVCR (14) y el DANB (15), además las siguientes características:

- (13) Conjunto de caballos, sea cual fuere su número, con especialidad cuando están destinados a un determinado objeto, como los que se echan por delante de un vehículo para remudar los de tiro, los que pertenecen a un cuerpo de ejército o los de una mensajería. (Granada, 1998[1890], s. v. *caballada*)
- (14) 1. Conjunto de caballos de que puede disponer un ejército, una tropa en marcha, etc. 2. Caballos disponibles en una estancia. (Guarnieri, 1957, s. v. *caballada*)
- (15) Conjunto de caballos, que no forman tropilla, especialmente cuando están reunidos para un determinado objeto. (Segovia, 1911, s. v. *caballada*)

En el DANB, donde el término aparece en los americanismos por ser conocido en América, pero no en España, se indica además la raíz patrimonial

caballo (Segovia, 1911, s. v. *caballada*). Mientras que el LRP adscribe el uso de *caballada* a toda Hispanoamérica, el NDAI lo clasifica como propiamente argentino y uruguayo (Bermúdez & Bermúdez, 1880-1947, s. v. *caballada*; Morínigo, 1998, s. v. *caballada*). Por lo tanto, *caballada* podría clasificarse como americanismo puro¹⁸, si tenemos en cuenta que, en el español peninsular, el término —si bien no es desconocido— se limita al ámbito rural y no está difundido en el habla diaria.

4.6. POTRERO

Al igual que el anterior, el término *potrero* constituye asimismo un derivado de una voz patrimonial, en este caso del sustantivo *potro*, y asimismo un americanismo puro. El lema *potrero* se encuentra en todos los diccionarios analizados, con excepción del IDA.

Como confirman las respectivas entradas en el LRP (16), el DRR (17), el DANB (18), el NVCR (19), el NDU (20) y en el NDAI (21), los significados de *potrero* pueden variar ligeramente:

- (16) 1. Subdivisión mayor o menor que se hace en un campo por medio de alambrados o zanjas anchas y profundas, destinada a la cría y alimentación del ganado. [...] 3. Espacio desprovisto de árboles en el interior de un bosque, pero rico en hierbas. 4. Sitio que, debiendo estar limpio, está cubierto de yuyos. (Bermúdez & Bermúdez, 1880-1947, s. v. *potrero*)
- (17) 1. Terreno cercado, para tener animales a mano, aquerenciar caballos, entropillar, desternear, etc. 2. Campo aparente para un pastoreo especial, por tener los mejores pastos, aguadas, etc. 3. Rinconada de buenos pastos. (Granada, 1998[1890], s. v. *potrero*)
- (18) 1. Dehesa potrill o terreno cercado, destinado para tener animales a mano, aquerenciar caballos, entropillar, etc., Ordinariamente tiene pasto y agua. 2. Rinconada de buenos pastos. 3. Campo aparente para una invernada u otro pastoreo especial, por tener los mejores pastos, aguadas, etc. (Segovia, 1911, s. v. *potrero*)
- (19) 1. Lugar donde se encierran los potros. 2. Porción de campo cercado, donde pastan los ganados. (Guarnieri, 1957, s. v. *potrero*)

¹⁸ La clasificación de *caballada* como americanismo puro se basa en la definición que da Company Company (2010, pp. XVII-XVIII), según la cual tales voces resultan inexistentes en el peninsular *general*. En este caso, *caballada*, suele emplearse en el lenguaje especial rural de España, pero no en el lenguaje general y cotidiano.

- (20) En un establecimiento rural, cada una de las fracciones cercadas en que se divide el campo, destinadas al pastoreo del ganado o a la agricultura. (Kühl de Mones, 1993, s. v. *potrero*)
- (21) 1. Prado o campo destinado a la cría y pastoreo de toda clase de animales. (Morínigo, 1998, s. v. *potrero*)

De este modo, las acepciones de *potrero* pueden abarcar una ‘porción de campo, destinado a la cría y alimentación del ganado’, un ‘terreno donde se encierran los potros’, una ‘subdivisión de un terreno rural perteneciente a una estancia’ o un ‘espacio de buenos pastos y hierbas’, siempre refiriéndose a las realidades rurales americanas.

Al igual que *caballada*, el DANB clasifica *potrero* como americanismo, siendo de este modo conocido en Hispanoamérica, pero no en España, e indica la base de derivación *potro* (Segovia, 1911, s. v. *potrero*). Dado que el NDAI también señala la difusión del término en toda América (Morínigo, 1998, s. v. *potrero*), cabe deducir que se trata de un americanismo común a toda Hispanoamérica que no está difundido en el español peninsular.

4.7. FLETE

Por último, *flete* —término cuyo origen se remonta al francés *fret* (‘precio del transporte de mercancías por aire, por mar, etc.’) y que forma parte del léxico español desde el siglo XV, ya que el CORDE testimonia las primeras apariciones del préstamo en traducciones al español de los años 1417 y 1427— está registrado en todos los diccionarios consultados.

Los significados de *flete* son varios, de ahí que pueda clasificarse como polisémico, como ilustran algunas de las entradas lexicográficas en el IDA (22), el DRR (23), el NVCR (24), el LRP (25) y el NDU (26):

- (22) Caballo ligero e infatigable para galopar, brioso, de buenas cualidades para paseo. (Barcia, 2006, s. v. *flete*)
- (23) Caballo bueno, ligero. (Granada, 1998[1890], s. v. *flete*)
- (24) Caballo hermoso que constituye el lujo del paisano. (Guarnieri, 1957, s. v. *flete*)
- (25) 1. Precio del alquiler del animal o vehículo empleado para el transporte de mercaderías, sean estas conducidas por mar o por tierra. 2. Carga que se transporta por mar o por tierra. 3. Valor del transporte. 4. Caballo de buena estampa, ligero e infatigable para el camino. (Bermúdez & Bermúdez, 1880-1947, s. v. *flete*)

- (26) 1. Transporte de cargas. 2. Carga que se transporta por tierra, por mar o por aire. 3. Precio que se paga por el transporte de objetos o animales en cualquier medio de transporte. 4. Vehículo que se contrata para transporta muebles u objetos, especialmente en ocasión de una mudanza. 5. Caballo de excelentes condiciones y calidad, especialmente el que se destina a paseo. (Kühl de Mones, 1993, s. v. *flete*)

Mientras que las entradas en los ejemplos (22), (23) y (24) indican como único significado el de ‘caballo de montar de buena calidad’, los demás diccionarios rioplatenses se deslizan semánticamente de este significado restringido y enumeran una serie más amplia que va desde ‘carga que se transporta’ o ‘vehículo que se contrata para el transporte’ hasta ‘valor del transporte’ o ‘precio del alquiler del animal o vehículo del transporte’, es decir, que van más allá de la calidad del caballo en sí, ya que este significado pareciera limitarse al Río de la Plata.

El DANB ofrece dos clasificaciones distintas de *flete*. Por una parte, Segovia (1911, s. v. *flete*) cataloga *flete* como argentinismo cuando significa “precio del transporte de mercaderías, etc., por agua o por tierra”, y por otra parte, como castellanismo como sinónimo de “*pingo*, en lenguaje *gaucho*”, siendo este gauchismo o ruralismo registrado como “caballo vivo, ligero, de buenas condiciones. Especie de caballo de regalo o bridón” (Segovia, 1911, s. v. *pingo*).

En lo que respecta a la difusión de *flete* y de sus acepciones, cabe distinguir tres aspectos. Por un lado, los significados referentes al transporte, su carga y precio, que se deslizan del significado propiamente rioplatense del ‘caballo de buena calidad’, resultan estar ampliamente difundidos en América por ser conocidos, además de en el Río de la Plata, también en “Colombia, Chile, Guatemala, Perú, Venezuela y P. Rico” (Bermúdez & Bermúdez, 1880-1947, s. v. *flete*). En el español peninsular, en cambio, el término se restringe casi exclusivamente al lenguaje técnico en el ámbito del transporte de carga (Chávez Fajardo, 2022, p. 48). Por otro lado, el significado más específico, propiamente rioplatense y, según el DANB, gauchesco, que constituye la única acepción mencionada en el IDA, el DRR el NVCR, es usado exclusivamente

en Argentina y Uruguay, así como en Bolivia y Colombia, según en LRP y el NDAI (Bermúdez & Bermúdez, 1880-1947, s. v. *flete*; Morínigo, 1998, s. v. *flete*). No obstante, tanto el NDA (27) como el NDU (28) señalan que esta acepción de *flete* suele usarse principalmente en el habla coloquial:

- (27) Coloq. Caballo rápido. (Kühl de Mones, 1993, s. v. *flete*)
- (28) Coloq. Caballo de excelentes condiciones y calidad, especialmente el que se destina a paseo. (Chuchuy & Hlavacka de Bouzo, 1993, s. v. *flete*)

Por lo tanto, cabe deducir que el significado restringido ‘caballo de montar de buena calidad’ resultara ser propio de la región del Río de la Plata —además de Bolivia y de Colombia—, pero limitado al habla coloquial, rural y gauchesca. En cambio, los demás significados relacionados con el transporte cuentan con una expansión más amplia en gran parte de América, mientras que, en España, *flete* se limita al ámbito propiamente técnico, de ahí que pueda clasificarse como americanismo semántico o incluso como americanismo puro¹⁹, dependiendo de las diferentes acepciones del término.

5. CONCLUSIONES

Para terminar, cabe repasar brevemente las observaciones y deducir las consecuencias en cuanto a la lexicografía rioplatense en el contexto americano, así como para el quehacer metalexicográfico panhispánico en general.

El comienzo de la lexicografía rioplatense en el siglo XIX se centró en la integración tanto de *indigenismos* provenientes de las lenguas autóctonas de la región, el guaraní, el quechua y el mapudungun, como de *ruralismos*, simbolizando la revaloración del legado indígena y de la vida campestre como partes de la identidad rioplatense en el léxico (Lauria, 2010, p. 187). Estas voces indígenas y referentes a la vida rural fueron presentados como

¹⁹ Véase la nota 18. En este caso, *flete*, suele emplearse con el significado ‘precio estipulado por el alquiler de una nave o de una parte de ella’ en el lenguaje especial técnico de España, pero no en el lenguaje general, lo cual permite su clasificación como americanismo puro, mientras que cabría clasificarlo como americanismo semántico en cuanto a la acepción de ‘caballo de buena calidad’.

propios de la región del Río de la Plata. Los respectivos diccionarios que se confeccionaron en Argentina y Uruguay solían ser diccionarios con criterios de uso y diferenciales, ya que describían el lenguaje cotidiano usado por los habitantes de ambas naciones en contraste con los empleos peninsulares.

En el siglo XX se inició un cambio de rumbo con la confección de diccionarios enfocados en los *americanismos* —tanto puros y semánticos como de frecuencia (Company Company, 2007; 2010; Ramírez Luengo, 2014)— con un panorama más amplio. No obstante, las delimitaciones de los conceptos como *americanismo*, *argentinismo*, etc. seguían siendo heterogéneas en la centuria pasada.

Como hemos podido confirmar en nuestro pequeño estudio cualitativo de naturaleza metalexicográfica, los siete *-ismos* analizados a base de un corpus epistolar dieciochesco en las obras lexicográficas rioplatenses de los siglos XIX y XX constituyen términos conocidos y usados o bien en amplias partes de América o bien principalmente en la región del Río de la Plata, siempre denominando realidades y referentes propios del continente americano o de la propia región. Por lo tanto, pueden clasificarse como americanismos semánticos (*cacique*, *estancia*, *flete*), a veces con una limitación a la región del Río de la Plata o a zonas más amplias de Sudamérica, y como americanismos puros (*chacra*, *pacú*, *caballada*, *potrero*, *flete*) cuando la voz resulta completamente desconocida en el español peninsular general o se restringe a un lenguaje especial. En el futuro, quedaría por ampliar el estudio, incorporando un número mayor de lexemas de origen indígena o patrimonial con el fin de desvelar los mecanismos detrás del uso de los diferentes *-ismos* en el léxico americano y rioplatense de la época, así como para llevar a cabo un análisis cuantitativo.

No obstante, no reina un acuerdo ni en cuanto a la clasificación ni en lo que se refiere a la indicación de la difusión de los respectivos *-ismos*, ya que no todos los diccionarios cuentan con la información acerca de la expansión y, en el caso de que sí la contengan, muchas veces no coinciden exactamente. En cuanto a las voces de origen indígena, tan solo algunos de

los diccionarios indican su procedencia, mientras que otros los tratan como los demás términos de origen patrimonial. Tanto el NDA como el NDU, por ejemplo, renuncian por completo a la mención de la procedencia autóctona de los tres indigenismos estudiados, *cacique*, *chacra* y *pacú*. Con respecto a las acepciones de cada uno de los términos estudiados, en cambio, sí puede afirmarse un acuerdo en los diccionarios rioplatenses, aunque los significados pueden variar ligeramente o enfocarse en aspectos distintos.

Las tres obras lexicográficas rioplatenses decimonónicas, sobre todo el LRP y el DRR, han resultado ser diccionarios de regionalismos que se centran en los términos propios del Río de la Plata ‘desde adentro’, es decir, elaborados desde la perspectiva regional de la zona geográfica estudiada. Los diccionarios del siglo XX, en cambio, abordan el léxico desde una perspectiva americana más amplia, con excepción del NVCR, que puede clasificarse como obra expresamente regionalista y ruralista que en muchas ocasiones recurre al posesivo *nuestro* para destacar los aspectos propios e identitarios rioplatenses, por ejemplo, en “pez de nuestros ríos y arroyos” (Guarnieri, 1957, s. v. *pacú*) o en “el cacique de nuestras parcialidades” (Guarnieri, 1957, s. v. *cacique*). De este modo, también el legado indígena es tratado como integrante cultural e identitario de la variedad rioplatense. Por una parte, el NDAI se caracteriza por siempre indicar la expansión de cada lema, dado que la obra abarca toda América y no solo el español regional rioplatense. El DANB, por otra parte, realiza un aporte innovador por su clasificación de los términos en *castellanismos*, *americanismos* y *argentinismos* en *sensu strictu*.

Los *argentinismos* y *uruguayismos* fueron presentados como conceptos diferenciales con vistas a la comparación y delimitación con otras variedades del español, es decir, que se usaba solo “hacia afuera” (Lauria, 2010, p. 183), o sea desde la perspectiva nacional propia hacia otras variedades.

En una perspectiva futura más amplia de la lexicografía histórica hispanoamericana, resultaría de gran interés abordar el tratamiento de los diferentes *-ismos* en la lexicografía de los siglos XIX y XX de diversas regiones

americanas con el propósito de comparar y contrastar tanto los diferentes conceptos, definiciones y tradiciones como también las ideologías y los componentes identitarios de cada una de las naciones de América a través de la metalexicografía. Este ángulo comparativo panamericano, así como la combinación de las perspectivas diacrónicas y sincrónicas hacia los diccionarios, permitirían en un futuro, además, un acercamiento a la lexicografía panhispánica actual y a los modelos pluricéntricos de la lengua española.

REFERENCIAS BIBLIOGRÁFICAS

AGN = Archivo General de la Nación. Fondos AR-AGN.DE/CJ, Sala 9.

Albers, M. (2025). *El español en la antigua Provincia del Paraguay. Comunicación epistolar jesuítica en el siglo XVIII* [Tesis de Doctorado, Universidad de Salzburgo]. Digitale Bestände der Universität Salzburg. <https://eplus.uni-salzburg.at/Abschlussarbeiten/content/titleinfo/13265361>

Bertolotti, V. (2015). *A mí de vos no me trata ni usted ni nadie. Sistemas e historia de las formas de tratamiento en la lengua española en América*. Universidad Nacional Autónoma de México.

Billaudelle, P. (1999). *Spanischsprachige Lexikographen im Spiegel der Prologe ihrer Werke (1780-1925)*. Kovač.

Bravo-García, E. & Cáceres-Lorenzo, M. T. (2011). *La incorporación del indigenismo léxico en los contextos comunicativos canario y americano (1492-1550)*. Peter Lang. <https://doi.org/10.3726/978-3-0352-0077-5>

Canale, G. & Coll, M. (2016). Historia y presente del yeísmo (rehilado) en el Uruguay. *Lexis*, 40(1), 5-40. <https://doi.org/10.18800/lexis.201601.001>

Carricaburo, N. (2015). *Las fórmulas de tratamiento en el español actual*. Arco Libros.

Chávez Fajardo, S. (2022). *Diccionarios del fin del mundo*. Fondo de Cultura Económica.

Coll, M. (2024). Los diccionarios del español de América. En S. Torner, P. Battaner & I. Renau (eds.), *Lexicografía hispánica. The Routledge Handbook of Spanish Lexicography* (pp. 481-493). Routledge. <https://doi.org/10.4324/9780429244353-36>

Company Company, C. (2007). *El siglo XVIII y la identidad lingüística de México. Discurso de ingreso a la Academia Mexicana de la Lengua*. Universidad Nacional Autónoma de México & Academia Mexicana de la Lengua.

Company Company, C. (2010). Introducción. En Academia Mexicana de la Lengua (ed.), *Diccionario de Mexicanismos* (pp. XV-XXIII). Siglo XXI.

- CORDE = Real Academia Española. *Corpus diacrónico del español*. <https://corpus.rae.es/cordenet.html>
- DAMER = Asociación de Academias de la Lengua Española (2010). *Diccionario de americanismos*. <https://www.asale.org/damer/>
- Da Rosa, J. J. & Lucián, E. (2016). Lenguaje del Río de La Plata: rescate de un diccionario dialectal desconocido. *Thesaurus*, 58, 128-154.
- DANB = Segovia, L. (1911). *Diccionario de argentinismos, neologismos y barbarismos: con un apéndice sobre voces extranjeras interesantes*. Coni.
- DRR = Granada, D. (1998[1890]). *Diccionario rioplatense razonado* (Ed. Ú. Kühl de Mones). Arco Libros.
- Enguita Utrilla, J. M. (2010). Indoamericanismos léxicos y estructuras discursivas en la *Relación* de Cristóbal de Molina. En P. Jiménez del Campo, P. Cuenca Muñoz & E. López Parada (eds.), *Relación de las fábulas y ritos de los incas* (pp. 199-216). Iberoamericana Vervuert. <https://doi.org/10.31819/9783954871766-010>
- Fontanella de Weinberg, M. B. (1992). *El Español de América*. Mapfre.
- Frago Gracia, J. A. (2011). El español de América en la Independencia. Adiciones gramaticales I. *Boletín de Filología*, XLVI(1), 47-74. <https://doi.org/10.4067/S0718-93032011000100002>
- Haensch, G. & Werner, R. (1993a). Introducción. En C. Chuchuy & L. Hlavacka de Bouzo (eds.), *Nuevo Diccionario de Americanismos. Tomo II: Nuevo Diccionario de Argentinismos* (Dir. G. Haensch & R. Werner) (pp. XIX-LXVII). Instituto Caro y Cuervo.
- Haensch, G. & Werner, R. (1993b). Introducción. En Ú. Kühl de Mones (ed.), *Nuevo Diccionario de Americanismos. Tomo III: Nuevo Diccionario de Uruguayismos* (Dir. G. Haensch & R. Werner) (pp. XVII-LVI). Instituto Caro y Cuervo.
- Huisa Téllez, J. C. (2024). Los diccionarios contrastivos de las variantes del español de América. En S. Torner, P. Battaner & I. Renau (eds.), *Lexicografía hispánica. The Routledge Handbook of Spanish Lexicography* (pp. 469-480). Routledge. <https://doi.org/10.4324/9780429244353-35>
- IDA = Barcia, P. L. (2006). *Un inédito Diccionario de argentinismos del siglo XIX*. Academia Argentina de Letras.
- Kühl de Mones, Ú. (1998). Introducción. En D. Granada, *Diccionario rioplatense razonado* (Ed. Ú. Kühl de Mones) (pp. V-XXXI). Arco Libros.
- Lapesa, R. (2008[1981]). *Historia de la lengua española*. Gredos.
- Lauria, D. (2010). Tratamiento de indigenismos en el *Vocabulario rioplatense razonado* de Daniel Granada (1889). *Ianua. Revista Philologica Romanica*, 10, 175-202.
- Lauria, D. (2011). Apuntes para una historia de la producción lexicográfica monolingüe en la Argentina: etapas del proceso de diccionarización y modalidades diccionarísticas entre 1870 y 1910. *Boletín de Filología*, XLVI(1), 105-151. <https://doi.org/10.4067/S0718-93032011000100004>

- LRP = Bermúdez, W. P. & Bermúdez, S. W. (1880-1947). *Lenguaje del Río de la Plata*, folios manuscritos o mecanografiados. Academia Nacional de Letras.
- NDA = Chuchuy, C. & Hlavacka de Bouzo, L. (1993). *Nuevo Diccionario de Americanismos. Tomo II: Nuevo Diccionario de Argentinismos* (Dir. G. Haensch & R. Werner). Instituto Caro y Cuervo.
- NDAI = Morínigo, M. A. (1998). *Nuevo Diccionario de Americanismos e Indigenismos*. Claridad.
- NDU = Kühl de Mones, Ú. (1993). *Nuevo Diccionario de Americanismos. Tomo III: Nuevo Diccionario de Uruguayismos* (Dir. G. Haensch & R. Werner). Instituto Caro y Cuervo.
- NVCR = Guarnieri, J. C. (1957). *Nuevo vocabulario campesino rioplatense*. Editorial Florensa.
- Pöll, B. (2018). *Spanische Lexikologie*. Narr Francke Attempto. <https://doi.org/10.24053/9783823391586>
- Ramírez Luengo, J. L. (2007). *Breve historia del español de América*. Arco Libros.
- Ramírez Luengo, J. L. (2014). Cómo el español de España genera americanismos: a propósito del americanismo *puto* 'homosexual'. *Lengua y habla*, 18, 1-12. <https://doi.org/10.22517/25393812.12271>
- Ramírez Luengo, J. L. & Quirós García, M. (2015). Observaciones sobre el léxico del español de Yucatán (1650-1800). *Revista de Filología Española*, XCV(1), 183-210. <https://doi.org/10.3989/rfe.2015.08>
- Sánchez Méndez, J. P. (2016). Las concepciones lingüísticas de la Ilustración hispanoamericana. En M. Guzmán Riverón & D. M. Sáez Rivera (eds.), *Márgenes y centros en el español del siglo XVIII* (pp. 21-42). Tirant Humanidades.
- Zimmermann, K. (2016). Missionarslinguistik in kolonialen Kontexten. Ein historischer Überblick. En T. Stolz, I. H. Warnke & D. Schmidt-Brücken (eds.), *Sprache und Kolonialismus. Eine interdisziplinäre Einführung zu Sprache und Kommunikation in kolonialen Kontexten* (pp. 169-192). De Gruyter. <https://doi.org/10.1515/9783110370904-008>



DE LA “A” A LA “AZ”. ESTUDIO DE LAS ANOMALÍAS FORMALES DEL TEXTO EN ESPAÑOL DEL *CAMBRIDGE DICTIONARY ONLINE* (ENGLISH-SPANISH/SPANISH-ENGLISH)

FROM “A” TO “AZ”. A STUDY OF FORMAL ANOMALIES OF THE SPANISH TEXT IN *CAMBRIDGE DICTIONARY ONLINE* (ENGLISH-SPANISH/SPANISH-ENGLISH)

Santiago Rodríguez-Rubio
Universidad Pablo de Olavide
srodmed@upo.es

RESUMEN

En el presente trabajo, se analiza el texto en español de las entradas correspondientes a la letra “A” de las partes English-Spanish y Spanish-English del [Cambridge Dictionary Online](https://www.cambridge.org/es/dictionary), en función de los siguientes tipos de anomalías formales: mayúscula, puntuación, diseño, acento ortográfico, espacio, *errata de no palabra* (ENP) y *errata de palabra* (EP). Los resultados en términos de incidencia y ubicación de los tipos/subtipos de anomalías demuestran el carácter sistemático de los defectos y reclaman una acción correctora acorde, bajo la premisa de una revisión integral. Más allá del análisis de las anomalías, se abordan aspectos relacionados con la calidad lexicográfica como: a) la posible relación entre los defectos y los métodos (manual, automatizado o semiautomatizado) de introducción de datos; b) el difícil acomodo entre la productividad y la calidad en un contexto de creciente complejidad en que la tecnología o el uso que se hace de ella sigue provocando ajustes; c) la colaboración de los usuarios en la introducción y revisión del contenido lexicográfico; y d) el efecto perjudicial que las imágenes publicitarias pueden tener en los usuarios de los diccionarios en línea gratuitos.

Palabras clave: anomalías formales, calidad lexicográfica, lexicografía en línea gratuita, revisión de diccionarios.

ABSTRACT

The present study analyses the Spanish text of entries corresponding to letter “A” in parts English-Spanish and Spanish-English of the [Cambridge Dictionary Online](https://www.cambridge.org/es/dictionary), according to the following formal anomaly types: capitalisation, punctuation, design, written accent, spacing, non-word error (*errata de no palabra*/ENP) and real-word error (*errata de palabra*/EP). The results in terms of number of anomalies and location of the anomaly types/subtypes reveal a system of anomalies and ask for appropriate corrective measures, on the premise of a thorough revision. In addition to anomaly analysis, the following lexicographic quality aspects are dealt with: a) the possible relationship between the defects and the data input methods (manual, automated or semi-automated); b) the difficult balance between productivity and quality, in a context of increasing complexity where there are still gaps in relation to technology (or the use humans make of that technology); c) user participation in the input and revision of lexicographic content; and d) the detrimental effect that promotional images in free online dictionaries may have on users.

Keywords: formal anomalies, dictionary quality, free online lexicography, dictionary revision.

Recibido: 20-04-2026
Aceptado: 19-06-2026

DOI: <https://doi.org/10.17561/rilex.9.2.10412>



1. INTRODUCCIÓN

En el presente estudio, se han analizado las anomalías formales del texto en español de las entradas correspondientes a la letra “A” de las partes English-Spanish y Spanish-English del [Cambridge Dictionary Online](#). En una lógica integral que desborda ampliamente la de las colaboraciones puntuales de usuarios, presentamos un necesario análisis de defectos que traslada la importancia de llevar a la práctica una de las ventajas teóricas de los diccionarios en línea: una más “rápida y fácil” revisión lexicográfica.

Las anomalías son fenómenos ampliamente observados que afectan a todo tipo de textos. Los errores tipográficos (una de las formas más comunes de ruido natural en los textos escritos creados por usuarios, normalmente debido a un tecleo erróneo accidental) pueden aparecer en modelos lingüísticos de gran tamaño (Liu, Zhao, Altinger, Schütze & Hedderich, 2025). Debido a la natural habilidad de las erratas para abrirse camino hacia el texto publicado (Wheatley, 1893), no solo los usuarios, sino también desarrolladores o editores pueden generar erratas cuando introducen datos en sus recursos. Además, la errata no solo puede introducirse a través de un error de actuación humano, sino también automáticamente por acción del corrector ortográfico. De hecho, Sharma y Gupta (2015, pp. 99-100) manifiestan que los avances tecnológicos (informáticos) han supuesto una escalada de *spelling errors*, afirmación que, salvando las distancias, concuerda con la lógica de que “as datasets become larger and more complex, it can be challenging to develop anomaly detection methods that can scale effectively to these larger datasets” (Gujral, 2023, p. 7).

Entre los problemas más acuciantes, se encuentra la asunción errónea de que los textos que se introducen (por ejemplo, aquellos que forman parte de modelos lingüísticos) están limpios, dejando el terreno de la robustez desatendido en cuanto al aspecto de las erratas (Liu, Zhao, Altinger, Schütze & Hedderich, 2025). Esa asunción contrasta en cierto modo con la noción de “Dirty Web” como fuente lexicográfica, es decir, el hecho de que el

contenido de Internet puede contener numerosos errores lingüísticos, a pesar de que el propio volumen de información funcione a menudo como filtro de corrección; por ejemplo, la forma *believe* aparece en 3,8 millones de búsquedas, frente a las 41 600 ocurrencias de *beleive* y las 1700 de *beleave*, lo que permite identificar fácilmente la forma correcta (Svensén, 2009, pp. 59-60).

2. CALIDAD DE LOS DICCIONARIOS ELECTRÓNICOS

2.1. INTRODUCCIÓN Y REVISIÓN DE CONTENIDO EN DICCIONARIOS ELECTRÓNICOS

Destacados especialistas han manifestado de manera constante en el tiempo que la tarea de revisión de un diccionario es, en realidad, inacabable (Landau, 2001; Klosa-Kückelhaus & Tiberius, 2024). Herold, Meyer y Wiegand explican que el marco de revisión de los diccionarios en línea suele situarse entre el de los diccionarios en papel (cuya reedición podía tardar años o incluso décadas) y el de la publicación inmediata de actualizaciones individuales (2024, p. 22). No obstante, Landau (2001, pp. 399-400) cuestionó que los diccionarios electrónicos se revisaran de manera continua y advirtió sobre el deterioro de la fase de revisión (*proofreading*) que la espiral y acumulación de tareas propiciada por los propios avances tecnológicos puede ocasionar. Jackson (2018, p. 836) significativamente usa el adverbio *potentially* cuando alude a la superior actualización de los diccionarios online respecto de los realizados en papel, a la par que Lew y De Schryver incluyen (como no podía ser de otra forma) el modal de posibilidad en su afirmación de que “Online dictionaries as well as dictionary apps can be updated as often as needed” (2014, p. 345). Sea como fuere, Herold, Meyer y Wiegand declaran: “In online reference works that are compiled with a limited set of editors, a dictionary entry will only be published after rigorous checking and approval” (2024, p. 22).

Svensén (2009, p. 421) distingue diversos *modus operandi* en cuanto a la introducción y revisión del contenido lexicográfico. En primer lugar, el uso de un procesador (especialmente cuando se acompaña de anotación

tipográfica) limita la búsqueda y la revisión automática de la coherencia y de la conformidad con los principios editoriales establecidos. En segundo lugar, el procedimiento basado en bases de datos léxicas tiene la virtud de que se pueden hacer búsquedas en función del tipo de campo, lo que facilita que el usuario pueda encontrar conexiones basadas en la ausencia de linealidad. Finalmente, se puede usar un lenguaje de marcado (por ejemplo, Extensible Markup Language, XML) que permite un acceso más rápido a los datos almacenados¹. En entradas complejas, la compilación y edición se desgrana en múltiples campos, entre los que se encuentran aquellos relativos al idioma de referencia; en diccionarios basados en lenguaje de marcado, puede usarse una sintaxis determinada para definir el formato en términos de negrita y cursiva, entre otras características (Abel & Meyer, 2024, pp. 233-234). Otros autores incluyen el tipo y tamaño de letra, la negrita, la cursiva, la mayúscula y signos de puntuación como las comillas dentro de la categoría *formal properties* (Koeva & Blagoeva, 2013, p. 59). Por su parte, Bloodgood y Strauss (2016, p. 79, pp. 81-82) expusieron seis métodos para marcar anomalías en diccionarios electrónicos basados en XML, entre los que se encuentran sistemas basados en expectativas sobre contenido de un determinado tipo de campo (*single-field*) y sobre contenido de múltiples campos interrelacionados (*tied-field*). El primer tipo de sistema descubrió la presencia de una barra (/) cuyo valor disyuntivo no encajaba en un campo que no admitía opciones. Se pusieron como ejemplos de campos interrelacionados la ortografía y la transcripción fonética, o la frase ilustrativa y su traducción. De manera análoga,

¹ Basándose en la *Quality Assurance Activity Statement* (World Wide Web Consortium, W3C), Rosenblum y Golfman (2004) manifestaron que el aseguramiento de la calidad es un aspecto fundamental del lenguaje XML y puede usarse para detectar automáticamente si se cumplen determinados requisitos (por ejemplo, que el nombre de una revista aparezca en cursiva o que contenga palabras que no incluyan letras iniciales mayúsculas). Los autores señalan un aspecto fundamental al referir que muchas organizaciones despliegan un aseguramiento de la calidad limitado o no lo despliegan en absoluto, ya que “quality assurance incurs extra cost, and many organizations are unwilling or unable to make that investment” (Rosenblum & Golfman, 2004, p. 5).

Rodrigues, Zajic, Doermann, Bloodgood y Ye (2011) ya habían explorado métodos que permitieran a los lexicógrafos detectar anomalías intraentrada e interentrada bajo el prisma de la regularidad estructural, mediante el uso de medidas de perplejidad. Se mencionó, por ejemplo, el hecho de que en los diccionarios bilingües el texto de los idiomas de referencia estuviera mezclado. Los autores emplearon un *Dictionary Manipulation Language* (DML, en Zajic, Maxwell, Doermann, Rodrigues & Bloodgood, 2011, p. 297), una de cuyas ventajas principales era la posibilidad de marcar los errores en el diccionario original.

En todo caso, según Svensén (2009, pp. 422-423), los programas de edición lexicográfica que se usen deben (entre otras funciones): alertar sobre redundancias de lemas; permitir la función copia-pegar de texto; promover que se respete el orden estructural de manera que los lemas aparezcan antes que las definiciones; posibilitar el uso de plantillas de entrada; y facilitar la consulta de un registro de modificaciones (especialmente útil en la fase editorial del diccionario).

No existe un sistema de edición lexicográfica perfecto. Así, según Simões y Salgado (2022, p. 425), el modelo TEI Lex-0 (versión avanzada de la *Text Encoding Initiative*) refuerza la interoperabilidad y permite un más fácil manejo computacional, pero puede ser más propenso al error cuando se realiza una edición manual de los recursos. Los autores proponen métodos de simplificación de la codificación de fenómenos (por ejemplo, de la sinonimia), al objeto de que el lexicógrafo pueda centrarse en su labor de análisis lingüístico. El modelo TEI Lex-0 facilita la inclusión de ejemplos tanto bibliográficos (entrecomillados) como inventados por el propio lexicógrafo (Simões & Salgado, 2022, p. 426, p. 428). Es de suponer que el lexicógrafo introduce los ejemplos de uso inventados de manera manual, o, al menos de manera híbrida manual-automática.

El lexicógrafo tiene a su disposición programas de gestión lexicográfica (*Dictionary Writing Systems*, DWS) que pueden incrementar la productividad

gracias al uso de plantillas y de la función de copia, además de reducir el riesgo de errores aleatorios (Koeva & Blagoeva, 2013, p. 65). Dichos programas pueden usarse en tareas de revisión o edición (Abel, 2012; Barbierik, Děngeová, Holcová, Jarý, Liška, Lišková & Virius, 2014), pero más recientemente se ha mostrado que, según una encuesta sobre prácticas lexicográficas realizada a nivel europeo, los DWS eran útiles para comprobar la cohesión, pero muchos de ellos no incluían un corrector ortográfico (*spellchecker*) (Kallas, Koeva, Kosem, Langemets & Tiberius, 2019). Algo más recientemente, frente a los problemas de calidad de los grandes conjuntos de datos de corpus, los creadores de la aplicación “Lexical Data API” basada en los *K Dictionaries* (KD) han defendido una combinación de revisión automática y manual de sus recursos, privilegiando la calidad sobre la cantidad (Lonke, Kernerman & Dzhuranyuk, 2022, p. 405).

Por otro lado, las tareas de revisión requieren de una elevada sistematicidad que tenga en cuenta conexiones y la multiplicidad de espacios en los que la misma información pueda encontrarse. Ese condicionante aparece en la literatura de manera constante en el tiempo: desde el estándar denominado *Multi-Dictionary Formatter* o MDF, en cuyo marco los errores deben corregirse no solo en la versión imprimible del diccionario, sino también en la base de datos léxica original (Coward & Grimes, 2000, p. 63), hasta las bases de datos lexicográficas centralizadas, que integran una “constelación” de recursos, con la consecuente complejidad en la gestión de los datos (Klosa-Kückelhaus & Tiberius, 2024, p. 90).

En relación con el uso de corpus, Geyken y Lemnitzer (2024, pp. 219-220) plantean que una de las desventajas de la integración de fuentes primarias en los diccionarios digitales es precisamente la presencia en ellas de errores, equivocaciones (*mistakes*) e incoherencias, a las que se suman “errores” que surgen “during the process of automatically annotating and analysing the data since no language technology tool can operate without making mistakes”. Los autores indican la pertinencia de lo anterior en términos de fiabilidad

(ya que los usuarios esperan encontrar información conforme a las normas lingüísticas) y apuntan que una de las discrepancias puede deberse a la diferencia entre formas extraídas de la fuente primaria y que se desvían de la norma ortográfica (*raw data*) y, por otro lado, las que han de consignarse en el diccionario con arreglo a dicha norma (*processed data*).

2.2. COLABORACIÓN DE LOS USUARIOS DE DICCIONARIOS Y APUNTES EMPRESARIALES

Podemos partir de la idea básica de que la interacción de los usuarios y lexicógrafos es fundamental para crear diccionarios fáciles de usar. Abel y Meyer amplían el foco, expresando que el grado de participación de los usuarios de diccionarios puede abarcar desde compilaciones completas hasta entradas individuales, incluyendo la corrección de errores, en un marco en que “user participation has become an increasingly important factor in the planning, development, and use of a dictionary” y en el que incluso las editoriales lexicográficas pueden ahorrar dinero (Abel & Meyer, 2024, pp. 228-230).

Atendiendo a lo expresado por Klosa-Kückelhaus y Tiberius (2024, pp. 71-73), la interacción de los usuarios con los editores (por ejemplo, reportando errores ortográficos) no solo está asentada en la práctica lexicográfica, sino que puede suponer un impulso para el desarrollo y la mejora de los diccionarios electrónicos. En función de los proyectos, editores jefes, directores de proyecto y lexicógrafos intervienen en varios estadios del proceso de preparación para corregir, entre muchos otros aspectos, errores ortográficos e inconsistencias. En algunos casos, el lenguaje de marcado o *mark-up language* “was subsequently refined in order to identify other units of information in the microstructure of the dictionary text that could be derived from typographical features (examples, for instance, are always coded in italics)” (Klosa-Kückelhaus & Tiberius, 2024, pp. 79-81, p. 83).

En relación con el tipo de diccionario objeto del presente estudio (es decir, un diccionario electrónico institucional respaldado por una editorial asentada), Abel y Meyer (2024) mencionan que no conocen ningún

proyecto activo de participación “directa” de los usuarios y afirman: “this indicates that this type of direct user participation was not successful. This can be due to the economic situation of the institution, the costs of running such a service, the quality of the submissions, or the amount of work that would be needed to make professional use of these contributions” (Abel y Meyer, 2024, p. 237). La participación “indirecta” implica sugerencias, solicitudes y corrección de errores. Las contribuciones conllevan ventajas e inconvenientes. Por un lado, pueden suponer un considerable trabajo para el editor, ya que, por ejemplo, no solo ha de verificar y categorizar la contribución, sino también comprobar si puede localizar los errores a partir de las indicaciones que los usuarios facilitan (Abel y Meyer, 2024, p. 238, p. 241). Por otro lado, desde una perspectiva más empresarial, las aportaciones pueden conllevar una mejora de la calidad de las obras o productos, una oportunidad de publicidad y de atraer a los usuarios hacia su marca y un valor añadido en forma de ahorro de tiempo y dinero (Abel y Meyer, 2024, p. 248, p. 250). No obstante, conviene recordar que el mantenimiento o la actualización a menudo son desatendidos en el campo lexicográfico (Lorentzen & Trap-Jensen, 2016, p. 138).

2.3. TIPOS DE ANOMALÍAS

El concepto de anomalía no ha sido delimitado y suele variar en función del campo de estudio (Foorthuis, 2021, p. 297). En el campo de las ciencias de la informática, Gujral lo define como “rare or unusual events, observations, or patterns that do not conform to the expected behavior or norm within a dataset” (2023, pp. 1-3). El autor señala que las anomalías pueden ser causadas por errores humanos y que su detección puede servir para hacer un seguimiento de los defectos en los procesos de fabricación. Adams y Maxwell establecieron la siguiente distinción: “an anomaly is an unusual structure that has been noticed [...] an error is a structure which a human expert has determined is incorrect. Errors range from typos to information which has been mis-tagged...” (2016, p. 23).

2.3.1. Erratas

Partimos de la distinción general entre *non-word error* (forma errónea que no está en un diccionario) y *real-word error* (forma errónea que adopta el aspecto de una palabra). La dificultad que los correctores ortográficos (*spell-checkers*) y aplicaciones informáticas tienen a la hora de detectar y corregir *real-word errors* (también denominados *grammatical errors*) ha sido referida de manera constante y sostenida en el tiempo (Peterson, 1986, pp. 633-634; Gardner, 1992, p. 51; Fossati & Di Eugenio, 2008, p. 896; Samanta & Chaudhuri, 2013, p. 211; Dashti, Bardsiri & Bardsiri, 2018, p. 488). Dando una vuelta de tuerca, estos últimos autores puntualizan que los propios correctores ortográficos pueden generar *real-word errors* cuando intentan corregir *non-word errors* cometidos por los humanos (p. 488).

Por otro lado, son cuatro las subcategorías de erratas ampliamente adoptadas en el estudio del modelaje del ruido y la evaluación de la robustez de los sistemas: sustitución, adición, supresión y transposición (Liu, Zhao, Altinger, Schütze & Hedderich, 2025). Como se expuso en otro lugar, estas subcategorías fueron sentadas por estudios de los campos de la sicolingüística y del procesamiento del lenguaje natural y, según Blecula (1983), se remontan a la antigüedad griega (Rodríguez-Rubio & Fernández-Quesada, 2021). Dado su carácter universal, son también las cuatro categorías principales que se tienen en cuenta en el campo de las bases de datos en línea (Gardner, 1992, p. 50)².

A pesar de ser universales, ciertos fenómenos motores responsables de equivocaciones del lenguaje escrito han recibido diversas denominaciones en función del período histórico y de la disciplina de estudio. En el campo de la crítica textual, se puso como ejemplo de “atracción” la forma *piereden*

² Desde la perspectiva del usuario, el impacto de las erratas (o, en su caso, de los errores ortográficos) puede variar desde una simple molestia, hasta errores científicos o con implicaciones legales, pasando por una percepción de dejadez que afecte a la credibilidad del comunicador científico/profesional o de la propia base de datos (Gardner, 1992, p. 52).

(por *pierden*) (Blecua, 1983). El fenómeno de la “influencia” entre letras como causa de las erratas se ha mencionado desde el PLN, por ejemplo, *évidemement* (por *évidemment*); *aéoroport* (por *aéroport*) o *bubget* (por *budget*) (Ren & Perrault, 1992, p. 411). En el marco de la sicolingüística, se expuso como ejemplo de “anticipación” (o “migración”) *bfiber* por *fiber* (Grudin, 1983, p. 133) y de “perseveración” *difffer* (por *differ*) (Logan, 1999). Con anterioridad a todos ellos, ya se había descrito desde ese mismo terreno la anticipación (entre palabras, a mayor abundamiento): *wrapid** *writing* (Lashley, 1951)³. Dichas manifestaciones se enmarcan en lo que en psicología se denomina *serial order*, fenómeno que se ha estudiado ampliamente en relación con la producción hablada, pero también con la escrita (Snyder & Logan, 2014, p. 1697).

Por último, poniendo el foco en uno de los elementos fundamentales del terreno metalexicográfico, no debe extrañar que en algunos estudios no se incluyan entre los requisitos para que un ejemplo sea bueno el hecho de que no contenga erratas (Kosem, Koppel, Kuhn, Michelfeit & Tiberius, 2019; Geyken & Lemnitzer, 2024, p. 212-214), pues, entendemos, se trata de una condición tan básica (léase, fundamental) que se da por sentada. Otros autores sí refieren expresamente ese requisito (Leroyer, 2018, pp. 445-446).

2.3.2. Otras anomalías (*mayúscula, acento ortográfico, signos de puntuación y diseño*)

El uso de mayúsculas iniciales en lemas es un aspecto importante que el lexicógrafo ha de tener presente al realizar su obra. Para evitar opciones no acordes, como escribir todos los lemas con letra inicial mayúscula, con letra inicial minúscula añadiendo una nota explicativa (v. gr. *america* por *America*) o con todas las letras en mayúsculas, Svensén formula: “the basic rule is

³ En otro lugar, se detectaron en diccionarios especializados bilingües numerosos ejemplos que revelaban problemas motores de los tipos descritos (algunos de ellos relacionados con el acento ortográfico): *calientamiento*, *inmunidad*, *economía agraria**, *también cambién**, *selección jurisdiccional** (Rodríguez-Rubio Mediavilla, 2021, p. 133, p. 141).

that the lemma should be presented in the graphic form normally used within a sentence in continuous text” (2009, pp. 109-110). La relevancia de las letras iniciales mayúsculas en relación con nombres propios fue expuesta anteriormente en el terreno del procesamiento del lenguaje natural (Ren & Perrault, 1992, pp. 409-410). Lorentzen y Trap-Jensen (2016, p. 140) mencionan la problemática de los cambios ortográficos que, en 2012 y en el contexto danés, afectaron al uso de la mayúscula en siglas como bnp/BNP.

En relación con las anomalías en las que interviene el acento ortográfico, Ren y Perrault realizaron la siguiente clasificación de los errores (coincidente con la tipología general que hemos referido arriba en relación con las erratas o errores ortográficos): *insertion* (éssai por essai), *deletion* (agées por âgées), *substitution* (évènement por événement) y *repositioning* (chomâge por chômage) (1992, p. 410). Los mismos autores también incluyeron en su estudio errores de puntuación que tenían que ver con el uso de los guiones en palabras compuestas o con el genitivo sajón (pp. 410-411).

En lo concerniente al diseño (nuevamente con el usuario en mente), cabe citar las palabras de Levy y Steel:

Users appreciate regular patterns of engagement with the text and this is encouraged through precise and, oftentimes, special use of particular font variants, for example bold for the headword, a smaller font for the definition, and italics for the examples etc. Familiarity and routine are important in design, and of course this principle applies for all dictionaries (2015, p. 179).

3. IMÁGENES PUBLICITARIAS Y MULTIMODALIDAD LEXICOGRÁFICA

En general, la idoneidad de las imágenes que aparecen en los diccionarios puede abordarse desde múltiples prismas, con referencia a los beneficios y utilidad de las imágenes, pero también a la sobrecarga cognitiva (*cognitive load*) que pueden acarrear (Dziemianko, 2022, 2024). Retomando conceptos de épocas en las que aún no se había desarrollado la multimodalidad lexicográfica digital, cabe mencionar también el enfoque de la efectividad (*effectivity*) (Swanepoel, 2001).

Los diccionarios electrónicos (aunque habría que matizar a qué tipo de diccionario electrónico se está aludiendo) permiten a los lexicógrafos pasar de una disposición estática a una dinámica, con una estructura de entradas lexicográficas basada en diferentes capas (Gouws, 2014, p. 164). Ahora bien, el dinamismo puede ser un arma de doble filo en lexicografía, ya que, por ejemplo, una animación puede desorientar al consultante (Lew & Doroszewska, 2009, pp. 253-254). Fuertes-Olivera postula que los datos lexicográficos deben poder ser convertidos en información lexicográfica “in a single cognitive process or click”. A continuación, el autor asegura que esa condición no se cumple en la mayoría de los diccionarios españoles, por mor de defectos en su tratamiento o presentación (2023, p. 31).

Dziemianko (2012, p. 197) alude a la merma en cuanto a la utilidad (*usefulness*), la recepción, producción y aprendizaje (Dziemianko, 2019, p. 5) y el tiempo de consulta (Dziemianko, 2020, p. 394) que los anuncios publicitarios generan en las versiones gratuitas de diccionarios en línea como *Longman Dictionary of Contemporary English* (LDOCE) y *Oxford Advanced Learner's Dictionary* (OALD9). No en todos los estudios la autora analizó todos esos diccionarios y, en ocasiones, los resultados interestudio no fueron concordantes (Dziemianko, 2020, p. 394). Por su parte, Ferrett y Dollinger (2021) manifiestan que los anuncios (algunos de los cuales cambian sin cesar e incluyen efectos llamativos como las animaciones) son una de las principales desventajas de los diccionarios electrónicos respecto de los diccionarios en papel: “It is clear that advertisements, ‘noise’, which are generally required for commercial reasons, make demands on presentation space that run counter to lexicographical needs” (Ferrett & Dollinger, 2021, p. 68, p. 74).

Las palabras de Dziemianko ponen de relieve que la cuestión de las imágenes publicitarias (más concretamente, cómo afectan en términos de uso) no se ha resuelto:

This is a significant research gap, as even top-quality free online dictionaries, including the most prestigious English monolingual dictionaries for foreign learners,

allow targeted advertisements. It should be remembered that online dictionaries are not like many other websites, which may be read for pleasure, joy or broad overviews. They are consulted to satisfy punctual linguistic needs (Tarp 2012: 111), which requires focused attention. Yet, at the same time, advertisements vie for dictionary users’ attention and may affect lexicographic information costs, i.e. the difficulty or inconvenience associated with dictionary lookup (Nielsen 1999: 111) (Dziemianko, 2020, p. 378).

4. MATERIALES Y METODOLOGÍA

4.1. MATERIALES

A los diccionarios de Cambridge que se encuentran en el portal <https://dictionary.cambridge.org/> se los ha llamado “*Cambridge Dictionaries Online*” (Jackson, 2018; Engelberg & Storrer, 2024, p. 56), “*Cambridge English Dictionaries, CaED*” (Geyken & Lemnitzer, 2024, p. 225) o “*Cambridge-Online*” (Engelberg & Storrer, 2024, pp. 49-50). En el presente trabajo, se usa la primera de las denominaciones, manteniendo el nombre inglés en singular y, cuando sea oportuno, la sigla “CDO”. En lugar de como diccionarios individuales, las direcciones “English-Spanish” y “Spanish-English” se consideran partes de un mismo diccionario (*Cambridge Dictionary Online*).

Sobre la base de la tipología de Engelberg y Storrer, se puede considerar el *Cambridge Dictionary Online* como un diccionario general (categoría de usuario *learners’ dictionary*) bilingüe, del tipo portal lexicográfico (híbrido entre un motor de búsqueda y una red de diccionarios o *dictionary network*), además de un portal léxico en virtud de la cobertura de aspectos culturales, enseñanza-aprendizaje de vocabulario, listas temáticas de palabras, juegos y concursos de preguntas, secciones de “palabra del año”, etc. (2024, pp. 31-59). A partir de las nociones *persistence* y *identity* (Herold, Meyer & Wiegand, 2024, pp. 24-26), cabe asociar a los diccionarios de Cambridge con una clara vocación de permanencia.

Las categorías de anomalías estudiadas en el presente trabajo son: “mayúscula”, “puntuación”, “diseño”, “acento ortográfico”, “espacio”, “errata de no

palabra” (ENP) y “errata de palabra” (EP)⁴. Para las ENP y EP, se van a usar las subcategorías sustracción, adición-repetición, sustitución o transposición de una o más letras y, por otro lado, sustracción, adición-repetición, sustitución o transposición de una o más palabras, respectivamente.

Las ubicaciones en las que se encuentran las anomalías del presente estudio son: equivalente, explicación, ejemplo (léase, frase ilustrativa), definición y lema.

4.2. METODOLOGÍA

En el presente estudio, no se ha empleado el término *anomalía* en sentido estricto tal y como suele entenderse en el terreno de las ciencias de la informática, sino general: “[defecto de forma o de funcionamiento](#)”.

Se ha realizado una revisión manual de los datos lexicográficos. Se analizaron las anomalías del texto en español de las entradas correspondientes a la letra “A” de las partes English-Spanish (desde *A & E* hasta *azygous*) y Spanish-English (desde *a* hasta *azuzar*) del [Cambridge Dictionary Online](#).

Cuando aparecieron varias anomalías seguidas, computó cada una de ellas y, si procedía, se desglosaron por tipo. Así, en la entrada “[A-bomb](#)” aparece “Bomba Atomica”, que contiene tres anomalías, dos relativas al uso de la mayúscula y una relativa al acento ortográfico (ver Anexo 1, parte English-Spanish).

Se analizaron solo las palabras en español, porque, de haber incluido también aquellas en inglés, la complejidad habría superado los márgenes del trabajo tal y como se concibió, ya que se habría tenido que incluir un aspecto más (en clave bilingüe) a los previstos: revisión de las anomalías en ambas partes del diccionario, análisis de los tipos generales y subtipos de anomalías y análisis de la ubicación de dichas anomalías. Como excepción a la norma de que solo se analizó el texto en español, cabe decir que las anomalías relativas al diseño sí incluyeron texto en inglés, ya que las implicaciones no

⁴ Como alternativas a las categorías del PLN *non-word errors* y *real-word errors*, se han mantenido las denominaciones en español acuñadas en otro lugar: *errata de no palabra* y *errata de palabra* (Rodríguez-Rubio Mediavilla, 2021).

son meramente lingüísticas. Por ejemplo, la Imagen 1 (captura de pantalla del CDO) revela la repetición del lema *a barrage of something*.

Las anomalías de diseño de la parte Spanish-English implican ausencia de espacio, pero, en realidad se trata de un solapamiento de segmentos o campos textuales, de ahí que se incluyeran en la categoría “diseño”. La Imagen 2 muestra que existen dos tipos de texto diferenciados en el contenido de

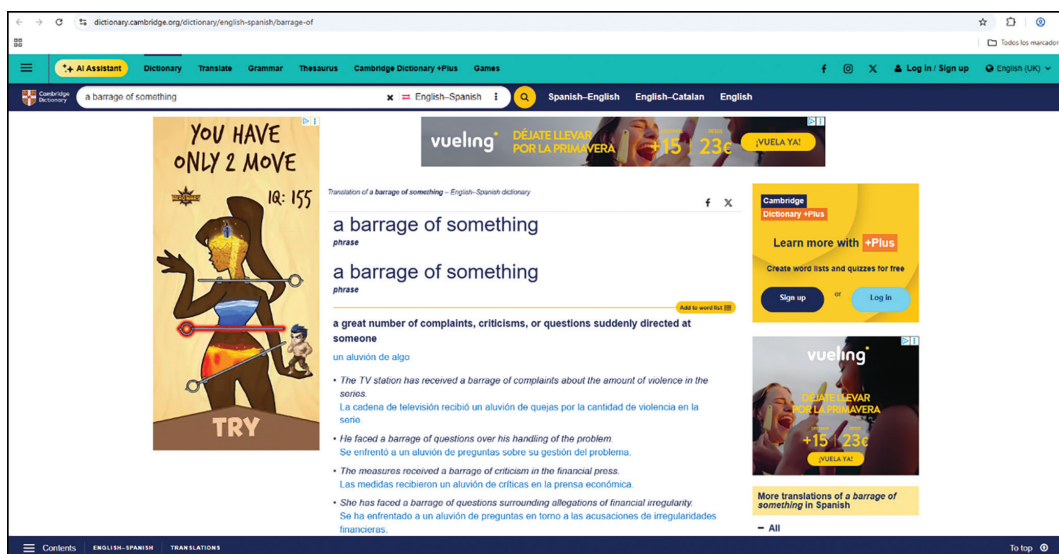


IMAGEN 1: Anomalía de diseño (repetición de lema)

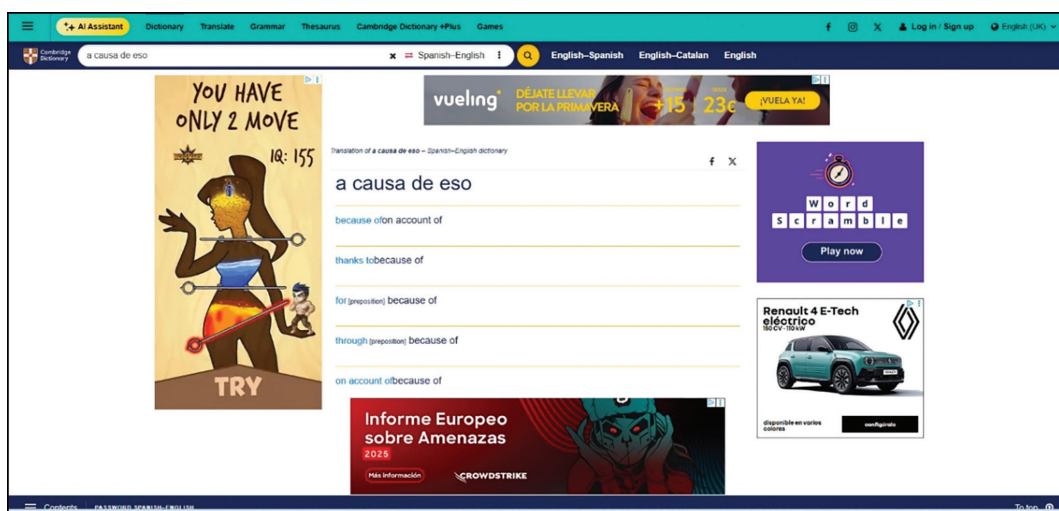


IMAGEN 2: Anomalía de diseño (solapamiento de segmentos textuales)

los equivalentes en inglés de la entrada “a causa de eso”: 1) uno escrito en letra azul, que reza por ejemplo “because of”, clicando en el cual se accede mediante hipervínculo a dicho lema de la parte English-Spanish; y 2) otro, escrito en negro y no vinculado que reza “on account of”. Al seleccionar con el cursor el fragmento “ofon” se observa un solapamiento entre ambos elementos. Los segmentos “thanks tobecause of” y “on account ofbecause of” también presentan el problema, no así “for [preposition] because of” y “through [preposition] because of”. Ver ejemplos similares en el Anexo 1.

En el *Cambridge Dictionary Online*, aparecen imágenes de muy diverso tipo, tanto publicitarias (por ejemplo, anunciando juegos electrónicos) como “lingüísticas” (asociadas a blogs o cuestionarios/juegos). Las publicitarias, ya sean estáticas o animadas, aparecen a menudo de manera errática e intercaladas entre los componentes microestructurales. En la parte Spanish-English en mayor medida, a veces las imágenes no se cargan y no aparecen (por tanto, dejando un vacío que genera una sensación de discontinuidad textual cuando el usuario se mueve hacia arriba y abajo en la página). Otras (si bien pocas), las imágenes centrales (no laterales) no encajan exactamente con el texto escrito, de forma que ambos textos (el escrito y el pictórico) se solapan. En el presente estudio, esos desajustes espaciotemporales no se han considerado como anomalías, pero invitamos a los editores o desarrolladores del diccionario a que revisen este aspecto.

Finalmente, no se ha tratado la problemática del uso de mayúscula o minúscula en función de mudanzas en la norma ortográfica (Lorentzen & Trap-Jensen, 2016, p. 140), sino los casos en los que: 1) un determinado elemento se desvía de la “norma” o costumbre interna del propio *Cambridge Dictionary Online* [en la definición de *aparte*, se indica “Que está separado o es distinto”, vulnerando la norma no escrita pero constatable de forma constante en el diccionario de que las definiciones empiezan con letra inicial minúscula (<https://dictionary.cambridge.org/dictionary/spanish-english/aparte>)]; 2) no se siguen normas asentadas [en la entrada “agent”, figura

“hollywood” con letra inicial minúscula (<https://dictionary.cambridge.org/dictionary/english-spanish/agent>), mientras que en “*Arab*”, aparece “*estados unidos*” en lugar de “*Estados Unidos*”.

5. RESULTADOS

En la Figura 1, los resultados de incidencia general de anomalías arrojan un número 2,6 veces mayor de anomalías en la parte English-Spanish (771) que en Spanish-English (290). En la parte English-Spanish, sobresalen las categorías de mayúsculas y puntuación (que representan en conjunto un 63,9 % de las anomalías de dicha parte), mientras que en Spanish-English destaca de manera sensiblemente más acusada la categoría diseño (84 %). La proporción de *erratas de palabra* (EP) respecto de las *erratas de no palabra* (ENP) es similar en ambas partes (el número de las EP es 2,1 y 2,5 veces superior al de ENP en la parte English-Spanish y Spanish-English, respectivamente). Las

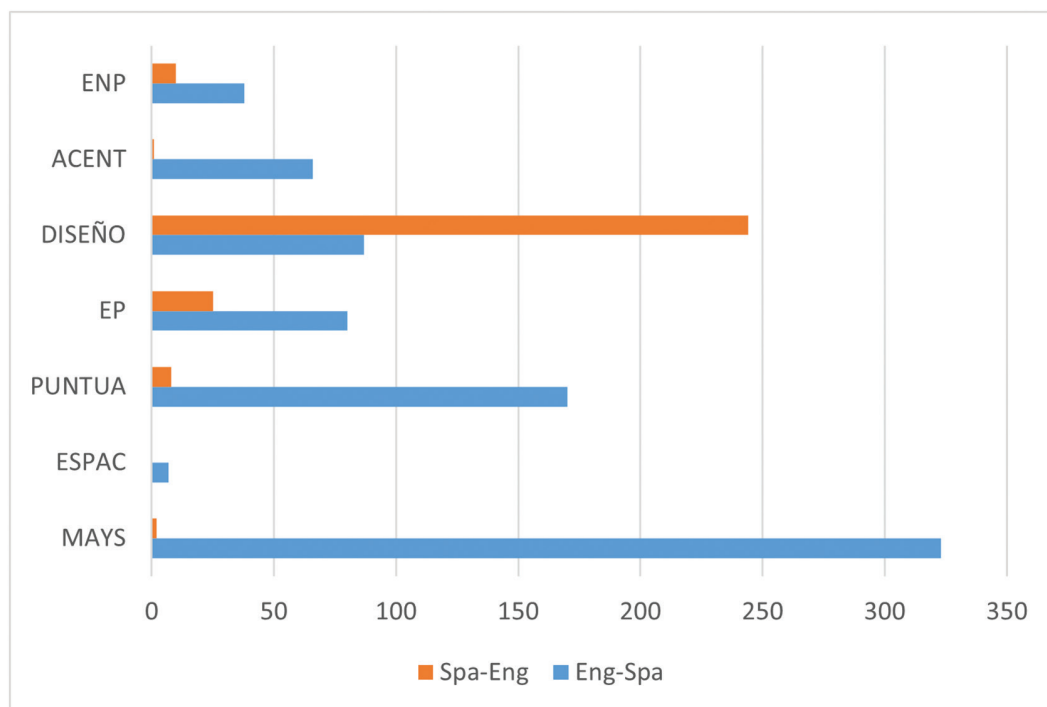


FIGURA 1: Incidencia de anomalías (general)

anomalías relativas al espacio ocupan en ambas partes la última posición (de hecho, en Spanish-English no se han detectado anomalías de este tipo).

Las Figuras 2 y 3 muestran una distribución cualitativamente homogénea de los subtipos de errata, con los siguientes resultados de mayor a menor incidencia en ambas partes: sustitución, sustracción, adición-repetición y transposición en cuanto a las *erratas de palabra* y, por otro lado, adición-repetición, sustracción, sustitución y transposición en lo que se refiere a las *erratas de no palabra*. El desglose en función de las partes revela una cifra parecida a la del dato de incidencia general, con un número agregado de EP y ENP 3,2 veces mayor en la parte English-Spanish (113) en comparación con Spanish-English (35).

En términos de ubicación, las Figuras 4 y 5 expresan claramente que la mayoría de las anomalías se sitúan en el campo del “equivalente”: 438 (56,8 %) en la parte English-Spanish y, de manera más acusada, 244 (84,1 %) en Spanish-English. En el suelo se sitúan las anomalías situadas en el campo “lema”, con 87 (11,3 %) en English-Spanish y un testimonial 1 (0,3 %) en Spanish-English. El hecho de haber

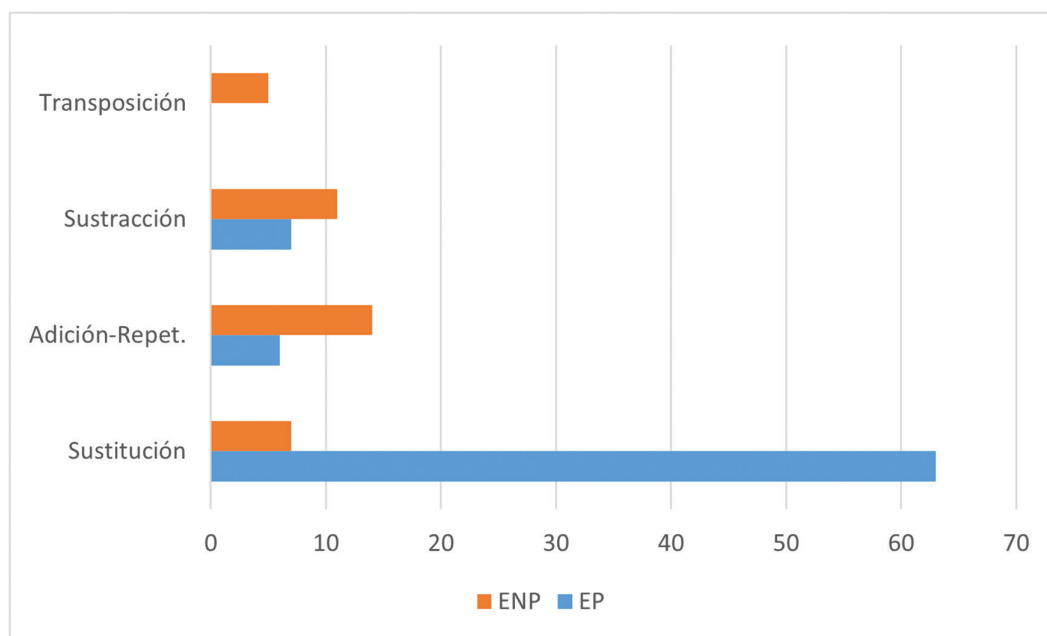


FIGURA 2: Incidencia de subtipos de errata (English-Spanish)

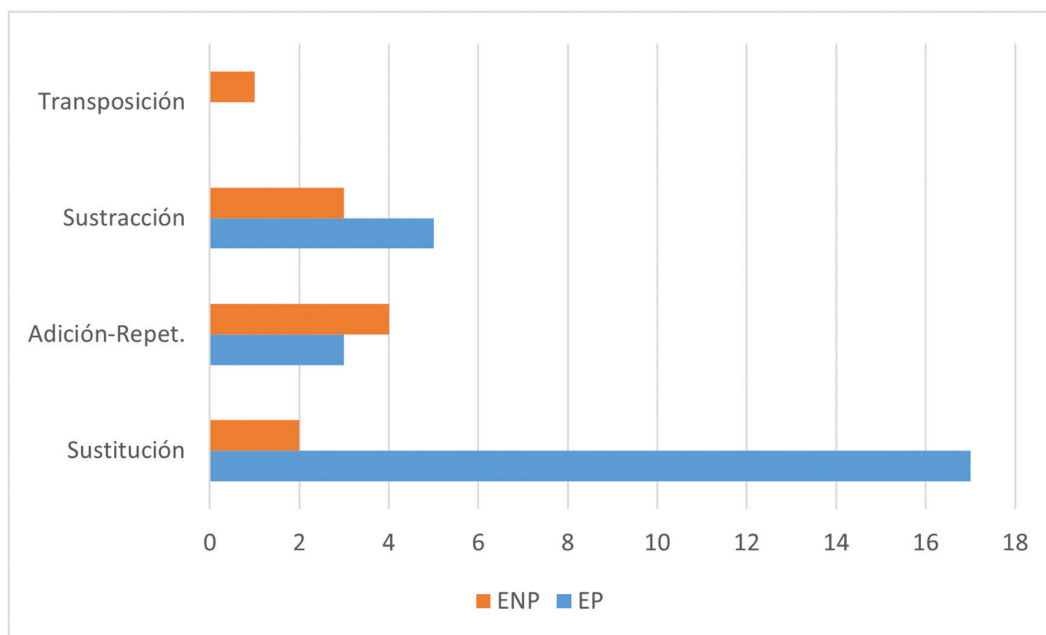


FIGURA 3: *Incidencia de subtipos de errata (Spanish-English)*

analizado el texto en español, pero no en inglés, hace que en English-Spanish no se hayan registrado anomalías en la ubicación “definición”. Por último, en English-Spanish, se observaron anomalías en todas las ubicaciones, pero en Spanish-English la ubicación “lema” es, como acabamos de explicitar, testimonial. El análisis desgranado de los elementos más salientes de las Figuras 4 y 5 muestra que, a menudo, no se trata de una acumulación de defectos inconexos. Así, en la parte English-Spanish:

- a) de las 323 anomalías de mayúsculas, 315 (97,5 %) responden al mismo problema, consistente en haber escrito con letra inicial mayúscula los equivalentes que, según la norma no escrita del propio CDO, deben empezar con minúscula;
- b) de las 129 anomalías en explicaciones, 128 (99,2 %) consisten en la indicación de dos puntos (:) y una coma (,) seguidos. Por ejemplo, en “forma abreviada de ‘also known as’; alias” de la entrada “[aka](#)”;
- c) las 87 (100 %) anomalías de diseño consisten en la repetición del lema de referencia.

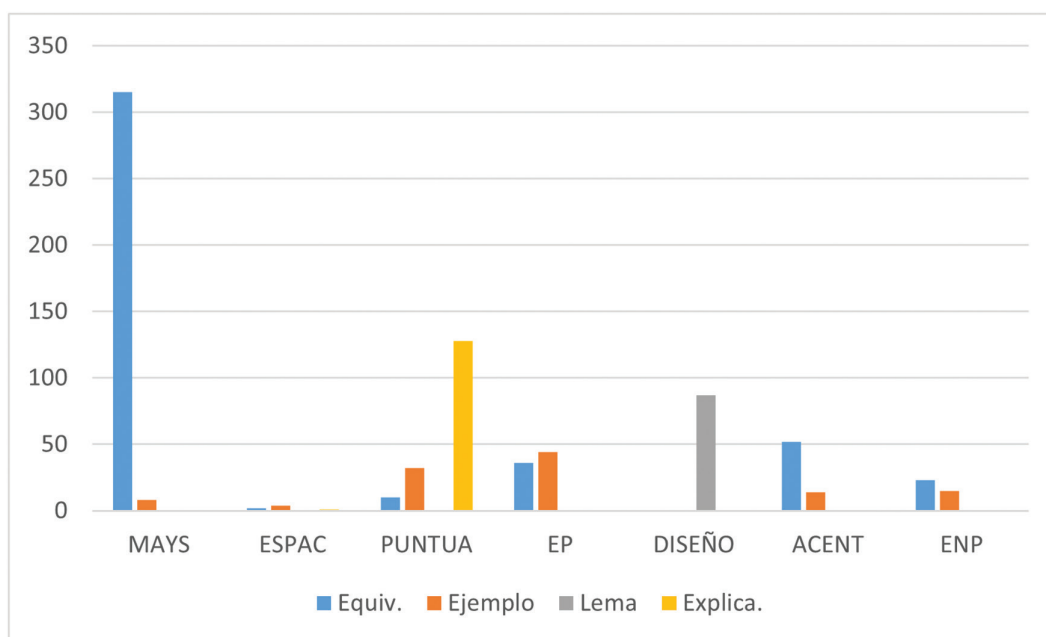


FIGURA 4: Ubicación de las anomalías (English-Spanish)

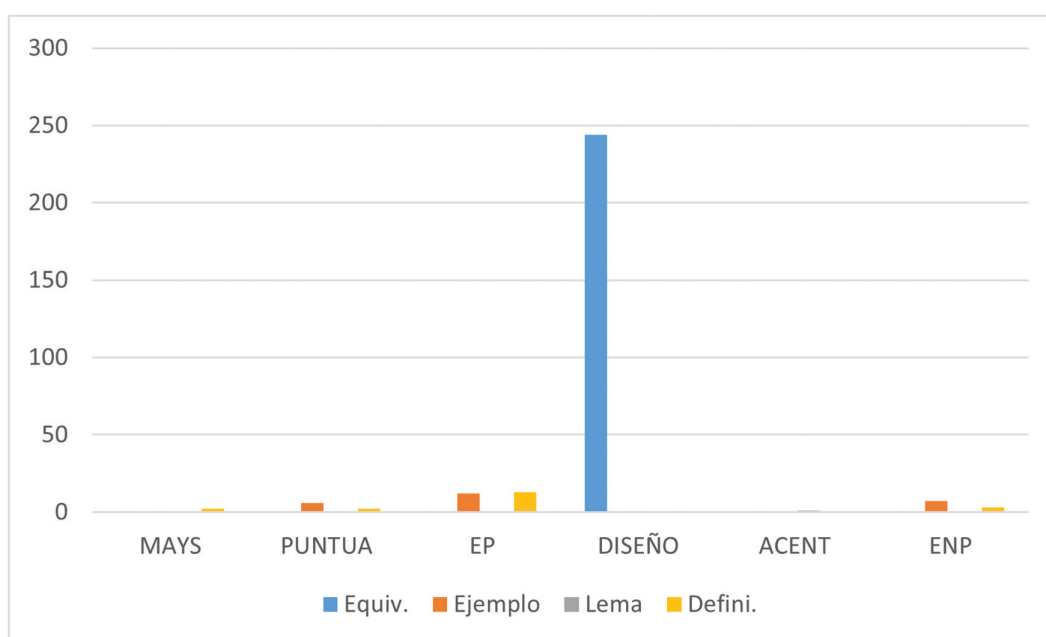


FIGURA 5: Ubicación de las anomalías (Spanish-English)

De manera similar, en la parte Spanish-English, las 244 anomalías de diseño revelan el mismo tipo de solapamiento de segmentos textuales explicado arriba.

En los resultados del presente estudio no podían faltar ejemplos de erratas causadas por fenómenos motores (dada la universalidad de estos últimos). La Tabla 1 contiene un ejemplo de anticipación y otro de perseveración (todos interpalabra) en cada una de las partes del CDO.

	Anticipación	Perseveración
English-Spanish	<i>terreno utilizad<u>a</u>* para la agricul- tura</i> (ejemplo de la entrada “agri- culture”: https://dictionary.cambridge.org/dictionary/english-spanish/agriculture)	<i>le enfad<u>ó</u> much<u>ó</u>*</i> (ejemplo de la entrada “angry”, sección “TRANSLATION of angry ”: https://dictionary.cambridge.org/dictionary/english-spanish/angry)
Spanish-English	<i>negocios<u>os</u>* sucios</i> (ejemplo de la acepción “desper- tar sospecha de ilegalidad” de la entrada “apestar”: https://dictio- nary.cambridge.org/dictionary/ spanish-english/apestar)	<i>una casa de campo de* muy apa- cible</i> (ejemplo de la acepción “agradable, sereno y tranquilo” de la entrada “apacible”: https://dictionary.cambridge.org/dictionary/spani- sh-english/apacible)

TABLA 1: Ejemplos de anticipación y perseveración (Eng-Spa/Spa-Eng)

Ver en el Anexo 1 otros ejemplos de anticipación o perseveración intra o interpalabra, ya sea en English-Spanish o en Spanish-English.

En clave bilingüe, se registraron aparentes interferencias, como:

- *adjectivo* (por *adjetivo*, posible influencia del vocablo inglés *adjective*), en la entrada “adj”.
- *objectivo* (por *objetivo*, posible influencia de *objective*) en la frase ilustrativa *comenzó con el objetivo* de acabar con la crisis* de la entrada “aim”.

Si bien no hemos aplicado un enfoque semántico en este estudio, cabe mencionar un ejemplo de doble negación, un fenómeno universal amplia-
mente observado en contextos de aprendizaje y comunicación tanto cuando
interviene alguna lengua extranjera como cuando nos limitamos a la propia
lengua materna. En las entradas del CDO “antiporn” y “antipornography”
encontramos la formulación “antipornografía, contra la antipornografía*”, que
debería rezar: “antipornografía, contra la pornografía”.

Además, aunque no se realizó un estudio de la posible reproducción de
las anomalías en varias entradas (o en el seno de una determinada entrada)

del diccionario, podemos mencionar que la formulación “con interés en lo artesana*” (por “con interés en lo artesano”) aparece en las entradas vinculadas “[arty-crafty](#)” y “[artsy-craftsy](#)”. Otros casos de aparente reproducción, como la ausencia de espacio en el mismo contenido de las entradas relacionadas “a gatas” y “a cuatro patas” (*on all fourson hands and knees*) o “a voz en cuello” y “a voz en grito” (*at the top of one’s voicevery loudly*) responden en realidad a anomalías de diseño (como se ha explicado en el apartado “Métodos”).

Entre los ejemplos de palabras que deberían tener letra inicial mayúscula en lugar de minúscula, al objeto de diferenciar la palabra del correspondiente nombre propio, se encuentran *Clara* (por *clara*, en el lema [albumen](#)) y *Angustia* (por *angustia*, en el lema [angst](#)).

6. DISCUSIÓN

En el presente estudio, se ha ido más allá de la (beneficiosa y seguramente necesaria) colaboración o interacción de los usuarios de diccionarios explicitada en Klosa-Kückelhaus y Tiberius (2024, pp. 71-73), para presentar un conjunto representativo de anomalías del *Cambridge Dictionary Online* (partes English-Spanish y Spanish-English). Se ha demostrado así la vigencia de la lucha contra las anomalías en los diccionarios electrónicos. Aunque en el presente trabajo solo se han analizado anomalías formales del texto en español, en un futuro podrían abordarse aspectos de trascendencia semántica (por ejemplo, desviaciones o dobles negaciones) o, en clave bilingüe, que tengan que ver con la traducción. A su vez, convendría explorar la posible reproducción de anomalías en diferentes contenidos del diccionario; aspecto, el de la reproducción de anomalías, que ya se abordó en otro lugar en relación con plantillas lexicográficas y erratas de diccionarios especializados bilingües en papel (Rodríguez-Rubio & Fernández-Quesada, 2020a).

Se desconoce el método de introducción y revisión de contenido lexicográfico que se usa en el caso del *Cambridge Dictionary Online*. Se trata

de un aspecto relevante, pues algunos *modus operandi* pueden facilitar la tarea de revisión, mientras que otros pueden dificultarla. Como se ha aclarado arriba, algunos modelos de edición pueden ser más propensos al error cuando se lleva a cabo una edición manual (Simões & Salgado, 2022, p. 425). Cabe suponer que algunos de los errores lingüísticos presentes en los diccionarios electrónicos pueden deberse precisamente a la edición manual.

A su vez, sería interesante explorar la utilidad de las funciones de revisión ortográfica y gramatical de los programas de gestión lexicográfica (DWS) de los diccionarios en línea y si los errores o anomalías están claramente marcados para que el lexicógrafo pueda localizarlos con relativa facilidad a nivel más profundo, pero también mediante una ojeada a la vista WYSISWYG (*what-you-see-is-what-you-get*). Más específicamente, como se sugirió en otro lugar, sería conveniente averiguar cómo consigue el corrector ortográfico diferenciar la granularidad de un texto bilingüe en microestructuras complejas, donde intervienen dos lenguas de revisión (Rodríguez-Rubio & Fernández-Quesada, 2020b, pp. 397-398). Los diccionarios electrónicos pueden incluir funciones de edición que sirven para detectar anomalías, pero sigue quedando una parte de la tarea que ha de realizar el revisor humano. A partir de aquí, una derivada a investigar en futuros estudios sería determinar con qué periodicidad y nivel de detalle se revisan los diccionarios electrónicos, aspecto que planteamos hace unos años en un estado de la cuestión (meta)lexicográfica (Fernández-Quesada & Rodríguez-Rubio, 2022). Esa pretensión se toparía, no obstante, con el impedimento de que, como sugieren de manera constante y extendida en el tiempo los investigadores, los editores lexicográficos tradicionalmente han brindado escasa información sobre sus procesos internos (Swanepoel, 2001, p. 170; Klosa-Kückelhaus & Tiberius, 2024, p. 62). Esa parca información hace, a su vez, que no se sepa en qué medida los ejemplos del *Cambridge Dictionary Online* son inventados o reutilizados.

Existen numerosos manuales de Cambridge, entre otros, un *Handbook 2025-Regulations and guidance for administering Cambridge exams* y una

Editorial Services Style Guide for Academic Books orientado a autores. Pero se echa en falta una guía de uso del *Cambridge Dictionary Online* en consonancia con un recurso de consulta académica de primer orden con vocación de permanencia. Tal vez la megaestructura se asocie en mayor medida a los diccionarios en papel, pero componentes como “advertencias de uso”, “artículos de muestra” o aquellos que expliquen el manejo de la obra también son útiles en el marco de la lexicografía electrónica (Ramírez Sánchez, 2023, p. 850). Esas consideraciones coinciden parcialmente con las de Klosa-Kückelhaus y Tiberius, quienes incluyen el manual de usuario (o *guided tour*) entre los elementos integrantes del proceso de elaboración de diccionarios digitales (*Internet dictionaries*) (2024, p. 67).

Partiendo del desconocimiento del número de editores de los diccionarios de Cambridge, la afirmación de que cuando hay pocos editores las entradas lexicográficas deben someterse a una rigurosa revisión antes de su publicación (Herold, Meyer & Wiegand, 2024) debe aplicarse *a fortiori* a los casos en los que hay más editores.

Como se ha explicado arriba, las imágenes publicitarias pueden provocar ruido. Sin embargo, los anuncios pueden contribuir decididamente a la continuidad de los recursos lexicográficos de acceso gratuito. En el caso del *Cambridge Dictionary Online*, muchas de las imágenes publicitarias móviles o animadas incluyen figuras (por ejemplo, personajes de juegos *online*) y colores en extremo llamativos, lo que puede, en efecto, desorientar o incluso molestar al usuario. Además, esas imágenes van cambiando sin cesar y su tiempo de carga produce a menudo un desacople temporal entre el momento (normalmente breve) de la consulta del usuario y la aparición de la imagen, de manera que quedan espacios aparentemente en blanco intercalados en el contenido lexicográfico de numerosas entradas (especialmente en la parte Spanish-English). A pesar de que se descartó computar como anomalías dichos espacios vacíos, y a pesar de que propiamente hablando esas imágenes no pueden considerarse datos lexicográficos de todas formas, conviene

apuntar que las disfunciones que provoca su comportamiento errático no concuerdan con la lógica de ausencia de defectos en el tratamiento o la presentación que Fuertes-Olivera aplica a los datos lexicográficos (2023, p. 31).

7. CONCLUSIONES

La proporción de 2,1 veces *erratas de palabra* (80) frente a *erratas de no palabra* (38) en la parte English-Spanish y de 2,5 veces *erratas de palabra* (25) frente a *erratas de no palabra* (10) en la parte Spanish-English concuerda con la idea general expresada en la literatura de manera sostenida en el tiempo de que los *real-word errors* son difíciles de detectar o corregir. A efectos prácticos, la distribución cualitativamente homogénea de los subtipos de errata puede servir de guía con vistas a una hipotética revisión del *Cambridge Dictionary Online* en su totalidad. Así, en ambas partes del diccionario, se ha podido diferenciar entre el “comportamiento” de las *erratas de palabra* (con una clara prevalencia de las sustituciones, seguidas de las sustracciones y adiciones-repeticiones) y el de las *erratas de no palabra* (mucho más repartidas, pero siempre siguiendo el mismo orden de adiciones-repeticiones, sustracciones, sustituciones y transposiciones).

Más interesantes para los editores o lexicógrafos del CDO sean tal vez las conclusiones respecto de la ubicación de las anomalías de la letra “A”, pues a partir de ellas se puede prever en cierta medida dónde se situarán las anomalías del resto de letras; máxime cuando se ha observado una coincidencia en las causas de los defectos de mayúscula, puntuación y diseño en la parte English-Spanish y de diseño en Spanish-English. Una sistematicidad de anomalías precisa de un enfoque integral que incluya una revisión y unas medidas correctivas acordes. Nuevamente, los resultados pueden dirigir u ordenar la revisión que se lleve a cabo.

Los avances tecnológicos ofrecen potentes recursos de gestión lexicográfica, pero encierran la necesidad de una creciente atención por parte del lexicógrafo/editor que atienda a la mayor complejidad de los procesos y

a la multiplicación, expansión e interconexión de recursos. En el campo de la revisión y de la calidad de los datos (como en tantos otros) se ha de encontrar un equilibrio entre productividad (léase, cantidad) y calidad, entre la posibilidad de crear productos lexicográficos avanzados y complejos y la pertinencia de realizar un trabajo fino de cohesión y corrección tanto en el fondo como en la forma. La posibilidad de modificar en tiempo real el contenido lexicográfico exige una revisión dinámica y flexible en consonancia, al tiempo que detallista y rigurosa.

Pese a la existencia de un marco tecnológico que ofrece medios avanzados para promover la calidad de los datos, el papel del revisor humano continúa siendo fundamental. En línea con lo manifestado en otro lugar (Fernández-Quesada & Rodríguez-Rubio, 2022, p. 24) la evidencia de que la tarea de revisión de los diccionarios es un proceso inacabable no debe ser óbice sino acicate para que se sigan realizando revisiones periódicas de las grandes obras de referencia. En el plano académico, los resultados del presente estudio en un diccionario en línea, sumados a los de nuestras anteriores investigaciones sobre erratas en diccionarios en papel, muestran claramente la pertinencia de seguir profundizando en la búsqueda de la corrección formal lexicográfica. Tal vez, de hecho, la alianza entre el mundo profesional y el académico pueda alumbrar una oportunidad de conseguir obras o recursos más afinados en una disciplina milenaria como la lexicografía que aún sigue en pie a pesar de las dificultades.

REFERENCIAS BIBLIOGRÁFICAS

- Abel, A. (2012). Dictionary writing systems and beyond. En S. Granger & M. Paquot (eds.), *Electronic Lexicography* (pp. 83-106). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199654864.003.0005>
- Abel, A. & Meyer, M. (2024). User Participation. En A. Klosa-Kückelhaus (ed.), *Internet Lexicography* (pp. 227-260). De Gruyter. <https://doi.org/10.1515/9783111233758-009>
- Adams, N. & Maxwell, M. (2016). Somali Spelling Corrector and Morphological Analyzer. En C. Soria, L. Pretorius, T. Declerck, J. Mariani, K. Scannell & E. Wandl-Vogt (eds.), *Proceedings of the LREC 2016 Workshop “CCURL 2016 – Towards an Alliance for Digital Language Diversity”* (pp. 22-29).

- Barbierik, K., Děngeová, Z., Holcová, M., Jarý, V., Liška, T., Lišková, M. & Virius, M. (2014). Simple and Effective User Interface for the Dictionary Writing System. En A. Abel, C. Vettori & N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus. Bolzano/Bozen, Italy* (pp. 125-136). EURAC research.
- Blecua, A. (1983). *Manual de crítica textual*. Editorial Castalia.
- Bloodgood, M. & Strauss, B. (2016). Data Cleaning for XML Electronic Dictionaries via Statistical Anomaly Detection. En *Proceedings of the IEEE Tenth International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 2016* (pp. 79-86). <https://doi.org/10.1109/ICSC.2016.38>
- Coward, D. F. & Grimes, Ch. E. (2000). *Making Dictionaries: A Guide to Lexicography and the Multi-Dictionary Formatter*. SIL International.
- Dashti, S. M., Bardsiri, A. K. & Bardsiri, V. Kh. (2018). Correcting real-word spelling errors: A new hybrid approach. *Digital Scholarship in the Humanities*, 33(3), 488-499. <https://doi.org/10.1093/llc/fqx054>
- Dziemianko, A. (2012). Why One and Two Do Not Make Three: Dictionary Form Revisited. *Lexikos*, 22, 195-216. <https://doi.org/10.5788/22-1-1003>
- Dziemianko, A. (2019). The role of online dictionary advertisements in language reception, production, and retention. *ReCALL*, 31(1), 5-22. <https://doi.org/10.1017/S0958344018000149>
- Dziemianko, A. (2020). Smart Advertising and Online Dictionary Usefulness. *International Journal of Lexicography*, 33(4), 377-403. <https://doi.org/10.1093/ijl/ecaa017>
- Dziemianko, A. (2022). The usefulness of graphic illustrations in online dictionaries. *ReCALL*, 34(2), 218-234. <https://doi.org/10.1017/S0958344021000264>
- Dziemianko, A. (2024). Pictures in Online Dictionaries: Shall We See Them? *GEMA Online - Journal of Language Studies*, 24(2), 31-53. <http://doi.org/10.17576/gema-2024-2402-03>
- Engelberg, S. & Storrer, A. (2024). A Typology of Internet Dictionaries and Portals. En A. Klosa-Kückelhaus (ed.), *Internet Lexicography* (pp. 31-59). De Gruyter. <https://doi.org/10.1515/9783111233758-003>
- Fernández-Quesada, N. & Rodríguez-Rubio, S. (2022). Estudio preliminar. El tratamiento del error textual y de la errata en la era digital: elogio de la corrección. En N. Fernández-Quesada & S. Rodríguez-Rubio (eds.), *Detección y tratamiento de errores y erratas: Un diagnóstico para el siglo XXI* (pp. 13-28). Dykinson. <https://doi.org/10.2307/j.ctv2gz3s5v.3>
- Ferrett, E. & Dollinger, S. (2021). Is Digital Always Better? Comparing Two English Print Dictionaries with Their Digital Counterparts. *International Journal of Lexicography*, 34(1), 66-91. <https://doi.org/10.1093/ijl/ecaa016>
- Foorthuis, R. (2021). On the nature and types of anomalies: a review of deviations in data. *International Journal of Data Science and Analytics*, 12, 297-331. <https://doi.org/10.1007/s41060-021-00265-1>

- Fossati, D. & Di Eugenio, B. (2008). I saw TREE trees in the park: How to Correct Real-Word Spelling Mistakes. En *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco (pp. 896-901). <https://doi.org/10.63317/4qig6sasadgh>
- Fuertes-Olivera, P. A. (2023). Structuring a Collection of Lexicographic Data for Different User and Usage Situations. *Lexikos*, 33(2), 22-42. <https://doi.org/10.5788/33-2-1832>
- Gardner, S. A. (1992). Spelling Errors in Online Databases: What the Technical Communicator Should Know. *Technical Communication*, 39(1), 50-53.
- Geyken, A. & Lemnitzer, L. (2024). The Automatic Extraction of Lexicographic Information. En A. Klosa-Kückelhaus (ed.), *Internet Lexicography* (pp. 191-226). De Gruyter. <https://doi.org/10.1515/9783111233758-008>
- Gouws, R. (2014). Article Structures: Moving from Printed to e-Dictionaries. *Lexikos*, 24, 155-177. <https://doi.org/10.5788/24-1-1256>
- Grudin, J. T. (1983). Error Patterns in Novice and Skilled Transcription Typing. En W. E. Cooper (ed.), *Cognitive Aspects of Skilled Typewriting* (pp. 121-143). Springer. https://doi.org/10.1007/978-1-4612-5470-6_6
- Gujral, E. (2023). *Survey: Anomaly Detection Methods*. Department of Computer Science and Engineering University of California Riverside (pp. 1-16). <https://www.semanticscholar.org/paper/Survey:-Anomaly-Detection-Methods-Gujral/75b78a54132e193b5a-c528645045e9ccbb98395d>
- Herold, A., Meyer, P. & Wiegand, F. (2024). The Technological Context for Internet Lexicography. En A. Klosa-Kückelhaus (ed.), *Internet Lexicography* (pp. 5-29). De Gruyter. <https://doi.org/10.1515/9783111233758-002>
- Jackson, H. (2018). English Lexicography in the Internet Era. En P. A. Fuertes-Olivera (ed.), *The Routledge Handbook of Lexicography* (pp. 835-856). Routledge. <https://doi.org/10.4324/9781315104942-34>
- Kallas, J., Koeva, S., Kosem, I., Langemets, M. & Tiberius, C. (2019). Lexicographic practices in Europe: A survey of user needs. *ELEXIS – European Lexicographic Infrastructure*, Deliverable D1.1. <https://bit.ly/2WJwSlC>
- Klosa-Kückelhaus, A. & Tiberius, C. (2024). The Lexicographic Process. En A. Klosa-Kückelhaus (ed.), *Internet Lexicography* (pp. 61-96). De Gruyter. <https://doi.org/10.1515/9783111233758-004>
- Koeva, S. & Blagoeva, D. (2013). The Dictionary Writing System Lexit and its Application in Bilingual Lexicography. *Cognitive Studies*, 13, 57-65. <https://doi.org/10.11649/cs.2013.003>
- Kosem, I., Koppel, K., Kuhn, T. Z., Michelfeit, J. & Tiberius, C. (2019). Identification and Automatic Extraction of Good Dictionary Examples: The Case(s) of GDEX. *International Journal of Lexicography*, 32(2), 119-137. <https://doi.org/10.1093/ijl/ecy014>
- Landau, S. I. (2001). *Dictionaries. The Art and Craft of Lexicography*. Cambridge University Press.
- Lashley, K. S. (1951). The Problem of Serial Order in Behavior. En L. A. Jeffress (ed.), *Cerebral mechanisms in behavior* (pp. 112-146). Wiley.

- Leroyer, P. (2018). The Oenolex Wine Dictionary. En P. A. Fuertes-Olivera (ed.), *The Routledge Handbook of Lexicography* (pp. 438-454). Routledge. <https://doi.org/10.4324/9781315104942-28>
- Levy, M. & Steel, C. (2015). Language learner perspectives on the functionality and use of electronic language dictionaries. *ReCALL*, 27(2), 177-196. <https://doi.org/10.1017/S095834401400038X>
- Lew, R. & De Schryver, G. M. (2014). Dictionary Users in the Digital Revolution. *International Journal of Lexicography*, 27(4), 341-359. <https://doi.org/10.1093/ijl/ecu011>
- Lew, R. & Doroszewska, J. (2009). Electronic Dictionary Entries with Animated Pictures: Lookup Preferences and Word Retention. *International Journal of Lexicography*, 22, 239-257. <https://doi.org/10.1093/ijl/ecp022>
- Liu, Y., Zhao, R., Altinger, L., Schütze, H. & Hedderich, M. A. (2025). Evaluating Robustness of Large Language Models Against Multilingual Typographical Errors. *arXiv*, 2510.09536. <https://arxiv.org/abs/2510.09536v2>
- Logan, F. A. (1999). Errors in Copy Typewriting. *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1760-1773. <https://doi.org/10.1037//0096-1523.25.6.1760>
- Lonke, D., Kernerman, I. & Dzhuranyuk, V. (2022). Lexical Data API. En A. Klosa-Kückelhaus, S. Engelberg, Ch. Möhrs & P. Storzjohann (eds.), *Dictionaries and Society: Proceedings of the XX EURALEX International Congress, Mannheim, Germany* (pp. 401-408). IDS-Verlag.
- Lorentzen, H. & Trap-Jensen, L. (2016). What, When and How? – the Art of Updating an Online Dictionary. En T. Margalitadze & G. Meladze (eds.), *Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity* (pp. 138-145). Tbilisi University Press.
- Peterson, J. L. (1986). A Note on Undetected Typing Errors. *Communications of the ACM*, 29(7), 633-637. <https://doi.org/10.1145/6138.6146>
- Ramírez Sánchez, I. (2023). Anatomía del diccionario: Análisis descriptivo y propuesta terminológica de las estructuras de las obras lexicográficas. *Boletín de la Real Academia Española (BRAE)* 103(328), 841-861. <https://revistas.rae.es/brae/article/view/622>
- Ren, X. & Perrault, F. (1992). The Typology of Unknown Words: An Experimental Study of Two Corpora. En Ch. Boitet (ed.), *COLING '92 Proceedings of the 14th Conference on Computational Linguistics, Vol. 1* (pp. 408-414). GETA (IMAG) & Association Champollion. <https://doi.org/10.3115/992066.992132>
- Rodrigues, P., Zajic, D., Doermann, D., Bloodgood, M. & Ye, P. (2011). Detecting Structural Irregularity in Electronic Dictionaries Using Language Modeling. En I. Kosem & K. Kosem (coords.), *New Applications for New Users: Proceedings of eLex 2011* (pp. 227-232).
- Rodríguez-Rubio, S. & Fernández-Quesada, N. (2020a). The Dynamics of Typographical Error Reproduction: Optimising Formal Correctness in Three Specialised Bilingual Dictionaries. *ELIA*, 20, 147-190. <https://doi.org/10.12795/elia.2020.i20.06>

- Rodríguez-Rubio, S. & Fernández-Quesada, N. (2020b). Towards Accuracy: A Model for the Analysis of Typographical Errors in Specialised Bilingual Dictionaries. Two Case Studies. *Lexikos*, 30, 386-415. <https://doi.org/10.5788/30-1-1606>
- Rodríguez-Rubio, S. & Fernández-Quesada, N. (2021). El tratamiento de las erratas desde la sicolingüística y desde el procesamiento del lenguaje natural. Un estado de la cuestión. *Tonos Digital*, 40, 1-19.
- Rodríguez-Rubio Mediavilla, S. (2021). *Estudio comparativo de erratas en diccionarios especializados inglés-español en papel de los ámbitos económico-financiero, jurídico e industrial. Una propuesta metodológica* [Tesis Doctoral no publicada]. Universidad Pablo de Olavide.
- Rosenblum, B. & Golfman, I. (2004). Automated Quality Assurance for Heuristic-Based XML Creation Systems. *Proceedings of the Extreme Markup Languages® 2004 Conference, Montréal, Quebec, Canada* (pp. 1-11). <https://www.semanticscholar.org/paper/Automated-Quality-Assurance-for-Heuristic-Based-XML-Rosenblum-Golfman/91948d5b64aee5a91f008249e74066f20ff27b22>
- Samanta, P. & Chaudhuri, B. B. (2013). A simple real-word error detection and correction using local word bigram and trigram. En *Proceedings of the 25th Conference on Computational Linguistics and Speech Processing (ROCLING 2013), Kaohsiung, Taiwan* (pp. 211-220). Association for Computational Linguistics and Chinese Language Processing (ACLCLP). <https://aclanthology.org/O13-1022/>
- Sharma, S. & Gupta, S. (2015). A correction model for real-word errors. *Procedia Computer Science*, 70, 99-106. <https://doi.org/10.1016/j.procs.2015.10.047>
- Simões, A. & Salgado, A. (2022). Smart Dictionary Editing with LeXmart. En A. Klosa-Kückelhaus, S. Engelberg, Ch. Möhrs & P. Storjohann (eds.), *Dictionaries and Society: Proceedings of the XX EURALEX International Congress, Mannheim, Germany* (pp. 423-434). IDS-Verlag.
- Snyder, K. M. & Logan, G. D. (2014). The Problem of Serial Order in Skilled Typing. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1697-1717. <https://doi.org/10.1037/a0037199>
- Svensén, B. (2009). *A Handbook of Lexicography*. Cambridge University Press.
- Swanepoel, P. (2001). Dictionary Quality and Dictionary Design: A Methodology for Improving the Functional Quality of Dictionaries. *Lexikos*, 11, 160-190. <https://doi.org/10.5788/11-0-846>
- Wheatley, H. B. (1893). *Literary Blunders: A Chapter in the "History of Human Error"*. Elliot Stock.
- Zajic, D., Maxwell, M., Doermann, D., Rodrigues, P. & Bloodgood, M. (2011). Correcting Errors in Digital Lexicographic Resources Using a Dictionary Manipulation Language. En I. Kosem & K. Kosem (coords.), *New Applications for New Users: Proceedings of eLex 2011* (pp. 297-301). Institute for Applied Slovene Studies.

ANEXO 1

Muestra de anomalías de las partes English-Spanish y Spanish-English del *Cambridge Dictionary Online*.

ENGLISH-SPANISH

Dice	Debería decir	Ubicación	Observaciones	Enlace
MAYÚSCULA				
Bomba Atómica	bomba atómica	Equivalente		https://dictionary.cambridge.org/dictionary/english-spanish/a-bomb
Lluvia Acida	lluvia ácida	Equivalente		https://dictionary.cambridge.org/dictionary/english-spanish/acid-rain
SIGNOS DE PUNTUACIÓN				
Es una pianista consumada., Es una pintora consumada., Es una experta jinete.	Es una pianista consumada. Es una pintora consumada. Es una experta jinete.	Ejemplo	Sobra la coma	https://dictionary.cambridge.org/dictionary/english-spanish/accomplished
acumular, acumularse acrecentar	acumular, acumularse, acrecentar	Equivalente	Falta la coma	https://dictionary.cambridge.org/dictionary/english-spanish/accrete
DISEÑO				
ablutions <i>phrase</i> ablutions <i>phrase</i>	ablutions <i>phrase</i>	Lema	Repetición de lema	https://dictionary.cambridge.org/dictionary/english-spanish/ablutions
ERRATA DE PALABRA (EP)				
El juez despejó la sala de todos menos ella misma y los testigos	La jueza despejó la sala de todos menos ella misma y el testigo	Ejemplo	Sustitución. En la frase inglesa original, se indica que es una jueza y se usa la palabra <i>witness</i> (no <i>witnesses</i>)	https://dictionary.cambridge.org/dictionary/english-spanish/all
Aunque que	aunque	Equivalente	Adición-repetición. Aparente perseveración del fragmento “que”	https://dictionary.cambridge.org/dictionary/english-spanish/albeit

ERRATA DE NO PALABRA (ENP)				
¡te sigo quie- rendo!	¡te sigo que- riendo!	Ejemplo	Transposición. Se han inter- cambiado los fragmentos “e” e “ie” de la palabra “que- riendo”	https://dictionary.cambridge.org/dictionary/english-spanish/as-bald-as-a-coot
Cabio* Dras- tico	cambio drás- tico	Equivalente	Sustracción.	https://dictionary.cambridge.org/dictionary/english-spanish/about-face
ACENTO ORTOGRÁFICO				
Agnostico	agnóstico	Equivalente	Esta anomalía aparece dos veces seguidas en la sección TRANSLATION of agnostic	https://dictionary.cambridge.org/dictionary/english-spanish/agnostic
Árbitro	árbitro	Equivalente	Esta anomalía aparece dos veces seguidas en la sección TRANSLATION of arbiter	https://dictionary.cambridge.org/dictionary/english-spanish/arbiter
ESPACIO				
por los viajes y las aventura s	por los viajes y las aventu- ras	Ejemplo	Sobra espacio entre la “a” y la “s”	https://dictionary.cambridge.org/dictionary/english-spanish/adventure
el sonido den- tal inicial [t]se pierde	el sonido den- tal inicial [t] se pierde	Ejemplo	Falta espacio entre “[t]” y “se”	https://dictionary.cambridge.org/dictionary/english-spanish/assimilate

SPANISH-ENGLISH

Dice	Debería decir	Ubicación	Observaciones	Enlace
DISEÑO				
on all four- son hands and knees	on all fours on hands and knees	Equiva- lente	Solapamiento de segmentos textuales	https://dictionary.cambridge.org/dictionary/spanish-english/a-cuatro-patas
on all four- son hands and knees	on all fours on hands and knees	Equiva- lente	Solapamiento de segmentos textuales	https://dictionary.cambridge.org/dictionary/spanish-english/a-gatas

ERRATA DE PALABRA (EP)				
<i>una casa de campo de*</i> <i>muy apacible</i>	<i>una casa de campo muy apacible</i>	Ejemplo	Adición-repetición. Aparente perseveración de “de”	https://dictionary.cambridge.org/dictionary/spanish-english/apacible
<i>E* madre asintió orgu- llosa</i>	<i>La madre asintió orgu- llosa</i>	Ejemplo	Sustitución. Discordancia de género	https://dictionary.cambridge.org/dictionary/spanish-english/asentir
ERRATA DE NO PALABRA (ENP)				
en una amplia e incontralada* abertura	en una amplia e incontrolada abertura	Ejemplo	Sustitución. Aparente anticipación de la letra “a”	https://dictionary.cambridge.org/dictionary/spanish-english/abertura
Se despide attenta- mente*	Se despide atentamente	Ejemplo	Adición-repetición. Posible interferencia de la voz inglesa <i>attentively</i>	https://dictionary.cambridge.org/dictionary/spanish-english/atentamente
SIGNOS DE PUNTUACIÓN				
Redactar un acta o documento oficial en por el que se certifica un suceso	Redactar un acta o documento oficial en/ por el que se certifica un suceso	Defini- ción	Falta barra inclinada entre “en” y “por”. También podría tratarse de una adición de “en” o de “por”	https://dictionary.cambridge.org/dictionary/spanish-english/acta
<i>¡Ab, qué bien!.</i>	<i>¡Ab, qué bien!</i>	Ejemplo	Sobra el punto tras el signo de exclamación de cierre	https://dictionary.cambridge.org/dictionary/spanish-english/ah
MAYÚSCULA				
Que está separado o es distinto	que está separado o es distinto	Defini- ción	Las definicio- nes del CDO empiezan con letra minúscula	https://dictionary.cambridge.org/dictionary/spanish-english/aparte
ACENTO ORTOGRÁFICO				
aliteracion <i>noun</i>	aliteración <i>noun</i>	Lema		https://dictionary.cambridge.org/dictionary/spanish-english/aliteracion



LEXICO-SEMANTIC CONSTRUCTION OF FEAR IN MAYA ANGELOU'S AUTOBIOGRAPHY OF GIRLHOOD

LA CONSTRUCCIÓN LÉXICO-SEMÁNTICA DEL MIEDO EN LA OBRA AUTOBIOGRÁFICA DE NIÑEZ DE MAYA ANGELOU

Blanca Fernández Soriano
Universidad Complutense de Madrid
blfern07@ucm.es

ABSTRACT

This article focuses on the lexico-semantic construction of fear in Maya Angelou's first autobiographical piece, *I Know Why the Caged Bird Sings* (1969). Through a lexicographic analysis and the application of the *FrameNet* framework, this study explores the emotion talk and semantic framing of fear. The analysis reveals that fear emerges as the most frequent emotion in the text, with a high degree of lexical variation and semantic nuance. Even though some general profile verbs, nouns, and adjectives are used—namely, *afraid*, *fear*, *scared* and *to fear*—, in Angelou's text, a medium and high degree of fear intensity comes up frequently through elements such as *dread*, *panic* or *terror*. It is suggested that fear is rarely framed as a momentary, instinctive reaction; rather, it is constructed as a pervasive emotional condition shaped by gender, race, and social contexts. Results show that the core elements composing the semantic frame of fear in Angelou's narrative must be present to verbalize the expression of fear and at least three of these are present in all instances of this text: [EXPERIENCER], [STATE] and [STIMULUS]/[TOPIC]. By integrating linguistic, cognitive, and socio-cultural perspectives, this study contributes to the expansion of the area of lexicology and cognitive lexicography of emotion, and foregrounds the complexity and cultural specificity of fear in Black autobiographical discourse of girlhood.

Keywords: emotion, fear, lexicology, semantic frame, Maya Angelou, black girlhood.

RESUMEN

Este artículo se centra en la construcción léxico-semántica del miedo en la primera obra autobiográfica de Maya Angelou, *I Know Why the Caged Bird Sings* (1969). A través de un análisis lexicográfico y de la aplicación del marco *FrameNet*, este estudio examina el discurso de emoción y el encuadre semántico del miedo. El análisis revela que el miedo es la emoción más frecuente en el texto, con un alto grado de variación léxica y matices semánticos. Aunque se emplean algunos sustantivos, verbos y adjetivos de perfil general (como *afraid*, *fear*, *scared* y *to fear*), en el texto de Angelou aparece con frecuencia un grado medio o alto de intensidad del miedo mediante elementos como *dread*, *panic* o *terror*. Se sugiere que el miedo rara vez se enmarca como una reacción momentánea o instintiva, sino que se construye como una condición emocional persistente, moldeada por el género, la raza y los contextos sociales. Los resultados muestran que los elementos nucleares que componen el marco semántico del miedo en la narrativa de Angelou deben estar presentes para verbalizar su expresión y que al menos tres de ellos aparecen en todos los ejemplos analizados: [EXPERIENCER], [STATE] y [STIMULUS]/[TOPIC]. Al integrar perspectivas lingüísticas, cognitivas y socioculturales, este estudio contribuye a la ampliación del campo de la lexicología y la lexicografía cognitiva de la emoción, y pone de relieve la complejidad y especificidad cultural del miedo en el discurso autobiográfico de la niñez negra femenina.

Palabras clave: emoción, miedo, lexicología, marco semántico, Maya Angelou, niñez Negra femenina.

Recibido: 13-08-2025
Aceptado: 05-03-2026

DOI: <https://doi.org/10.17561/rilex.9.2.9849>



1. INTRODUCTION

Autobiographical memory is the core of identity.
To a large extent, we are the stories we tell about ourselves.
ROBYN FIVUSH (2013)

Emotions are deeply embedded in language, shaping how individuals experience, express, and conceptualize their inner thoughts. Particularly in autobiographical texts, language of emotion becomes a key site through which personal and cultural identities are negotiated. Even though there has been an increasing interest in the description of emotion from linguistics (Ortony, Clore & Collins, 1988; Kövecses, 1990, 2000, 2015; Wierzbicka, 1999, 2009; Peräkylä & Sorjonen, 2012; Tash, Kolesnikova, Ahani & Sidorov, 2024), autobiographical discourse has been ultimately neglected as a valid source for emotion analysis, especially in relation to life-writing of Black girlhood (hooks, 1989; Douglas, 2022). Existing literature on emotion and language highlights the continuing challenge that the analysis of emotional expression poses to scholarship due to its complex and subjective nature, as well as its core presence in human experience.

Among the most salient linguistic methodologies that have contributed to this area are Kövecses's work on conceptual metaphors of emotion (1986, 1990, 2000, 2002, 2015), Wierzbicka's emotion primitives through Natural Semantic Metalanguage (1990, 1992, 1999, 2009), or the evaluation systems proposed by Systemic Functional Linguistics (Martin & White, 2005; Alba-Juez & Thompson, 2014; Alba-Juez & Mackenzie, 2016, 2019). Some approaches have developed definition structures for emotion terms, such as Ostermann's cognitive lexicography (2012, 2015), which draws mainly on metonymical and metaphorical language from a cognitive linguistic approach to explain the way we express and structure our thinking and feeling. Yet, one of the most relevant issues regarding the relationship between language and emotion is the following: emotional experiences cannot be reduced to their linguistic labels. Social, cultural, and individual variables affect the meaning of

emotion terms, misleadingly taken at times as synonyms or equivalents in different languages.

Most previous work on emotion has been based on corpora, such as the BNC Consortium (2007), COCA (Davies, 2008) or GloWbE (Davies, 2013), among others, which have been of vital importance for the development of cognitive and lexico-semantic studies (Bednarek, 2009; Güldenring, 2017, 2020; Maternová, 2024). Nonetheless, this paper advocates for the value of autobiographical texts as a relevant corpus for the linguistic study of the construction of emotion. This sort of discourse allows for the observation of the cultural, economic, political, and individual elements that influence the perception and expression of emotion. At the same time, autobiographical memories are influenced by emotions, as they impact the encoding and retrieval processes. In this sense, for instance, Wardell, Madan, Jameson, Cocquyt, Checknita, Liu and Palombo researched on how the process of verbalizing emotions affects the actual experience of emotion (2021). As D'Argembeau, Comblain, and Van der Linden (2003) observed, emotional memories seem to receive more attention, be better remembered, and more frequently reactivated. Some authors have worked on emotional expression through memoir, like Holland and Kensinger's research on memory and autobiographical writing (2010) or Besemeres's work on bilingualism (2010).

In view of this, the present study attempts to bridge some of these gaps which concern the study of autobiographical discourse as a valid source for linguistic analysis of early emotional experience with a focus on Black girlhood. Considering the expressions and concepts which people use to talk and understand their emotions as "language- and culture-specific" (Wierzbicka, 2009, p. 4), this work proposes a lexicographic and contextual semantic analysis of the emotion talk of fear in *I Know Why the Caged Bird Sings* (1969), the first memoir by African American author Maya Angelou, in order to observe the intertwining of personal experience, appraisal, and context in the expression of emotion for a Black girl in the 1960s United States of America.

2. FEAR IN LANGUAGE

Parting from the complexity and multidimension of emotion, fear is generally considered one of the basic human emotions (Ekman, 1992a, 1992b, 2004), as well as primitive and evolutively necessary (Van der Kolk, 2014). Several disciplines have paid great attention to its study, particularly from the perspectives of biology (Daniel-Watanabe & Fletcher, 2021; Fanselow & Hoffman, 2024; Landová, Polák, Janovcová, Štolhoferová, Peterková, Chomik & Frynta, 2025) and psychology (Öhman & Mineka, 2001; LeDoux, 2012, 2014; Chen, Hussein, Nasri, Yao, Qin, Zhao & Zuo, 2024). Research has shown that it is a fundamental emotion which has been shaped by evolution to protect the individual and respond effectively to potential risks, such as natural disasters (Prokop, 2016), dangerous animals (LoBue & DeLoache, 2008), or some environmental factors, like darkness or heights (Peléšková, Polák, Janovcová, Chomik, Sedláčková, Frynta & Landová, 2024). Overall, it seems to be cognitively fixed with a negative connotation, a human response to the uncertain, the dangerous, or the unpredictable (Miroshnychenko, 2020).

As introduced, one linguistic framework stands out in the frame of cognitive semantics, which considers the role of culture and thought to comprehend the expression(s) of emotion. Wierzbicka has been recognized as one of the major contributors to emotion theory: her work emphasizes the significance of language and cross-cultural comparison in order to provide insights in the underlying universals of human cognition (Wierzbicka, 1999). One of her most influential theories is Natural Semantic Metalanguage (NSM), a linguistic framework that seeks to identify the core, culture-independent meanings that are thought to inhabit the vocabulary of all languages. NSM is rooted in the idea that language reflects and shapes the way we think about the world. Furthermore, NSM considers the notion of “prototypical” emotional events, which refers to the basic components of an emotional experience shared cross-culturally. Using these prototype scenarios, Wierzbicka (1992,

p. 539) suggests that emotion can be detected in terms of thought, want, and feeling. By identifying these prototypical components, the underlying universals of emotional experiences might be better understood and empirically observed, as well as the ways in which they are expressed and conceptualized in different cultures.

Following Wierzbicka's semantic theory, fear is about something bad happening (either to us or in general) (1999, p. 363). It is differentiated from anger or sadness, for the bad thing is predicted to happen or will probably happen in the future. Fear is also accompanied by a feeling of weakness or helplessness. Accordingly, the basic concept of fear covers three criteria: 1) something bad will/can happen (to me); 2) I don't want this; and 3) I can't prevent it (or: I don't know what I can do about it). She describes the prototype scenario of fear (noun and verb) as follows:

X feels something (when X thinks of Y)
sometimes a person thinks something like this:
 I don't know what will happen
 something very bad can happen
 I don't want this
 because of this, I would want to do something
 I don't know if I can do anything
because of this, this person feels something bad
X feels like this

On the other hand, with the aim of proposing a methodology for emotion analysis, which disentangles its linguistic complexity, Bednarek distinguishes between "emotion talk" and "emotional talk" (2008, 2009). Emotion talk constitutes the expressions that "directly and explicitly name a particular emotional response" (2009, p. 396), whereas emotional talk concerns verbal or non-verbal expressions that are typically used to indirectly signal some kind of emotional state. As emotional talk involves conceptual language, expressed mainly by metaphors and metonymies, which will be dealt with in further analyses, the focus of this paper will be emotion talk, which includes the lexical realisations of fear.

When it comes to the domain of fear, emotion talk in English covers a number of lexical items such as *afraid*, *dread*, *fear*, *terrified* or *terror*. Lazarus (1991) claims that fear is the name of a family of related emotions, which includes, for instance, anxiety fright, or horror. Thus, a diverse range of lexical units are expected to be found that form the complex picture of fear in this memoir. Figure 1 schematizes the main lexical items belonging to the semantic field of fear emotion talk according to the dictionary *Merriam Webster* (retrieved from Zhgun, 2019). With the notion *fear* as centre, the first periphery contains its closest analogues, whereas the outer periphery includes words usually observed in association with this domain.

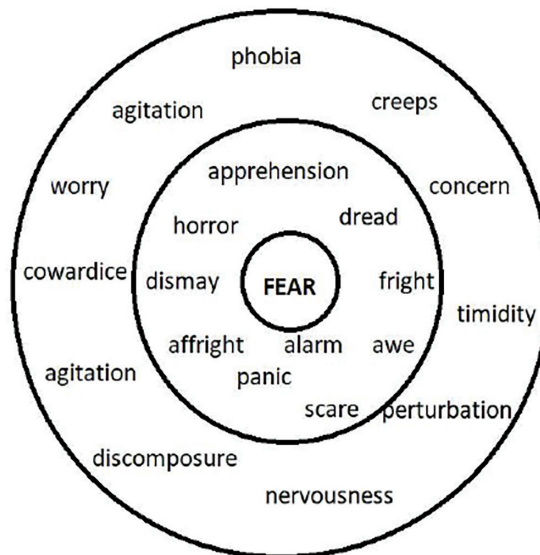


FIGURE 1: *Semantic field of the domain of fear (Zhgun, 2019, p. 248)*

In this sense, cognitive lexicography, while rather a young branch in linguistics, has proven useful and fruitful in the study of emotion terms. Within this framework, the *FrameNet* project (Ruppenhofer, Ellsworth, Petruck, Johnson, Baker & Scheffczyk, 2016) proposes a ‘feeling’-frame in order to describe the components of emotions. It theorizes an emotion or emotional state, an experiencer, and an evaluation of the emotional state for a certain emotion. Following this foundation, Ostermann defines fear cognitively as “the bad

and threatening feeling slowly coming up inside you when you think something bad is going to happen and you are in danger" (2012, p. 496). For instance, when you are by yourself in the dark, fear may make your heart beat faster, make you shiver or tremble, or want to run away from the situation.

As it will be further developed in Section 3, fear is composed by certain core elements: an experiencer (i.e., the person or sentient entity that experiences or feels the emotions), an expressor (i.e., the body part, gesture, or other expression of the experiencer that reflects their emotional state, denoted by an adjective or noun), a state (i.e., the abstract noun that describes a more lasting experience by the experiencer), a stimulus (i.e., the person, event, or state of affairs that evokes the emotional response in the experiencer), and the topic (i.e., the general area in which the emotion occurs, which indicates a range of possible stimulus). Non-core elements of fear include circumstances, degree, empathy target, explanation, manner, parameter, and time.

As it is suggested, the classification and codification of fear is not as simple as it might seem since it may take different psychological, physiological, social, and linguistic realizations depending on the context and the individual, as well as their background, culture, and language. Hence, it is relevant to carry out analyses that take these variables into consideration to attain a comprehensive understanding of the phenomenon of fear.

3. METHODOLOGY

Bearing in mind the aforementioned, this paper focused on Maya Angelou's first memoir *I know why the caged bird sings* (1969), a text of approximately 78,000 words tracing Maya's¹ childhood from age three through seventeen, from her shy beginnings in the racially segregated South of the United States of America to the birth of her first son in San Francisco. Through a bottom-up

¹ This paper distinguishes two recognizable voices, because of the memoiristic nature of the text: Maya, which will be used to refer to the young girl in the narration, and Angelou, the retrospective voice of the author.

analysis of emotion of the text, all the instances of fear were extracted. An instance refers to a clause, sentence or segment where fear is, either explicitly expressed or inferable from context, the main emotion of a character. This means that, within the same segment, fear can be attributed to different characters. As suggested by previous research, fear is one of the most frequent emotions in childhood (Tremewan & Strongman, 1991; Mash & Barkley, 2014) and it was also the case in this text, where fear instantiated on 75 occasions.

As discussed in Section 2, this paper distinguished between *emotion talk* and *emotional talk* for its analysis, even though the focus was solely on emotion talk. Thus, this study comprised mainly a manual qualitative analysis of all the samples that express fear in the text, collecting the lexical realisations of the domain of fear. Due to the written nature of the object of study of this project, only verbal expressions of fear were included, though they may describe non-verbal reactions of the individuals (e.g., “The certainty made my knees weak”, p. 235). As a general definition of the notion of *fear* is not enough to contain its complete meaning, which involves personal experience, appraisal, and context in which it is expressed, a lexicographic and contextual semantic analysis was needed, which allowed to describe the lexico-semantic clusters of fear in this text (Kolshansky, 2007). All the lexicalisations of emotion talk of fear were collected, defined according to the *Oxford English Dictionary* (from now on, *OED*) (2025) and the *Merriam Webster* (from now on, *MW*) (2025), and semantically analysed in interrelation and based on the influence of words in context.

Besides, the *FrameNet* cognitive lexicographic framework was used to determine the core and non-core elements of fear in this context. *FrameNet* suggests that fear is composed by the following core elements and non-core elements, collected in Table 1 (definitions retrieved from <https://framenet.icsi.berkeley.edu/frameIndex>).

Out of these components, this paper did not include the notions of [PARAMETER] and [CIRCUMSTANCES], since they were considered unclear in

Core elements	Experiencer	The person or sentient entity that experiences or feels the emotions
	Expressor	The body part, gesture, or other expression of the Experiencer that reflects their emotional state, denoted by an adjective or noun
	State	The abstract noun that describes the experience by the Experiencer
	Stimulus	The person, event, or state of affairs that evokes the emotional response in the Experiencer
	Topic	The general area in which the emotion occurs, which indicates a range of possible Stimulus
Non-core elements	Circumstances	The condition(s) under which the Stimulus evokes its response. In some cases it may appear without an explicit Stimulus.
	Degree	The extent to which the Experiencer's emotion deviates from the norm for the emotion.
	Empathy target	The individual or individuals with which the Experiencer identifies emotionally and thus shares their emotional response.
	Explanation	The explanation for why the Stimulus evokes a certain emotional response.
	Manner	Any description of the way in which the Experiencer experiences the Stimulus which is not covered by more specific Frame Elements, including secondary effects and general descriptions comparing events. Manner may also describe a state of the Experiencer that affects the details of the emotional experience.
	Parameter	A domain in which the Experiencer experiences the Stimulus.
	Time	The Time when the Experiencer, Expressor, or State can be described as having said emotion.

TABLE 1: *Core and non-core elements of fear*

their definitions and examples, and were deemed to overlap with the [TOPIC] element, which already includes the context and environment in which the emotion takes place and provides a range of possible stimuli for fear. All the remaining elements were annotated and quantified to analyse general trends and differentiated in two different categories: Maya's own fear and someone else's fear.

A model of analysis was needed for this kind of study which covered all the mentioned elements that compose the domain of fear. On the basis of all previous academic research regarding fear, Table 2 shows the model of analysis that this paper originally proposes to analyse in depth an emotion, particularly in this case, fear.

Linguistic expression of fear	Fear as a SEMANTIC FRAME		Core elements	
			Non-core elements	
	The language of fear	EMOTION TALK	Fear Vocabulary	Adjectives
				Adverbs
				Nouns
				Verbs
	EMOTIONAL TALK			

TABLE 2: *Model of linguistic analysis of fear*

4. RESULTS AND DISCUSSION

4.1. EMOTION TALK

Two different experiences of fear can be distinguished in this text: one's own fear (Maya's fear) and someone else's fear (e.g., Bailey, the Black kids, etc.). The following figures (Figures 2 and 3) schematize the main lexical items that form the emotion talk of FEAR in this text divided in the described perspectives and organized in terms of frequency.

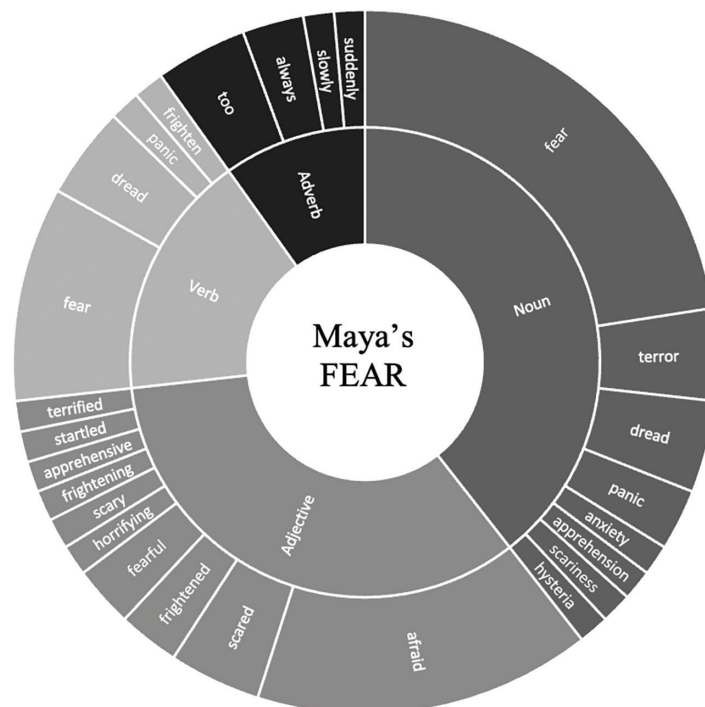


FIGURE 2: *Semantic field of the domain of Maya's FEAR*

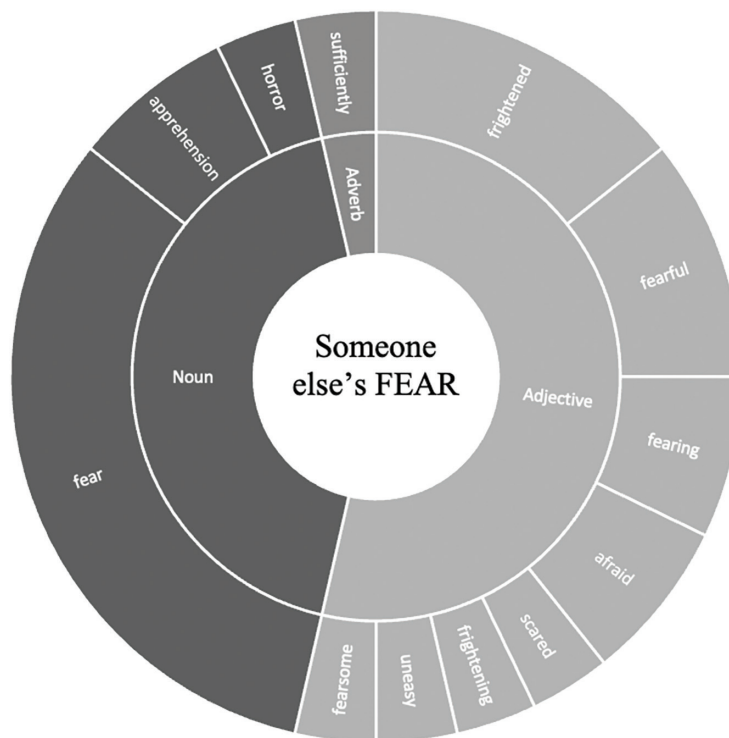


FIGURE 3: *Semantic field of the domain of someone else's FEAR*

Expectedly, the instances of the emotion talk of Maya's fear (N=71) are higher than those of someone else's fear (N=28), and the lexical units referring to her own emotions are also more varied in all grammatical categories, probably due to the focus of the memoir on her inner world and her emotional journey.

4.1.1. Nouns

In both cases, the noun *fear* leads the emotion talk of fear (16 occurrences for Maya's fear and 9 for someone else's). As anticipated, this noun takes varied meanings and uses. According to the *OED*, *fear* can be used as "the emotion of pain or uneasiness caused by the sense of impending danger, or by the prospect of some possible evil" (2025, n., sense 2.a), i.e., a (negative) emotional experience, as illustrated by (1). This emotion can be personified as in extract (2), or display a cumulative instance of it, like in (3). Besides, it can also be viewed "with regard to an object" (*OED*, 2025, n., sense 3), such as Maya and the other kids' fear of death that is implanted "under their skins" by the priest (4).

- (1) Momma put her hand on my arm and said, "You shaking, Sister. What's the matter?" My skin still rippled from the experience of *fear*. (p. 166)
- (2) An awkward *fear* crept up slowly as I contemplated sitting in the car all night alone. (p. 237)
- (3) My *fears* were so powerful I couldn't move to look at the center of the bed. (p. 289)
- (4) After he had put the *fear* of the cold grave under our skins, he began to speak of Mrs. Taylor. (p. 161)

Other nouns are used in this domain to refer to different intensities of the emotion, such as *dread*, *panic* or *terror*. These items are usually referred to as synonyms of *fear* in English; nonetheless, semantic nuances may be observed between them, in their definitions as well as in the contexts of the memoir. According to the *OED*, *panic* is a "sudden feeling of alarm or fear of sufficient intensity or uncontrollableness as to lead to extravagant or wildly unthinking behaviour" (2025, n., sense 1.a), as exemplified by (5) where Maya's control is almost lost, and a higher degree of intensity is involved.

- (5) A fog of *panic* nearly suffocated me. (p. 235)

Next, *dread* adds the idea of "extreme fear; deep awe or reverence; apprehension or anxiety as to future events" (*OED*, 2025, n., sense 1). Instance (6) also displays a higher degree of intensity, as well as a specific fear of facing a certain situation, in this case, the futility of life.

- (6) The *dread* of futility has been my lifelong plague. (p. 248)

Then, *terror* is defined to be "the state of being terrified or extremely frightened; intense fear or dread" (*OED*, 2025, n., sense 1) and "the most extreme degree of fear" (*MW*, 2026, n.). This degree of fear is encountered several times in this work. Example (7) displays the gradience of fear increasing until it reaches a natural force-like terror that swallows and paralyzes Maya.

- (7) It was a *fear* distantly related to the earlier *panic*. *Terror* did not engulf me wholly, but crawled along my mind like a tedious paralysis. (p. 235)

Another semantic nuance is added when nouns such as *anxiety*, *apprehension* or *hysteria* are used. First, *anxiety* is defined by the *OED* to convey a “worry over the future or about something with an uncertain outcome” (2025, n., sense 1.a) or a “pathological state characterized by inappropriate or excessive apprehension or fear, which may be generalized or attached to particular situations” (*OED*, 2025, n., sense 4). Example (8) displays an anxious Maya, scared for the possibility of a bombing and being relaxed by her grandmother.

- (8) I ran all the way home. Not too sure I wouldn't be bombed before I reached Bailey and Mother. Grandmother Baxter calmed my *anxiety* by explaining that America would not be bombed. (p. 208)

These items are related to nervousness, especially of future events. Likewise, *apprehension* is defined as “fear as to what may happen” (*OED*, 2025, n., sense II.12) or “suspicion or fear especially of future evil” (*MW*, 2026, n., sense 1), and considered here a fear of mild intensity. In instances (9) and (10), young Maya is able to grasp the *apprehension* of others (Mamma and the other Black kids, respectively) from their non-verbal language, which later turns in both cases into an explicit fear.

- (9) Her *apprehension* was evident in the hurried movements around the kitchen and in her lonely *fearing* eyes. (p. 114)
- (10) Their faces, which showed *apprehension* before reaching the coffin, revealed, on the way down the opposite aisle, a final confirmation of their *fears*. (p. 161)

Hysteria is one of the most culturally-loaded items, for it was believed originally to be a disease unique to women and caused by disfunction of the uterus (see Paulon, 2022). It is currently understood as an “overwhelming, uncontrollable emotion or agitation” (*OED*, 2025, n., sense 1) or “behaviour

exhibiting overwhelming or unmanageable fear or emotional excess" (MW, 2026, n., sense 2). In example (11), *hysteria* is a high-intensity emotion of fear that takes over Maya's body and results in uncontrolled tears.

- (11) Not a damn, I decided, and opened the flood gates for *hysteria*.

Once the tears began, there was no stopping them. (p. 237)

Next, the case of *scariness* is a remarkable one, since it does not convey the temporary state of feeling scared, like the previous examples, but it describes the "quality of being scary" (MW, 2026, n.). This meaning is closer to a personality trait, as example (12) implies, in which it mixes with a vivid imagination characteristic of children that tends to provide scary but imaginary scenes.

- (12) It is hard to understand why my vivid imagination and tendency toward *scariness* didn't provide me with gory scenes of bloody crashes on a risco de [sic] Mexico. (p. 238)

Finally, *horror* is one of the highest-intensity degrees of fear and it only appears once in the memoir describing another's fear, Bailey's fear (13). It is defined as a "painful emotion compounded of loathing and fear; a shuddering with terror and repugnance; strong aversion mingled with dread" (OED, 2025, n., sense 3.a). Indeed, this item is used in a quite extreme situation, after Bailey carried the corpse of a Black man and the horror sticks with him, like a burden.

- (13) Bailey couldn't let go of the *horror*. "I picked up a side of the sheet and walked right in the calaboose with the men. I walked in the calaboose carrying a rotten dead Negro". His voice was ancient with shock. He was literally bug-eyed. (p. 198)

4.1.2. Adjectives

Adjectives seem to be the richest grammatical form to compose the emotion talk of FEAR in this memoir, expressing different degrees of intensities as well as agency. The most frequent adjective to express one's fear is *afraid* (11 instances for Maya's fear and 2 for someone else's fear), which is defined

as being “alarmed, frightened; in a state of fear or apprehension, moved or actuated by fear” (*OED*, 2025, adj., sense 1.a) and considered of medium intensity. Like “fear”, this adjective can be used to refer to a situation, be it real or hypothetical (14), object or person (15), or afraid to perform a certain action (or not to perform it) (16).

- (14) I never had the nerve to go up to him. I was quite *afraid* that if I tried to say, “Hello, Reverend Thomas,” I would choke on the sin of mocking him. (pp. 35-36)
- (15) I tried to staunch the flood of fear. Why was I *afraid* of the Mexicans? (p. 237)
- (16) Under the most common of circumstances I knew it did no good to argue with Momma. So I walked down the steep stairs, *afraid* to look back and *afraid* not to do so. (p. 189)

Other adjectives considered of medium intensity present in this work, like *afraid*, are *fearful*, *fearing*, *fearsome*, *scared*, *scary* and *startled*. Deriving from *fear*, *fearful*, *fearing* and *fearsome* present several semantic differences and uses. First, *fearful* is dual in its agency as it can describe the element that is “frightened, timorous, timid, apprehensive” (*OED*, 2025, adj., sense II.3.a), like Bailey and Maya in (17) or Maya’s secret in (18), or the agent that causes the fear, as the sounds of the chimney in (19).

- (17) We were both *fearful* of Mother’s coming and impatient at her delay. (p. 59)
- (18) That evening, in the bosom of the now-dear family home I uncoiled my *fearful* secret and in a brave gesture left a note on Daddy Clidell’s bed. (p. 288)
- (19) The chimney made *fearful* sounds of protest as it was invaded by the urgent gusts. (p. 153)

Next, *fearing* defines the element that fears, whether it is a person or an object. In example (20), *fearing* describes Momma’s eyes and, through her eyes, it describes her own fearful state. According to the *OED*, it is often used

“in combination with prefixed object” (2025, adj.), as in (21). Nonetheless, when it comes to God or religion, *fearing* can also describe a fear in relation to a “reverential awe” (MW, 2026, n., sense 2).

(20) Her apprehension was evident in the hurried movements around the kitchen and in her lonely *fearing* eyes. (p. 114)

(21) “Well, now, understand me, Sister Henderson, I don’t hold this against you, I knows you a *God-fearing* woman”. (p. 151)

On the other hand, *fearsome* can describe both the subject that fears and the one which inspires the fear (OED, 2025, adj., sense 1 and 2). This adjective only appears once in the memoir and describes a more permanent awe-inspiring personality trait in some of Maya’s family members that had been raised in fear and violence (22).

(22) With that kind of encouragement, backed by explosive tempers, it was no wonder they became *fearsome* characters. (p. 66)

Next, *scared* can suppose a more lasting trait (23) or a momentary state of fear (24), while *scary* describes the element (person, event, object) that causes the fear (25).

(23) Uncle Willie muttered, “They don’t really hate us. They don’t know us. How can they hate us? They mostly *scared*.” (p. 197)

(24) Uncle Willie laughed and said, Maybe she was *scared* to go in the kitchen. (p. 166)

(25) Surely she didn’t mean the long fork that hung on the wall behind the kitchen stove —a *scary* million miles away. (p. 165)

The item *startled* adds the idea of surprise, alarm, or shock to the fear (MW, 2026, v.). In instance (26), the moment of shock when she is sexually abused by Mr. Freeman incapacitates Maya to react and run away from the situation.

- (26) Whatever it was, I hadn't encountered the sensation in all the years of sleeping with Momma. It didn't move, and I was too *startled* to. (p. 72)

Moving on to high-intensity fear adjectives, the memoir comprises the items *frightened*, *frightening*, *horrifying* and *terrified*. To start with, *frightened* describes someone that is *put into* or *affect with fright* (OED, 2025, adj., sense a). Curiously enough, even though it is used in its adjectival and verbal derivatives, *fright* does not appear in the memoir as a noun. It is defined by the OED as a "sudden fear, violent terror or alarm" (2025, n., sense 1.b), presenting a high intensity with an idea of suddenness or violence. In many cases where this adjective is used in the text, it is related to childlike confusion and lack of understanding extreme situations, as illustrated by (27), (28) and (29).

- (27) Years later I discovered that the United States had been crossed thousands of times by *frightened* Black children traveling alone to their newly affluent parents in Northern cities. (p. 5)
- (28) Even Bailey was *frightened*. He sat all to himself, looking at a man's death —a kitten looking at a wolf. Not quite understanding it but *frightened* all the same. (p. 87)
- (29) I was confused and *frightened*. He was going to get a whipping and maybe he had done something terrible. (p. 116)

Contrastingly, *frightening* describes the element that frightens. In the analysed examples, the idea of childlike confusion or ignorance of the world (e.g., knowledge of death) seems to be maintained, as illustrated by (30) and (31). Similarly, *horrifying* characterizes something that causes to feel horror, another quite intense fear. In example (32), Maya suffers from nightmares created by her own childish imagination that mentally afflict her.

- (30) At eleven years old, death is more unreal than *frightening*. (p. 160)
- (31) It meant that he had seen or heard of something so ugly or *frightening* that he was paralyzed as a result. (pp. 195-96)

- (32) Because of the lurid tales we read and our vivid imaginations and, probably, memories of our brief but hectic lives, Bailey and I were afflicted—he physically and I mentally. He stuttered, and I sweated through *horrifying* nightmares. (p. 72)

Then, *terrified* defines an individual “seized with terror; affected or actuated by terror” (*OED*, 2025, adj., sense 1). Like *terror*, it expresses a rather high-level intensity of fear, which would explain why there is only one example of this adjective. In instance (33), Maya feels this sudden terror after realizing they were going to meet their mother.

- (33) Were we going to see Her? I thought we were going to California. I was suddenly *terrified*. (p. 58)

Finally, there are only two instances of mild-intensity fear adjectives. *Apprehensive* expresses the idea of fear or anxiety for an unknown future: “anticipative of something adverse; fearful of what may be about to happen” (*OED*, 2025, adj., sense 6.a). In example (34), *apprehensive* is used in relation to the stronger intensity of *afraid*.

- (34) He said, “Just stay right here, Ritie, I ain’t gonna hurt you”. I wasn’t afraid, a little *apprehensive*, maybe, but not afraid. (p. 72)

On the basis of this fearful anxiety, *uneasy* describes the state of being “uncomfortable or disturbed in mind; anxious, apprehensive” (*OED*, 2025, adj., sense 4.a), whether it is “physical or mental” discomfort (*MW*, 2026, adj., sense 1). Example (35) displays a Maya that detects in her uncle’s nervous laugh everyone’s discomfort to enter the unknown of the dark kitchen.

- (35) His high little laugh didn’t fool me. Everyone was *uneasy* at being beckoned into the unknown. (p. 166)

4.1.3. Verbs

Curiously enough, not many verbs seem to form the emotion talk of FEAR in this text. There are only 12 verbal utterances and they all refer to Maya’s fear. It comes as no surprise that “fear” is the most frequent verb to express fear (N=7). Considering all previous definitions of the item,

this verb is deemed as a medium intensity expression of fear. Among others, instance (36) is interesting in that it narrates from the perspective of time that young Maya was afraid of the potential consequences of a man she disliked marrying her grandmother, arguing that children are built for survival. Next, occurrence (37) displays the element of the potential unconsciousness of fear and example (38) refers to fear of the future and losing home.

- (36) Maybe due to the fact that I was in my romanticist period, or because children have a built-in survival apparatus, I *feared* he was interested in marrying Momma and moving in with us. (pp. 154-155)
- (37) Their eyes, language and customs belied the white skin and proved to their dark successors that since they didn't have to be *feared*, neither did they have to be considered. All this was decided unconsciously. (p. 211)
- (38) In those miles, along the twisted roads beside the steep mountain, I *feared* that I would never get back to America, civilization, English and wide streets again. (pp. 232-233)

In increasing order of intensity, other verbs that appear are *frighten* (39), *dread* (40), and *panic* (41). While *frighten* is performed by the group that causes the fear on Maya (in this case, some white children), *panic* is used in a causative structure where the darkness provokes in Maya the panic. Finally, carries Maya as the subject that suffers the fear.

- (39) I wanted to beg her, "Momma, don't wait for them. Come on inside with me. If they come in the Store, you go to the bedroom and let me wait on them. They only *frighten* me if you're around". (p. 30)
- (40) I *dreaded* laughing in church. If I lost control, two things were certain to happen. I would surely pee, and just as surely get a whipping. (p. 43)

- (41) One foot into the darkness and the sense of detachment from reality nearly made me *panic*. The idea came to me that I might never get out into the light again. (p. 166)

4.1.4. *Adverbs*

Although no adverbs refer directly to the expression of fear (e.g., *alarmedly*, *fearfully*, *hysterically*, *scaringly*, etc.), a few items are used in this text to modify the intensity or nuance of the expression of fear. Notably, most of these adverbs add a meaning of duration, either short or long, or intensity. The most frequent adverbial item is *too*, which expresses such a high intensity that stops Maya from reacting or performing the desired action, as examples (42), (43) and (44) exemplify.

- (42) It didn't move, and I was *too* startled to. (p. 72)
- (43) I was *too* afraid to splash water or even to cry and take a chance of drowning out Bailey's pleas for help, but the pleas never came and the whipping was finally over. (p. 117)
- (44) If Dad stayed a very long time in the house, I was *too* afraid to go to the door and ask for him, and besides, my feminine training would not allow me to walk two steps with blood on my dress. (p. 248)

Next, *sufficiently* is used to decrease the intensity of the fear to a level enough to not perform an action and be deemed polite by other people, as in (45).

- (45) Anyone, I reckoned, *sufficiently* afraid or sufficiently dull could be polite. (p. 173)

In terms of duration of the emotion, *always* increases the time of a generally temporary state into a more lasting trait, as examples (46) and (47) show, which may potentially become a personality attribute.

- (46) As I had *always* feared, no, known, the trials had been for nothing. (p. 248)
- (47) I was *always* afraid when I found him [her father] watching me, and wished I could grow small like Tiny Tim. (p. 55)

Suddenly adds the idea of unexpectedness, abruptness, or short time, in the building up of the emotion and as well as its duration (48).

(48) I was *suddenly* terrified. (p. 58)

Contrastingly, *slowly* describes a fear that is built up gradually at a slow pace and lasts longer in time (49).

(49) An awkward fear crept up *slowly* as I contemplated sitting in the car all night alone. (p. 237)

Briefly summarizing the emotion talk of fear in Angelou's narrative, the text is rich in lexical elements of fear, as synthesized in Table 3. General profile verbs, nouns and adjectives are used, illustrated by *afraid*, *fear*, *scared* and *to fear*. These words seem to neutrally refer to the emotion of fear, with no extra nuances. However, in Angelou a medium and high degree of fear intensity comes up frequently through elements such as *dread*, *panic* or *terror*. Fear is projected as something that may become a lasting trait or an emotion that may overcome the experiencer abruptly.

FEAR EMOTION TALK	INTENSITY		
	HIGH	MILD	LOW
VERBS	Dread Panic Frighten	Fear	
NOUNS	Dread Terror Panic Hysteria Anxiety Horror	Fear Scariness	Apprehension
ADJECTIVES	Terrified Frightened Frightening Horrifying Dreaded	Afraid Startled Scared Scary Fearful Fearing Fearsome	Apprehensive Uneasy
ADVERBS	Suddenly, always, too, slowly		

TABLE 3: *Intensity of Emotion talk of FEAR*

4.2. CORE AND NON-CORE ELEMENTS

As anticipated in Section 2, the framework of cognitive lexicography can be very useful to carry out an in-depth linguistic analysis of emotion, as this paper intends. On the basis of the *FrameNet* project, which collects the different core and non-core elements of a variety of concepts, this section analyses the linguistic components of all the instances of fear in this memoir.

First, [STATE] is the most present element in the linguistic realization of fear. In 97.33% of the instances (73 out of 75), there was at least one explicit verbalization of the state of the emotion, i.e., a noun, verb, or adjective which identified the emotion as fear. Only in two cases there is no explicit mention of the emotion, but it can be glimpsed in the text in metaphorical terms, which makes its interpretation more subjective. In example (50), Maya and Bailey enter the white part of town and fear is understood as opposite to pleasure and a dangerous expedition for which they are defenceless.

- (50) But the *pleasure fled* [Exr] when we reached *the white part of town* [Top]. After we left Mr. Willie Williams' Do Drop Inn, the last stop before whitefolksville, we had to cross the pond and adventure the railroad tracks. *We* [Exp] *were explorers walking without weapons into man-eating animals' territory* [Exr]. (p. 25)

Then, in instance (51), there is no explicit element of [STATE], but there is an [EXPRESSOR] which reflects the altered emotional state and a metaphorical explanation through which this state can be inferred. Fear is here expressed in terms of paralysis and lack of breath, though Maya does not relate it to herself, but the world which cannot move out of fear.

- (51) The *tears* [Exr] that had slipped down *my* [Expl] dress left unsurprising dark spots, and made the front yard blurry and even more unreal. *The world had taken a deep breath and was having doubts about continuing to revolve* [Exp]. (p. 31)

The [EXPERIENCER] is present too in 97.33% of the occurrences (73 of 75), since in most of the cases Angelou specifies whether it is herself or others

who feel the fear. The cases in which no explicit [EXPERIENCER] is encountered correspond to the descriptions of the situations that cause fear in general, stepping into the dark in example (52) and a potential bombing in instance (53). It could be argued that Maya is included in the people that experience fear in these situations, since she is describing them as such. Nonetheless, it is not made explicit in these expressions, which favour an impersonal structure.

(52) “Ma’am?” Surely she didn’t mean the long fork that hung on the wall behind the kitchen stove —*a scary* [Sta] *million miles away* [Stim]. (p. 165)

(53) *The undertone of fear* [Sta] *that San Francisco would be bombed* [Stim] which was *abetted by weekly air raid warnings, and civil defense drills in school* [Expl], heightened my sense of belonging. (p. 212)

Next, the [EXPRESSOR] element relates to the physiological reactions and behaviours of fear, that is, the observable components that express the emotion. Even though it is a core element through which the [STATE] can be inferred and may provide different descriptions of the intensity and other features of the emotion, it is not mandatory for it to appear. In general, the [EXPRESSOR] is only made explicit in 37.33% of the examples (28 out of 75). It appears more often in the cases of Maya’s fear (41.18% of her instances). This might be due to the fact that she is the main [EXPERIENCER] in this text and narrates and feels in first-person all the physiological effects of fear in her own body, whereas other people’s fear can generally be identified through external physiological effects or reactions, such as crying or physical paralysis.

Then, the [STIMULUS] or [TOPIC] are verbalized in every case in the expression of fear and only in one instance they appear, but are not specified. In general, in 65.33% of the cases a [STIMULUS] for the emotion is made explicit and in 66.67% of the occurrences a [TOPIC] is displayed, which gives an idea of potential Stimuli for the fear. No significant differences can be

observed between Maya's fear and someone else's fear, for Angelou knows and provides context for most of the instances. The only occurrence where neither of them is specified belongs to a case of Bailey's fear: the reason for his fear is not known by Maya and not verbalized by him either. In return, [EXPERIENCER], [EXPRESSOR], and [STATE] are specified for it to be identified as fear (54).

- (54) "What's the matter, Bailey Junior?" *He* [Exp] *said nothing* [Exr]. I knew when I saw him that it would be useless to ask anything while *he was in that state* [Sta]. It meant that he had seen or heard *of something so ugly or frightening* [Stim] that he was *paralyzed as a result* [Exr]. (pp. 195-196)

(While there are at least three core elements in all occurrences of fear (most frequently [EXPERIENCER], [STATE], and [STIMULUS]/[TOPIC]), non-core elements are not required to be specified and only provide the expression of emotion with extra information and nuances of time, manner, or intensity, among others. There are not many significant differences in the trends of non-core elements between Maya's fear and someone else's fear, except in terms of [EXPLANATION] and [MANNER].

As regards [EXPLANATION], this component provides an explicit reason for the [STIMULUS] to cause the fear and is verbalized more frequently in the case of Maya's fear (there are 13 instances of [EXPLANATION] for Maya's fear and 4 for someone else's fear). As in the case of the [EXPRESSOR] component, this might be due to her being the main Emoter of the memoir and being more aware of certain elements and contextual evidence. For instance, passage (55) displays Maya's stream of thought explaining why she was scared while she was being sexually abused by Mr. Freeman.

- (55) Then he dragged me on top of his chest with his left arm, and his right hand was moving so fast and his heart was beating so hard that *I* [Exp] *was afraid* [Sta] *that he would die* [Stim]. *Ghost stories revealed how people who died wouldn't let go of whatever*

*they were holding. I wondered if Mr. Freeman died holding me
how I would ever get free.* [Expl] (p. 73)

Then, [MANNER] adds the description of the way the emotion is experienced. It is specified in 5 instances of Maya's fear and only 1 of someone else's fear, similarly to the previous differences between these categories. This component is mainly performed by the adverbs covered in 4.1.4, such as *slowly* or *suddenly*, which modify the way that fear takes place or is felt. Another adverb comprises too the notion of [TIME] (which could be deemed as involving the [DURATION] of the fear), which appears only in three occasions (all for Maya's fear) and all refer to a longer period of time or even permanent trait: *always*. *Suddenly* could be considered as well to refer to a matter of time, since the emotion comes abruptly and may also be shorter in time.

As regards [DEGREE], this component is mainly defined too by adverbs which modify the intensity of fear. It is specified only on occasion, as there are only 4 occurrences for someone else's fear and 5 instances for Maya's own fear. Most of the cases of the element of [DEGREE] intend to enhance the intensity of the emotion, using adverbial phrases such as *so*, *that much* or *too*. Nonetheless, there are two instances which refer to different degrees of fear, such as "How can they hate us? They *mostly* scared" (p. 197), which refers to the mix of fear with another feeling (namely, hate), and "Anyone, I reckoned, *sufficiently* afraid or sufficiently dull could be polite" (p. 173), which lower the degree of fear to a level enough for people to consider it polite.

Finally, [EMPATHY TARGET] is an interesting category, since it makes reference to the emotional identification of the [EXPERIENCER] with another (feeling) individual with whom an emotional response can be shared. Although it could be argued whether this is a cross-cultural and universal component of fear, especially when it comes to children, five occurrences of this aspect were found in this text. Significantly enough, all the instances of this component in the memoir correspond to Maya being empathetic towards people close to her. This makes sense in that it is those people, namely

Bailey and her family, whom she feels closer to, knows best, and can, thus, relate emotionally. Out of the five instances, four of them display Bailey as the [EMPATHY TARGET]. This may be due to the fact that they are siblings –thus, raised together and similarly–, and close in age, so they share similar emotional responses, which Maya can identify as the corresponding emotion. This can be inferred from instances such as “He said nothing. I knew when I saw him that it would be useless to ask anything while he was in that state” (pp. 195-196).

5. CONCLUSION

The purpose of this study was to provide a comprehensive account of the linguistic construction of Black girlhood FEAR in Maya Angelou's first memoir, *I know why the caged bird sings*, understood as a socio-cultural, individual, and subjective phenomenon. Through a bottom-up approach analysis of the lexico-semantic components of the language of fear in the text, this paper has observed that fear constitutes a negative personal emotion and a disruptive force in the life of a Black girl living in 1960s USA. Fear is most frequently presented as a temporally restricted process, although there are occasions where fear is experienced as persistent, something closer to an emotional disposition. In Angelou's text, fear is commonly presented as triggered by a dangerous or unpleasant situation, such as violence, racial discrimination, social rejection, or sexual assault, accompanied by a set of physiological and behavioural reactions.

Lexically, fear is presented as a family of different realizations of one emotion: anxiety, fright, horror, panic, worry. The emotion talk of fear in this memoir is composed mainly by the noun *fear*, the verb *to fear* (only regarding one's own fear), and the adjective *afraid*. Nonetheless, fear takes many different lexicalizations which add different nuances of intensity (*terror*), agency (*frightening*) or urgency (*anxiety*). In all these samples, analysed lexicographically, fear is deemed a negative emotion for the Emoter. Besides, adjectives seem to be the most productive grammatical form to directly

express fear. Even though the adverbs in the corpus do not convey fear by themselves, they also add semantic nuances to the expression of fear in terms of its duration (*always*), intensity (*too*), or manner (*suddenly*).

Using the cognitive lexicography *FrameNet* project, the core and non-core elements of fear in this memoir were retrieved to observe which were the most frequent and necessary elements in the expression of fear. It is noteworthy that, on the one hand, core elements, as defined by the *FrameNet* project, must be present to verbalize the expression of fear and at least three of these are present in all examples of this text ([EXPERIENCER], [STATE], and [STIMULUS]/[TOPIC]). On the other, non-core elements, although not mandatory, help to define and provide with different nuances the expression of the emotion, like its intensity, its manner, or duration. Through these components, a different approach is attained which manages to break down all the elements that compose this emotion and uncover certain socio-cultural elements, such as the stereotypical duration of fear, the perception of someone else's fear, or the identification with the emotional responses of other individuals.

This study was limited by the lack of information on the linguistic expression of emotion as an interrelated system of lexical, semantic, cognitive, and socio-cultural elements, but it attempted to originally propose a model that agglutinates all previous knowledge of these areas.

Notwithstanding these limitations, this can be a fruitful area for further work from the perspective of linguistics. A natural progression of this work would be to analyse other emotions, perhaps primary emotions which are less culturally variable (like joy or anger) or secondary emotions, which could offer larger socio-cultural variations (such as shame). It would also be interesting to conduct this study in other languages, as most emotion research in linguistics has focused so far on English, and to use other memoirs and autobiographies to expand our knowledge of life narration. Another question raised by this research could be the comparison between young Black and

white girls, other sociocultural groups (like Latino women), and the potential different perception of fear between Black children and adults.

REFERENCES

- Alba-Juez, L. & Mackenzie, J. (2016). *Pragmatics: Cognition, Context and Culture*. McGraw Hill.
- Alba-Juez, L. & Mackenzie, J. (eds.) (2019). *Emotion in Discourse*. Benjamins Publishing Company.
- Alba-Juez, L. & Thompson, G. (2014). *Evaluation in Context*. John Benjamins.
- Angelou, M. (1969). *I Know Why the Caged Bird Sings*. Random House.
- Bednarek, M. (2008). *Emotion Talk across Corpora*. Palgrave Macmillan.
- Bednarek, M. (2009). Emotion talk and emotional talk: Cognitive and discursive perspectives. In H. Pishwa (ed.), *Language and Social Cognition: Expression of the Social Mind* (pp. 395-431). Mouton De Gruyter.
- Besemeres, M. (2010). Emotions in bilingual life narratives. In V. Cook & B. Bassetti (eds.), *Language and Bilingual Cognition* (pp. 479-506). Psychology Press.
- BNC Consortium. (2007). *The British National Corpus*. XML Edition, Oxford Text Archive, <http://hdl.handle.net/20.500.14106/2554>
- Chen, C., Hussein, S. Z. B., Nasri, N. W. M., Yao, J., Qin, Y., Zhao, Z. & Zuo, K. (2024). Fear of childbirth among pregnant women: A concept analysis. *Journal of Advanced Nursing*, 80(11), 4476-4487. <https://doi.org/10.1111/jan.16218>
- D'Argembeau, A., Comblain, C. & Van der Linden, M. (2003). Phenomenal characteristics of autobiographical memories for positive, negative, and neutral events. *Applied Cognitive Psychology*, 17(3), 281-294. <https://doi.org/10.1002/acp.856>
- Daniel-Watanabe, L. & Fletcher, P. C. (2021). Are Fear and Anxiety Truly Distinct? *Biological Psychiatry Global Open Science*, 2(4), 341-349. <https://doi.org/10.1016/j.bps-gos.2021.09.006>
- Davies, M. (2008). *The Corpus of Contemporary American English (COCA)*. <https://www.english-corpora.org/coca/>
- Davies, M. (2013). *Corpus of Global Web-Based English: 1.9 billion words. from speakers in 20 countries*. Available at: <http://corpus.byu.edu/glewbe/>.
- Douglas, K. (2022). *Autobiography and Childhood*. *Oxford Bibliographies*. Retrieved 26 June 2025. <https://n9.cl/riaiz>
- Ekman, P. (1992a). Facial expressions of emotion: New findings, new questions. *Psychological Science*, 3, 34-38.
- Ekman, P. (1992b). An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.
- Ekman, P. (2004). *Emotions Revealed*. Henry Holt.

- Fanselow, M. S. & Hoffman, A. N. (2024). Fear, defense, and emotion: A neuroethological understanding of the negative valence research domain criteria. *American Psychologist*, 79(5), 725–734. <https://doi.org/10.1037/amp0001354>
- Fivush, R. (2013). Autobiographical Memory. In E. Keightley & M. Pickering (eds.), *Research Methods for Memory Studies* (pp. 13-28). Edinburgh University Press.
- Güldenring, B. A. (2017). Emotion metaphors in new Englishes. *Cognitive Linguistic Studies*, 4(1), 82–109. <https://doi.org/10.1075/cogls.4.1.05gul>
- Güldenring, B. A. (2020). *Emotion Metaphors in New Englishes: A Corpus-Based Study of Emotion Concepts in Institutionalized Second-Language Varieties of English*. Philipps-Universität Marburg. <https://doi.org/10.17192/z2020.0480>
- Holland, A. C. & Kensinger, E. A. (2010). Emotion and autobiographical memory. *Physics of Life Reviews*, 7(1), 88–131. <https://doi.org/10.1016/j.plrev.2010.01.006>
- hooks, b. (1989). *Talking Back: Thinking Feminist, Thinking Black*. South End Press.
- Kolshansky, G. (2007). *Contextual Semantics*. LKI.
- Kövecses, Z. (1986). Metaphors of anger, pride and love: a lexical approach to the structure of concepts. *Pragmatics and Beyond*, VII (8). John Benjamins.
- Kövecses, Z. (1990). *Emotion Concepts*. Springer-Verlag.
- Kövecses, Z. (2000). *Metaphor and Emotion: Language, Culture, and Body in Human Feeling*. Cambridge University Press.
- Kövecses, Z. (2002). *Metaphor: A Practical Introduction*. Oxford University Press.
- Kövecses, Z. (2015). *Where Metaphors Come from. Reconsidering Context in Metaphor*. Oxford University Press.
- Landová, E., Polák, J., Janovcová, M., Štolhoferová, I., Peterková, Š., Chomik, A. & Frynta, D. (2025). Imprint of ancestral and modern threats in human mind - experience of fear, disgust, and anger. *Frontiers in Psychology*, 15, 1-16. <https://doi.org/10.3389/fpsyg.2024.1520224>
- Lazarus, R. S. (1991). *Emotion and Adaptation*. Oxford University Press.
- LeDoux, J. E. (2012). Evolution of human emotion: a view through fear. *Progress in Brain Research*, 195, 431–442. <https://doi.org/10.1016/B978-0-444-53860-4.00021-0>
- LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences*. 111, 2871–2878. <https://doi.org/10.1073/pnas.1400335111>
- LoBue, V. & DeLoache, J. S. (2008). Detecting the snake in the grass: attention to fear-relevant stimuli by adults and young children. *Psychological Science*, 19, 284–289. <https://doi.org/10.1111/j.1467-9280.2008.02081.x>
- Martin, J. & White, P. R. (2005). *The Language of Evaluation: Appraisal in English*. Palgrave Macmillan.
- Mash, E. J. & Barkley, R. A. (2014). *Child Psychopathology*. The Guildford Press.

- Maternová, T. (2024). *HAPPINESS in American and British English: A Corpus-Based Metaphorical-Pattern Analysis*. Masaryk University. Available from: <https://theses.cz/id/naqsgc/>
- Merriam-Webster.com. (2026). *Merriam-Webster Dictionary*. Retrieved 12 August 2025. <https://www.merriam-webster.com/>
- Miroshnychenko, P. (2020). Conceptualization of fear in the news about the coronavirus. *Scientific Notes of Institute of Journalism*, 2(77).
- Öhman, A. & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483–522. <https://doi.org/10.1037/0033-295X.108.3.483>
- Ortony, A., Clore, G. & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press.
- Ostermann, C. (2012). Cognitive lexicography of emotion terms. In R. V. Fjeld & J. M. Torjusén (eds.), *Proceedings of the 15th EURALEX International Congress* (pp. 493–501). Department of Linguistics and Scandinavian Studies, University of Oslo.
- Ostermann, C. (2015). *Cognitive Lexicography: A New Approach to Lexicography Making Use of Cognitive Semantics*. De Gruyter.
- Oxford University Press. (2025). *Oxford English Dictionary*. Retrieved 12 August 2025. <https://www.oed.com/>
- Paulon, M. (2022). Hysteria: rise and fall of a baffling disease. A review on history of ideas in medicine. *Journal of Psychopathology*, 28(3), 152–161. <https://doi.org/10.36148/2284-0249-461>
- Pelěšková, Š., Polák, J., Janovcová, M., Chomik, A., Sedláčková, K., Frynta, D. & Landová, E. (2024). Human emotional evaluation of ancestral and modern threats: fear, disgust, and anger. *Frontiers in Psychology*, 14, 1–14. <https://doi.org/10.3389/fpsyg.2023.1321053>
- Peräkylä, A. & Sorjonen, M. L. (2012). *Emotion in Interaction*. Oxford University Press.
- Prokop, P. (2016). Universal human fears. In T. K. Shackelford & V. A. Weekes-Shackelford (eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1–5). Springer International Publishing.
- Ruppenhofer, J., Ellsworth, M., Petruck, M., Johnson, C., Baker, C. & Scheffczyk, J. (2016). *FrameNet II: Extended Theory and Practice*. Available at: <https://n9.cl/xq916>
- Tash, M. S., Kolesnikova, O., Ahani, Z. & Sidorov, G. (2024). Psycholinguistic and emotion analysis of cryptocurrency discourse on X platform. *Scientific Reports*, 14, 8585. <https://doi.org/10.1038/s41598-024-58929-4>
- Tremewan, T. & Strongman, K. T. (1991). Coping with fear in early childhood: Comparing fiction with reality. *Early Child Development and Care*, 71(1), 13–34. <https://doi.org/10.1080/0300443910710102>

- Van der Kolk, B. (2014). *The Body Keeps the Score: Brain, Mind, and Body in the Healing of Trauma*. Allen Lane, Penguin Books.
- Wardell, V., Madan, C. R., Jameson, T. J., Cocquyt, C. M., Checknita, K., Liu, H. & Palombo, D. J. (2021). How emotion influences the details recalled in autobiographical memory. *Applied Cognitive Psychology*, 35(6), 1454–1465. <https://doi.org/10.1002/acp.3877>
- Wierzbicka, A. (1990) The semantics of emotions: Fear and its relatives in English. *Australian Journal of Linguistics*, 10(2), 359-375.
- Wierzbicka, A. (1992). *Semantics, Culture, and Cognition: Universal Human Concepts in Culture-Specific Configurations*. Oxford University Press.
- Wierzbicka, A. (1999). *Emotions across Languages and Cultures*. Cambridge University Press.
- Wierzbicka, A. (2009). Language and metalanguage: Key issues in emotion research. *Emotion Review*, 1, 3-14.
- Zhgun, D. (2019). The semantic content of the notions fear and anger and the conditional changes of these emotions. *Lengua y Habla*, 24, 243-257.



¿CARCAMUSAS PARA LOS CARCAS Y LAS MUSAS? LA HISTORIA DEL NOMBRE DEL PLATO TOLEDANO

CARCAMUSAS FOR THE CARCAS ('OLD FOGEYS') AND THE MUSAS ('MUSES')? THE HISTORY OF THE NAME OF TOLEDAN DISH

Francisco Escudero Paniagua
Universidad Rey Juan Carlos
francisco.escudero@urjc.es

RESUMEN

Este trabajo aborda la historia y la etimología de la voz *carcamusa(s)*, que actualmente hace alusión a un plato toledano. Los objetivos de este trabajo son, por un lado, comprobar la verosimilitud de las propuestas etimológicas planteadas y, por otro lado, rastrear el uso histórico de la voz en español (primeras documentaciones de la voz, significados y étimo inmediato). Para ello se ha procedido a buscar las posibles variantes de la voz *carcamusas* en corpus históricos y en repositorios digitales. La etimología más difundida es que se trata de un compuesto *carca* + *musa(s)*. Esta hipótesis parece una etimología popular y no es aceptada por todos. Otra etimología, difundida recientemente en algunas páginas web y publicaciones de prestigio, relaciona la voz con el hebreo *car-camush*, supuestamente 'carne arrugada', de origen romance o hebreo. Esta propuesta carece totalmente de base documental que la apoye y puede ser descartada. La revisión de los corpus y repositorios digitales, en cambio, arrojan más luz sobre la posible historia de la palabra. Desde 1805 y durante todo el siglo XIX se documentan casos de *carcamusa* cuyos usos se asemejan a los de *cancamusa* descrito por el *DEA* y por el *DLE*. En el siglo XX no se vuelve a documentar hasta 1984, cuando se alude por primera vez al plato toledano. Aunque aparentemente los usos de *carcamusa* del siglo XIX no guardan relación con el uso actual, resulta extraño que existan dos palabras homónimas tan peculiares sin relación entre sí; si bien no está del todo claro, lo más probable es que, de manera directa o indirecta, la voz *carcamusa(s)*, usada para referirse al plato toledano, provenga de los usos decimonónicos de *carcamusa*, que se trata, a su vez, de una variante del sustantivo *cancamusa*.

Palabras clave: carcamusas, cancamusa, léxico gastronómico, etimología, historia de la lengua.

ABSTRACT

This paper addresses the history and etymology of the word *carcamusa(s)*, which currently refers to a dish from Toledo. The objectives of this paper are, on the one hand, to verify the plausibility of the proposed etymologies and, on the other hand, to trace the historical use of the word in Spanish (first documented uses of the word, meanings and etymology). To this end, we have searched for possible variants of the word *carcamusas* in historical corpora and digital repositories. The most widespread etymology is that it is a compound of *carca* + *musa(s)*. This hypothesis seems to be a popular etymology and is not accepted by everyone. Another etymology, recently disseminated on some websites and in prestigious publications, relates the word to the Hebrew *car-camush*, supposedly 'wrinkled meat', of Romance or Hebrew origin. This proposal lacks any documentary basis to support it and can be dismissed. However, a review of digital corpora and repositories sheds more light on the possible history of the word. From 1805 and throughout the 19th century, there are documented cases of *carcamusa* whose uses resemble those of *cancamusa* described by the *DEA* and the *DLE*. In the 20th century, it is not documented again until 1984, referring to the Toledan dish. Although the uses of *carcamusa* in the 19th century appear to be unrelated to its current use, it is strange that there are two such peculiar homonymous words with no relation to each other. Although it is not entirely clear, it is most likely that, directly or indirectly, the word *carcamusa(s)*, used to refer to the dish from Toledo, comes from the nineteenth-century uses of *carcamusa*, which is, in turn, a variant of the noun *cancamusa*.

Keywords: carcamusas, cancamusa, gastronomic lexicon, etymology, history of Spanish language.

Recibido: 12-03-2026
Aceptado: 29-04-2026

DOI: <https://doi.org/10.17561/rilex.9.2.10333>



1. INTRODUCCIÓN

Carcamusa —o *carcamusas*— es el nombre que recibe un plato típico de la ciudad de Toledo y alrededores. Se trata de un estofado que contiene principalmente magro de cerdo, tomate y guisantes, entre otros ingredientes. La voz no está recogida en diccionarios de uso común, como el *Diccionario de la lengua española* (DLE), de la RAE (s. f.), o el *Diccionario del español actual* (DEA), de Seco, Andrés y Ramos (2023), aunque no existe duda de su significado actual. No existen estudios lingüísticos sobre su etimología, pero sí existen varias propuestas creadas en ámbitos no académicos, algunas de ellas muy difundidas y aceptadas en la sociedad toledana y en el ámbito culinario. Sin embargo, muchas de ellas se contradicen entre sí, ninguna se basa en documentación comprobable y o bien parecen producto de la etimología popular, o bien no llegan a explicar verdaderamente el origen del nombre. Solo en un caso, que se tratará de manera pormenorizada en el apartado 3.2, se propone un origen de la voz.

Así, el objetivo principal de este trabajo es arrojar luz sobre la etimología de la palabra que designa el plato toledano. Para ello se han revisado las hipótesis más difundidas sobre el origen del vocablo, se ha estudiado su fundamentación, se ha investigado en corpus lingüísticos académicos (*CORDE*, *CREA*, *CORPES XXI*, *CDH*, *NTLLE*) y se ha buscado documentación en la *Biblioteca Digital Hispánica*, en la *Hemeroteca Digital*, en la *Biblioteca Virtual de Prensa Histórica* y en *Google Books*. Finalmente, se ha trazado una historia de la voz, desde las primeras documentaciones de la palabra hasta las más recientes, se han descrito sus significados y se ha analizado su evolución. Con estos datos, se ha realizado una propuesta etimológica e histórica de *carcamusa(s)* que explica, hasta donde ha sido posible, el origen del término.

2. CONTEXTO CULINARIO

Una parte importante de las etimologías propuestas suele vincular la aparición del plato con la aparición del sustantivo *carcamusa(s)*, es decir, que

supuestamente se inventó la palabra para denominar ese estofado nuevo. Otras etimologías, en cambio, no llegan a explicar el verdadero origen del nombre. Simplemente se explica que alguien bautizó la comida como *carcamusa(s)*, pero no se hace alusión a si se trataba de una invención *ad hoc* o si existía previamente la palabra con otro significado.

Para indagar en el origen de la voz es necesario, por tanto, aclarar antes el origen del plato —aunque este sea igual de oscuro o más que el origen del sustantivo—. En otras palabras, conviene realizar una aproximación a la historia de la gastronomía para llevar a cabo este trabajo de historia de la lengua. Ambas perspectivas son necesarias para el análisis de las propuestas etimológicas y para reconstruir convenientemente la historia de la voz.

La propuesta más extendida y aceptada es que las carcamusas (el plato) fueron creadas a mediados del siglo XX en el bar Ludeña, ubicado en el casco histórico de la ciudad de Toledo. Se suele atribuir su invención al dueño del bar por aquel entonces, José Ludeña, y a su esposa, M.^a Mercedes Herrera. Esto lo afirmaba la propia M.^a Mercedes Herrera en una entrevista que realizó Carlos Iserte en el programa *Variotinto*, de Castilla-La Mancha Media:

Bueno, pues las carcamusas, mire usted, las inventamos mi marido y yo [...]. Y entonces queríamos poner una cosa de que no en Toledo que no hubiese esta cosa [sic] y al poner el bar dice “tengo que hacer por ahí una cosa que no tiene...” aunque co...empezamos a pensar aquí sentaditos los dos y “pongo esto, pongo lo otro” y ya dice “María, ¿y poner *carcamusas*?” digo “¡Ah, pues sí, me gusta, fíjate! [...]” (Castilla-La Mancha Media, 2016. La transcripción es nuestra).

Otro testimonio diferente se encuentra en el apartado de “Discusión” de la entrada de “Carcamusa” de la Wikipedia. En el texto, redactado y firmado por Chema Marín (periodista) el 26 de mayo de 2017¹, se afirma que el plato lo inventó Rufo Herrera Jiménez, camarero y familiar de José Ludeña y M.^a Mercedes Herrera.

Al leer el origen de las Carcamusas toledanas he de hacer algunas aclaraciones que he contrastado con los familiares del camarero que inventó este plato tan arraigado

¹ Esta información se ha tomado del historial de “Discusión: Carcamusa” de Wikipedia.

en la capital toledana. En primer lugar ratificar que el Restaurante Ludeña fue el primer bar de Toledo donde se sirvió este plato. Sin embargo el inventor del mismo no fue el propietario del bar, sino un camarero familiar suyo, Rufo Herrera Jiménez. Las Carcamusas surgieron cuando Rufo pretendía realizar una receta tradicional de venado de San Pablo de los Montes, probando con carne de cerdo e incorporando tomate junto al resto de ingredientes (Marín, 2017)².

Esta versión contrasta con lo narrado por M.^a Mercedes Herrera, pero añade un dato relevante sobre el origen del plato: no se trata de una invención total, como afirmaba la viuda del dueño, sino una versión de un plato ya existente, procedente de la región de los Montes de Toledo y, concretamente, de San Pablo de los Montes. Efectivamente, en la provincia de Toledo y especialmente en zonas montañosas, son comunes los estofados de venado.

La idea de que las carcamusas fueron más una versión innovadora que una novedad gastronómica también es recogida por el cocinero profesional Jorge Hernández Alonso (conocido también como Chef Koketo) en su página web³.

Carcamusas son las habituales magras de mi madre
He de reconocer que no conocía este plato, que sin embargo he comido toda mi vida en casa de mis padres y que con mayor o menor acierto era llamado en el lar: “magro con tomate” [sic]. Me sorprende que una asturiana utilice los mismo [sic] ingredientes que los toledanos para esta elaboración, pero como veréis “las magras” son conocidas en diferentes regiones. Este mismo alias materno, las magras, lo recibe un plato navarro de origen aragonés. Pero a diferencia de su primo castellano no va acompañado de los omnipresentes guisantes [...]. También son bien conocidas de la cocina murciana donde corren la misma suerte en tascas y bares, ocupando un lugar privilegiado como tapa o aperitivo de Cartagena. [...]. No podemos olvidar las magras turolenses. En “Teruél [sic] también existe” la carne del cerdo no es fresca sino curada, un jamón serrano que acompaña a la salsa de tomate y es imprescindible además un huevo frito (Hernández Alonso, 2018).

Así, el plato toledano no sería un plato nuevo, sino una versión local más de las magras con tomate, que se encuentran por la geografía española. Esto, lejos de ser anecdótico, resulta relevante para la historia de la voz.

² Chema Marín afirma que la fuente de esta información es la familia Herrera.

³ Este cocinero también recoge la versión de la familia Herrera de que los creadores del plato fueron José Ludeña y Rufo Herrera, sin mencionar a M.^a Mercedes Herrera. Su fuente es Ana Herrera, nieta de Rufo Herrera.

Si las carcamusas no fueron una invención gastronómica, sino una versión toledana de un plato ya existente, ¿las versiones más antiguas también se denominaban *carcamusas*? Pues bien, en una línea parecida, una de las propuestas etimológicas de la voz parece sugerir la existencia de comidas o platos antiguos de los que toman su nombre las carcamusas. No obstante, se darán más detalles de esta propuesta etimológica en el siguiente apartado.

3. PROPUESTAS ETIMOLÓGICAS

Como se ha podido comprobar, M.^a Mercedes Herrera explicaba que durante la invención del plato su marido dijo “¿Y poner *carcamusas*?”, propuesta que ella aceptó. Por el contrario, Chema Marín, en su nota a la entrada de Wikipedia, señala que:

En un primer momento [Rufo Herrera] las bautizo [sic] como Lisardas, pero más tarde el farmacéutico del barrio, Don Francisco propuso el nombre de Carcamusas y con él llegó hasta nuestros días (Marín, 2017).

Sin embargo, ninguno de los dos testimonios aporta información sobre el proceso de formación de la voz. Es más, no queda claro si la palabra *carcamusas* se creó para denominar el plato toledano o si existía antes de él. Teniendo esto en cuenta, podemos agrupar las siguientes propuestas etimológicas en dos tipos: i) las que proponen que la palabra se inventó para designar la comida; y ii) las que proponen que la palabra existía antes de su invención.

3.1. LA PALABRA SE INVENTA CON EL PLATO: LA COMPOSICIÓN CARCA + MUS(AS)

Dentro del primer grupo están las propuestas más extendidas y aceptadas socialmente, que presentan las características propias de etimologías populares. La etimología popular es un mecanismo de creación lingüística que “procura otorgarle al elemento incomprensible por parte del oyente-hablante un determinado sentido que en definitiva supone una reinterpretación” (García Manga, 2006, p. 106). En otras palabras, es una búsqueda de una motivación lingüística que explique el origen de una palabra extraña al hablante y/o la

dote de un significado más transparente. Esto es lo que ocurre precisamente con *carcamusas*; esta palabra anómala para muchos se ha interpretado a partir de elementos ya conocidos.

García Manga (2006, pp. 108-109) señala varios factores que favorecen la aparición de etimologías populares. Dos de ellos se dan en las diferentes propuestas etimológicas que se ven en este apartado: i) ha de existir identidad o similitud de formas fónicas y ii) el campo de las palabras compuestas y derivadas es propenso a ser objeto de etimologías populares. Así, se va a tender a interpretar *carcamusas* como un compuesto de *carca* + *musas*.

La etimología de *carcamusas* más famosa es la que afirma que *carca* hace referencia a hombres mayores y *musa(s)*, a mujeres jóvenes. Esta propuesta es, asimismo, la más citada en la mayoría de las entradas de páginas web sobre el plato (si no en todas). Se presentan a continuación algunos ejemplos:

El nombre podría hacer referencia a los clientes que frecuentaban el bar Ludeña, hombres de mediana edad (“carcas”) y por algunas señoritas más jóvenes, que ellos veían como sus “musas”. Como el plato era del gusto de ambos grupos, el nombre dado al mismo fue *carca-musas* (Alonso, 2026).

Dicen que fueron sus artífices quienes cambiaron el nombre del plato al contemplar la barra del bar llena de jubilados “carcas” que se entretenían siguiendo las conversaciones de las jóvenes universitarias “musas”. De ahí su nombre “Carcamusas” (Moratilla, 2022).

Cuenta la leyenda que las *carcamusas* es un plato ideado en un bar típico de Toledo que, durante mediados del siglo XX solían frecuentar hombres de cierta edad, a los que llamaban los *carcas* y mujeres jóvenes que ellos consideraban sus *musas*. Así, don José Ludeña, dueño y chef de esta receta, decidió denominarla *carca-musas* en honor a ambos grupos, ya que maravillaba a sendos (RTVE, 2023).

Se verá que, según la fuente, o bien los inventores del término no se mencionan, o bien se atribuye su invención a los creadores del plato o, directamente, a José Ludeña. Se podría pensar, por tanto, que esta propuesta etimológica complementa lo afirmado por M.^a Mercedes Herrera, pero en la entrevista mencionada anteriormente niega que el origen del término sea este.

Circula una segunda versión que también considera la voz como un compuesto, pero, en este caso, de *carca* (hombre mayor) y *mus* (juego de cartas):

Mientras disputaban sus partidas de mus y degustaban su aperitivo preferido, de los pocos alimentos que se servían ya que en la barra se dispensaban solo conservas. Los protagonistas seguían siendo los “Carcas” que jugaban a los naipes febriles partidas de “mus” (Hernández Alonso, 2018).

Por último, una tercera versión afirma que *carca* hace referencia a los jubilados que jugaban al dominó y *musas*, a las magras:

Existe otra, más plausible si cabe que la primera, y es que los ociosos jubilados pasaban la mañana jugando al dominó en el Bar Ludeña, hasta que sus estómagos eran avasallados por el aroma de las magras de cerdo entomatadas con las que cogían fuerzas para continuar la partida al tiempo que espabilaban del letargo de “la siesta del carnero”. Así eran los “carcas” del dominó quienes se comían a sus “musas” (las magras), en un tiempo en el que de las cocinas de los bares no salían más aromas que el de alguna tortilla de patata (Moratilla, 2022).

3.2. LA PALABRA PRECEDE AL PLATO: EL PRESTIGIO DE UNA PROPUESTA PSEUDOLINGÜÍSTICA

En la entrevista de Carlos Iserte a M.^a Mercedes Herrera, el periodista menciona que la palabra aparece en la letra de una canción⁴. Esta propuesta demostraría la preexistencia del vocablo antes de la creación del plato, pero no hemos podido encontrar ni el título, ni la letra ni una versión de dicha canción. Tal vez no exista, pues no es una propuesta etimológica muy conocida o extendida.

Diferente suerte ha tenido otra propuesta etimológica que apunta directamente a un étimo antiguo y que ha ido obteniendo validación y prestigio entre quienes han escrito sobre el origen de las carcamusas y/o su etimología. Se trata de una propuesta etimológica anónima, publicada en el blog Celtiberia.net, el 7 de noviembre de 2005. Lo único que sabemos del (de la) autor(a), es su nombre de usuario, Aromarti. Su propuesta aparece en un comentario a una entrada de blog escrita por él o ella, titulado “Del romance a la lengua universal”, publicado el 1 de noviembre de 2005. Tanto en la entrada como

⁴ Cita en concreto el verso “si quieres que te baile la carcamusa” (Castilla-La Mancha Media, 2016) (la transcripción es nuestra).

en el comentario detalla la historia del castellano y de lenguas semíticas en la península ibérica y expone numerosas etimologías de palabras castellanas de origen hebreo. Entre ellas, aparece *carcamusas*, que sería un compuesto de *carne* y *camush*, ‘arrugado’:

CARCAMUSA. (Carne de caza, generalmente venado, guisada y plato popular y típico en los Montes de Toledo). Camush (con shim final) significa en romance “arrugado”. Car-camush es carne arrugada (el que coma carcamusa lo sabe) (Aromarti, 2005)⁵.

Nótese que el étimo *camush* se escribe o se pronuncia “con shim final”. Aunque no lo deja muy claro en la etimología, parece referirse a la letra del alefato (shin, ש, [ʃ] o [s]), por lo que se deduce que se trataría de un hebraísmo. Conviene, no obstante, hacer una serie de aclaraciones sobre esta etimología antes de darle cualquier grado de credibilidad.

En primer lugar, no se ha encontrado documentación en los corpus históricos analizados que certifique la presencia de *carcamusa(s)*, *carcamush* o de un supuesto *camush* o similar en la lengua española antes del siglo XIX; ni siquiera los usos documentados en ese siglo tienen relación con ningún tipo de alimento. En segundo lugar, no se ha podido comprobar que *camush* significase ‘arrugado’ en romance o en hebreo. Además, hay que señalar que Aromarti muestra una acusada inclinación por atribuir un origen hebreo a palabras del español que, en realidad, tienen otro origen. Aquí algunos ejemplos (Aromarti, 2005):

CALCAR (en sentido de copiar). En hebreo CALCAL es nutrir o abastecer en cantidad suficiente para que no falte. El verbo es LECALCAL. De aquí sale la palabra “clavado” cuando algo está copiado o calcado de manera perfecta pues CaLAL es lo igual, lo completo.

CLUB: viene de Cluv: literalmente jaula, cuarto pequeño y sucio. De aquí se deriva cubil. ¿porqué club es una palabra internacional? PORque es hebrea,

CAÑÍ. Relacionado con lo anterior. CANUÍ es el “llamado” o “el apellidado”. Cañi-zares. Cañ-ete. Cañ-avate ¿Más?

⁵ Dada la cantidad de erratas ortotipográficas que existen en la redacción del usuario o usuaria Aromarti, se prescinde de añadir la marca [sic] en este y en los siguientes fragmentos citados redactados por él o ella. Puesto que se trata de textos copiados y pegados directamente de la página web, cualquier errata del texto citado está en el original.

CASA... En hebreo CASA es un sitio cubierto, algo que está cubierto o tapado. CASAI.

ALCACHOFA.... AL es “en lo alto” “por encima de ” y CASÚF significa plateado y también “por encima”. Con este juego de palabras nace la alcachuf (una planta de alcachofa tiene una pubescencia blanca toda ella preciosa).

Por último, tanto la entrada como el comentario escrito por Aromarti carecen del más mínimo atisbo de rigor histórico. Véanse, como ejemplos, los siguientes fragmentos sobre la “historia” del castellano (Aromarti, 2005):

¡Claro que la lengua española es una mezcla de otras lenguas romances! ¡Claro! Porque todas las lenguas romances comparten un tronco común: el hebreo. La península ibérica estaba poblada de íberos o iberís. Y qué curioso que hebreo en hebreo se diga ibrí o iberí (depende de como vocalices la b).

A partir de su llegada [de los judíos] a Iberiá se convierten en los Íberos (son íberís ó hebreos). Conviven pacíficamente godos e íberos, visigodos e íberos, celtas e íbero, etc. etc.,

Como Hispania era el nombre romano, los hebreos, maestros de la lengua, cogen la S la P/F y la N (parecida en escritura a la R latina) y la llaman SeP/FaRad... Sefarad, de ahí sefardí: el habitante de Hispania o Iberia.

Pese a ello, la propuesta etimológica romance/hebreo de *carcamusa*, ha sido muy difundida en fechas recientes. Aromarti redactó la etimología de *carcamusa* el 7 de noviembre de 2005, pero el 13 de julio de 2015 un autor anónimo la incluyó de manera literal en el artículo de Wikipedia sobre las carcamusas:

En realidad la palabra carcamusa (Carne de caza, generalmente venado, guisada y plato popular y típico en los Montes de Toledo) tiene su etimología en Camush (con shim final) significa en romance “arrugado”. Car-camush es carne arrugada (el que coma carcamusa lo sabe) (Carcamusas, 2015).

Este fragmento se ha modificado ligeramente en varias ocasiones y, actualmente, aparece lo siguiente con un enlace a los textos de Aromarti:

Según otras opiniones, la palabra carcamusa tendría su etimología en *camush* (con *-sh* final), que significa, en romance, “arrugado”, de modo que *car-camush* sería “carne arrugada” (Carcamusas, 2026).

Esta entrada de Wikipedia parece ser la fuente directa de los redactores de las páginas web que se hacen eco de esta propuesta etimológica, quienes la atribuyen a veces a “estudiosos del lenguaje” o a “la intelectualidad”:

Otros señalan que el origen de esta receta proviene de la lengua romance *car-camush* que significa carne arrugada, término que define muy bien cómo queda la carne tras ser cocinada de esta manera (RTVE, 2023).

Estudiosos del lenguaje aseguran que el vocablo “carcamusa” deriva etimológicamente de la lengua romance “car-camush” que vendría a ser “carne arrugada” y no se refieren a la de los carcas, más bien a la del puerco al fuego (Hernández Alonso, 2018).

La intelectualidad, tiene una tercera hipótesis, que deja fuera jóvenes, dominó y carcas, atribuyendo su nombre al lenguaje romance en el que la palabra “Camush” significa “arrugada”, por lo que podría venir de “car-camush” lo que vendría a ser “carne arrugada” (Moratilla, 2022).

Recientemente, esta pseudoetimología se ha hecho un hueco en publicaciones que gozan de mayor prestigio lingüístico, como es el *Diccionario gastronómico-literario de la cocina tradicional española (de la A la Z)*, realizado por Archiletras en colaboración con Paradores (2025).

4. HISTORIA DE LA VOZ A PARTIR DE LA DOCUMENTACIÓN DISPONIBLE

Dado que ninguna de las etimologías propuestas para *carcamusa(s)* parte de un estudio profundo, con base documental, que permita trazar verazmente la historia de la palabra, se ha procedido a recopilar testimonios de uso de voces similares o iguales a *carcamusa(s)* a lo largo de la historia y se ha analizado su significado.

Aunque la palabra alude a un término gastronómico, la búsqueda de testimonios no se ha ceñido a los textos culinarios: en primer lugar, la voz puede haberse documentado en textos de otra tipología, por lo que conviene atender a fuentes de todo tipo (Eberenz, 2014, pp. 30-33; Torres Martínez, 2022, p. 118); en segundo lugar, no está claro que la palabra *carcamusa(s)* siempre haya estado relacionada con el ámbito gastronómico, por lo que es posible que no se haya registrado en textos culinarios hasta épocas muy recientes.

En consecuencia, para esta investigación se han consultado diversos corpus de la RAE y en otros recursos digitales. La idoneidad del uso de fuentes digitales para la investigación del español de orientación diacrónica ha sido destacada por Campos Souto en varios trabajos (2018a, p. 238; 2018b, pp.

33-41) y, como se comprobará a lo largo de este apartado, son uno de los recursos fundamentales para el rastreo y el trazado histórico de voces de escasísima frecuencia, como *carcamusa(s)*.

Tras haber buscado en los corpus académicos (*CREA*, *CORDE*, *CORPES XXI*, *CDH* y *NTLLE*), solo se han encontrado cuatro testimonios en el *CORPES XXI*: dos en *La pasión según las fieras*, de Fernando Royuela y dos en *Los templos de la tapa. Los mejores locales de la Comunidad de Madrid*, de Sara Cucala, libros publicados en 2003 y 2009 respectivamente. Se trata, por tanto, de testimonios muy recientes, que hacen alusión al plato toledano. En cambio, en las fuentes digitales (*Biblioteca Digital Hispánica*, *Hemeroteca Digital Hispánica*, *Biblioteca Virtual de Prensa Histórica* y *Google Books*) se ha detectado *carcamusa(s)* (o formas similares) en unos pocos textos desde 1588 hasta 1984, con seis significados distintos. Se presentan a continuación de forma cronológica con la fecha de la primera aparición documentada:

1. Ariete romano (1588)
2. Nombre propio (1603)
3. Murmullo / habladuría, engaño (1805)
4. Nombre propio de personajes literarios (1866)
5. Sustantivo en aposición de sentido enfático (1884)
6. Plato de comida (12/02/1984)

El primero de los sentidos hacen alusión a voces latinas, no existentes aparentemente en castellano y, por tanto, no relacionadas con el plato toledano. El segundo es un apellido español, que nada tiene que ver con el plato. Los tres siguientes también son voces españolas, parecen estar relacionadas entre sí, pero, *a priori*, tampoco parecen estar relacionadas con las carcamusas toledanas.

4.1. ARIETE ROMANO

La voz *carcamusas* y *carcamusa* se documenta por primera vez en latín con el sentido de ‘ariete’ o ‘tipo de ariete’ romano en *Annalium et historiae francorum* de Pierre Pithou (1588):

Sub lucem, reuehut crates sua ad oppida furtim,
Arrietes carcamusas vulgo resonatos
Dimisere duos, pallos vetuit remouere
Quos nostri capiunt gaudenter depeculantes,
[...] (1588, p. 455).

Posteriormente, se atestigua en varios textos escritos en diferentes lenguas: *Analecta siue observationes reliquae ad militam et hosce libros*, de Justo Lipsio (1602), *Les antiquités et histoires gauloises et françoises*, de Claude Fauchet (1611), *Histoire des Comtes de Carcasonne*, de Guillaume de Bessé (1645), *Glossarium ad scriptores mediae et infimae latinitatis*, de Charles du Fresne (1710), *Il tesoro delle antichità sacre e profane*, de Agostino Calmet (1746), *A New and Enlarged Military Dictionary*, de Charles James (1805) o en el número 220 del periódico *El Vapor* (1852). Aparece por primera vez documentada en un texto en español en 1828, en el *Diccionario militar español-francés* de Federico Moretti:

CARCAMUSA. s. f. ant. *Carcamuse*. Maquina antigua de guerra, que algunos autores creen ser el ariete, ó muy parecida a aquel (1828, p. 412).

4.2. NOMBRE PROPIO: DIEGO DE CÁRCAMO

En el segundo tomo de *Hispaniae Illustratae* (1603), recopilación de varios textos en latín escritos por varios autores, se incluye la *Historia de la unión del reino de Portugal a la Corona de Castilla*, de Girolamo Franchi Conestaggio. En el libro VI de su obra, aparece una mención a un Iacobus Carcamus (1603, p. 1162) que es la traducción al latín de Diego de Cárcamo, funcionario real y secretario del prior de Crato (BRASILHIS Database, 2015), tal y como se puede comprobar en la traducción de la obra realizada por Luis de Bavia (1610, pp. 143-144). Por lo tanto, nada parece tener que ver el apellido latinizado con el plato toledano.

4.3. MURMULLOS Y RUMORES: EL ÉTIMO INMEDIATO CANCAMUSA

En diferentes textos del siglo XIX se documenta *carcamusa* con significados muy parecidos entre sí. La voz se documenta por primera vez en la *Colección de poesías pósthumas de D. Manuel Rincón*, publicadas por su padre, Genaro Faustino Rincón en 1805, con un significado aproximado a ‘murmullo o ruido molesto’.

Jua. No hay enfermo,
Pero callar es preciso. *En voz baja.*
Nar. ¿Por qué? *En voz alta.*
Jua. Porque no se puede.
Nar. Pero sin duda hay motivo. *Se sienta.*
Jua. Pero es preciso callar.
Man. ¡Qué carcamusa! ¡Qué ruido!
No puede salir un verso (1805, p. 79).

Posteriormente, se documenta con un significado aproximado a ‘rumor falso’ o ‘engaño’ en un artículo titulado “El glorioso” y publicado en el número 14 de *Guindilla* el 1 de septiembre de 1842:

Y qué hacen estos ganforros? Prometen miel, almíbar y piñoncitos mondados; pero lo que hacen es cebarse como los otros en las entrañas de su patria y las desgarran como tigres por saciar su ambicion. El egoismo es el móvil de sus infames carcamusas, la dilapidacion su ejercicio cotidiano. Nada les importa que perezca su patria, que la libertad se hunda (1842, p. 214).

Trece años después, el 24 de abril de 1855 vuelve a aparecer en *El Clamor Público*, con un significado similar al anterior:

Ningun medio omiten los cuestores de Monistrol para embaucar á sus víctimas. Se hacen dispensadores de indulgencias, crean ad libitum santuarios y hospitales que nadie ha visto ni ellos tampoco, hablan de manutencion de sacerdote, y se atraen con sus carcamusas numerosos cofrades (1855, p. 1).

Nuevamente, varias décadas después, en 1887 José Ramón Mélida la usa con el mismo significado en *A orillas del Guardarza*:

Todas esas *carcamusas* me huelen á que por ahí, por esas Universidades, anda suelto Lucifer engañando á los sencillotes como tú (1887, p. 120).

En vista de todo ello, es probable que *carcamusa* sea una variante de *can-camusa*, documentada por primera vez en *La vida y hechos de Estebanillo González* de 1646 (Corominas & Pascual, 1985, p. 195; *CDH*). El *Diccionario de autoridades* (1729) recoge la voz con la siguiente definición:

CANCAMUSA. S. f. Intencion fingída, con que se tira à deslumbrar à alguno, para que no se entienda el enaño que se le vá à hacer: y assi se suele decir, Yá le entiendo la cancamúsa, No entendió la cancamusa.

Actualmente sigue presente en el *DLE* con las marcas *coloquial* y *desusado* y la siguiente definición⁶:

1. *f. coloq. Desus.* Dicho o hecho con que se pretende desorientar a alguien para que no advierta el engaño de que va a ser objeto.

En este diccionario también se recoge *recancamusa* con una definición más simple, pero semánticamente parecida y sin la marca *desusado*⁷.

1. *f. coloq.* Maña o artificio para encubrir un engaño.

Ambas definiciones coinciden en gran medida con los usos de *carcamusa* de los textos de 1842, 1855 y 1887. En cambio, el de 1805 parece estar relacionado con la definición de *cancamusa* del *DEA*:

(*desp*) Monserga (cosa fastidiosa y que no merece atención).

Evidentemente, la coincidencia entre *cancamusa* y *carcamusa* no solo es semántica, sino también formal. La única diferencia está en un sonido o, mejor dicho, en el cambio [n] > [r]. Este tipo de rotacismo no es frecuente, pero sí posible en español. Rost Bagudanch (2009, 2016, 2024) ha documentado casos de alófonos róticos para el fonema /n/ (también de /l/) en español y se ha documentado en otras lenguas (2024, p. 115). En su trabajo de 2016, la investigadora analizó 818 casos de realizaciones de /n/, de los cuales 50 (6,11%) fueron realizaciones róticas (2016, p. 114) y en su trabajo de 2024, obtuvo un porcentaje similar (6,2%, 30 ocurrencias de 486).

Según estas investigaciones de partida, el rotacismo surgiría por “un proceso de lenición por *undershoot* y reducción temporal” (Rost Bagudanch, 2016, p.

⁶En el *DLE* no se hace alusión a la etimología de *cancamusa* ni de *recancamusa*. Sin embargo, Corominas y Pascual (1985, p. 195) sugieren que *cancamusa* tal vez sea un cruce de *canca-murria* y *musa*, y las relaciona, a su vez, con *recancamusa*. Por su parte, Rodríguez-Navas (1906, pp. 286 y 1227) afirma que *recancamusa* proviene de *cancamusa*, que, a su vez, proviene de *cáncana*, ‘banquillo’, y este del griego *karkinos*, ‘medio para atrapar’.

⁷*Recancamusa* se atestigua varias décadas antes de *cancamusa*, en 1626, en *Cuento de cuentos* de Quevedo, y más adelante, en 1874, en *Tradiciones peruanas* de Ricardo Palma, según se ha podido comprobar en el *CDH*. Se documenta en textos lexicográficos desde el *DRAE* de 1899, y durante todo el siglo XX, como sinónimo de *cancamusa*.

117); De hecho, en la mayoría de casos estudiados por Rost Bagudanch, de este tipo de rotacismos, el cambio se realiza por róticas aproximantes, por lo que se mantiene una estructura formántica similar a la de los alófonos nasales y laterales por los que se conmutan las róticas. Así, al reducirse la duración de la consonante y mantenerse la estructura formántica, el alófono presenta las características acústicas y perceptivas de una rótica (*ibid.*).

Aunque la investigadora señala que los alófonos róticos son mucho más numerosos en posición intervocálica, es posible un cambio [n] > [r] en posición de coda, como el de la [n] de *cancamusa*. Así, parece altamente probable que *carcamusa* sea una variante de *cancamusa*, surgida, posiblemente, a finales del siglo XVIII.

4.4. APELATIVOS Y APOSICIONES

En el mismo siglo XIX también encontramos un par de testimonios de *carcamusa* diferentes a los vistos en el subapartado anterior, pero posiblemente relacionados con ello; se registra como un nombre apelativo en el *Collar del diablo (memorias de un resucitado)* de Manuel Fernández González (1866):

Agárrenme ahora y préndanme al tío Carcamusa por embustero y trapalón y alar-mador del vecindario pacífico, diciendo que venían cuarenta salteadores á pegar fuego al pueblo; y no me replique usted, tío Carcamusa, porque me tiene usted muy lleno; y como me tartagee usted una sola palabra, el rompo á usted la justi-cia en las espaldas; y pocas bromas conmigo. Diez dias de cárcel y diez duros de multa, tío Carcamu (1866, p. 447).

Asimismo, José Ramón Mélida, que usó la voz *carcamusa* en su obra *A ori-llas del Guardarza*, de 1887, también la había usado anteriormente en *El demonio con faltas (memorias de un gato)*, de 1884:

Y todavía siguió la carcamusa de la tía Micaela hasta después de salir Manolo, quien al irse dio un portazo que me hizo entremecer. Fuyé cosa de alquilar bal-conaes para oír los elocuentísimos apóstrofes con que la tía Micaela avergonzó en privado y afrentó en público, es decir, ante las visitas, á su desgraciada sorbina (1884, p. 77).

Aquí *carcamusa* se usa como sustantivo o como adjetivo valorativo en un grupo nominal apositivo del tipo *el traidor de Bruto* o *el tonto de Juan* (véase

la *NGLE* de la Real Academia Española, 2009, 12.14b). Por lo tanto, *la carcamusa de la tía Micaela* significaría que *la tía Micaela es una carcamusa*.

Sin embargo, la existencia de ambos testimonios no hace sino reforzar la hipótesis de que *carcamusa* es una variante de *cancamusa*. En los dos casos, *carcamusa* es usado como adjetivo o sustantivo con el que se refiere a personas que dicen carcamusas o cancamusas (esto es, mentiras, engaños). Así, la tía Micaela se califica de mentirosa o engañosa y el tío Cancamusa, acusado de decir mentiras, lleva como sobrenombre la característica por la que se le culpa.

4.5. PLATO TOLEDANO

La primera vez que se documenta *carcamusas* para hacer referencia a la versión toledana de las magras con tomate es el 12 de febrero de 1984 en *La Prensa Alcarreña* IV, n.º 705:

Por lo que se refiere a la caza, es la perdiz la gran especialidad de Toledo. También aquí el cochifrito (cordero, ajos, tomate, aceite, huevos, azafrán, perejil, vino blanco), o las carcamusas, otra de las tapas que se pueden degustar antes del almuerzo en los numerosos bares que se brindan al visitante (1984, p. 4).

El testimonio nos revela que para 1984 las carcamusas ya se cocinaban y se consumían por toda la ciudad, no solo en el Ludeña, donde se había inventado tres décadas antes. Eso sí, casi cien años distan entre el último testimonio encontrado de *carcamusa* del siglo XIX (1887) y el primero del siglo XX. En otras palabras, casi cien años separan el significado ‘dicho o hecho engañoso’ del significado ‘estofado de carne magra con tomate y guisantes típico de Toledo’. Dos motivos pueden explicar este hecho:

1. **Es un fortuito caso de homonimia:** la voz del siglo XIX y la del siglo XX serían palabras con significados diferentes y etimologías distintas pero iguales en género, ortografía y pronunciación. *Carcamusa* y *cancamusa* del siglo XIX cayeron en desuso (tal y como aparece en el *DLE*) y, por eso, no hay registros posteriores, mientras que la palabra *carcamusa(s)* que aparece en la segunda mitad del siglo XX se inventó en Toledo.

2. ***Carcamusa* es una voz demasiado coloquial e infrecuente:** *carcamusa* parece una voz demasiado coloquial (incluso vulgar), cargada de connotaciones negativas y peyorativas, sumamente infrecuente y no está recogida en ningún repertorio lexicográfico. Por tanto, ha permanecido en la oralidad, pocas veces se ha puesto por escrito y, posiblemente por este motivo, la voz aparece en textos que no se han digitalizado todavía (obras literarias consideradas menores u olvidadas, periódicos de menos impacto, correspondencia privada, etc.) Es posible, incluso, que no existan más testimonios escritos de la voz en el siglo XX hasta la difusión del plato toledano.

El primer motivo rechaza la relación etimológica entre *cancamusa* y *carcamusa* ('dicho o hecho engañoso') con *carcamusa* (plato). Por el contrario, el segundo ni la rechaza ni la afirma. Por el momento, en esta investigación no se defiende la homonimia fortuita y sí la relación entre los usos decimonónicos y el uso actual. De ser así, el significado actual de *carcamusa(s)* sería un neologismo semántico, esto es, un nuevo significado para una voz que ya existía (Álvarez de Miranda, 2009, pp. 135-136; Adelstein, 2018, p. 19), aunque muy distante del (o de los) que tenía el lexema en el siglo XIX.

Surgen, por tanto, nuevas preguntas sobre la historia de la palabra: ¿Cómo pasó la voz de designar un dicho o hecho engañoso al plato de comida? ¿Hubo varios cambios semánticos hasta llegar a designar el plato? Ambas dudas solo podrán resolver si existen y se encuentran nuevos testimonios, especialmente entre 1887 y 1984.

5. RESUMEN Y CONCLUSIONES

El sustantivo *carcamusa(s)* ha sido objeto de etimologías populares de hondo calado cultural, especialmente la que atribuye su origen a la composición de *carca(s)* y *musa(s)*. También existen otras propuestas, entre las que destaca la del origen romance-hebreo. Aunque esta hipótesis hebraísta carezca totalmente de base documental que la apoye y pueda ser descartada (por este

y por otros motivos), es un buen ejemplo de cómo una propuesta pseudo-lingüística puede tener eco incluso en publicaciones de prestigio científico, como el *Diccionario gastronómico-literario* de Archiletras, gracias a internet y a páginas web de gran afluencia e impacto como Wikipedia. En cualquier caso, ninguna de estas propuestas es sólida.

Dejando al margen, pues, las elucubraciones etimológicas, lo que se sabe es que el origen del plato toledano parece ubicarse en el bar Ludeña y su creación debió de suceder hacia la segunda mitad del siglo XX —bien como invención original, bien como variante de un plato ya existente—, pero la voz *carcamusas* se documenta en español en el siglo XIX, con dos significados atribuidos a *cancamusa* y *recancamusa* por parte del *DEA* y del *DLE*: el de ‘Monserga (cosa fastidiosa y que no merece atención)’ (*DEA*), el de ‘Dicho o hecho con que se pretende desorientar a alguien para que no advierta el engaño de que va a ser objeto’ (*DLE*) y, para *recancamusa*, el de ‘Mañana o artificio para encubrir un engaño’ (*DLE*). Así pues, se puede afirmar con bastante certeza que *carcamusa* era una variante (posiblemente vulgar) de *cancamusa*.

No es hasta mediados del siglo XX cuando se vuelve a atestiguar su uso de nuevo haciendo alusión al plato toledano. Dada la poca documentación de *carcamusa(s)*, resulta difícil de explicar el cambio semántico de la palabra y, por lo tanto, asegurar la relación etimológica entre los usos decimonónicos y el uso actual. Sin embargo, es una palabra demasiado peculiar como para considerar que se trate de un fortuito caso de homonimia.

En definitiva, la etimología de *carcamusa(s)* como plato de comida sigue sin estar del todo clara, pero lo más probable es que, de manera directa o indirecta, provenga de los usos decimonónicos de *carcamusa*, que a su vez es una variante del sustantivo *cancamusa*. Quizás un estudio pormenorizado sobre las voces *cancamusa* y *recancamusa* o un afortunado descubrimiento de nuevos testimonios de la variante *carcamusa* en la primera mitad del siglo XX puedan arrojar luz sobre estas cuestiones y confirmar la propuesta etimológica que aquí se ha defendido.

Por el momento, hasta aquí llega lo que sabemos, lo que suponemos y lo que imaginamos del origen y la historia de *carcamusas*.

REFERENCIAS BIBLIOGRÁFICAS

- Adelstein, A. (2018). Tipologías de neología semántica nominal. En R. Bein, J. E. Bonnin, M. di Stefano, D. Lauria, M.^a C. Pereira (coords.), *Homenaje a Elvira Arnoux: Estudios de análisis del discurso, glotopolítica y pedagogía de la lectura y la escritura. TOMO III: Lectura y escritura* (pp. 17-36). Universidad de Buenos Aires.
- Alonso, J. L. (07 de enero de 2026). *Receta de carcamusas: Un plato típico de Toledo (2026)*. Leyendas de Toledo. <https://www.leyendasdetoledo.com/carcamusas-un-plato-tipico-de-toledo/> [Fecha de consulta: 03 de febrero de 2026].
- Álvarez de Miranda, P. (2009). Neología y pérdida léxica. En E. de Miguel Aparicio (ed.), *Panorama de la lexicología* (pp. 133-156). Ariel.
- Archiletras. (14 de mayo de 2025). *Diccionario gastronómico-literario de la cocina tradicional española (de la A a la Z)*. <https://www.archiletras.com/gastronomico/diccionario-gastronomico-literario-de-la-cocina-tradicional-espanola-de-la-a-la-z/> [Fecha de consulta: 03 de febrero de 2026].
- Aromarti. (1 de noviembre de 2005). *Del romance a la lengua universal*. Celtiberia.net. <https://www.celtiberia.net/es/biblioteca/?id=1595#r21429> [Fecha de consulta: 03 de febrero de 2026].
- Bavia, L. (1610). *Historia de la union del reyno de Portugal a la corona de Castilla: de Geronimo Franchi Conestagio [...] traducida de lengua italiana en nuestra vulgar castellana*. Sebastian de Cormellas. <https://digibug.ugr.es/handle/10481/24332>
- Bessé, G. (1645). *Histoire des comtes de Carcassonne*. Arnaud Estradier. https://www.google.es/books/edition/Histoire_des_Comtes_de_Carcassonne/khxDAAAACAAJ
- BRASILHIS Database. (2015). Diogo de Cárcamo (Carcome). <https://brasilhis.usal.es/es/personaje/diogo-de-carcamo-carcome> [Fecha de consulta: 03 de febrero de 2026].
- Calmet, A. (1746). *Il tesoro delle antichità sacre e profane* (2.^a ed., t. II). Francesco Pitteri / Dionisio Ramanzini. https://www.google.es/books/edition/Il_tesoro_delle_antichità_sacre_e_profana/qtg4u3vMMOgC
- Campos Souto, M. (2018a). Bibliotecas y hemerotecas digitales en el NDHE. *Cuadernos del Instituto Historia de la Lengua*, 11, 237-255. <https://doi.org/10.58576/cilengua.vi11.56>
- Campos Souto, M. (2018b). Las bases documentales del NDHE: Entre la realidad y el deseo. En D. Corbella, A. Fajardo & J. Langenbacher-Liebgott (eds.), *Historia del léxico español y Humanidades digitales* (pp. 19-46). Peter Lang.
- Castilla-La Mancha Media. (22 de abril de 2016). *Carcamusas, plato oficial de Toledo Capital Gastronómica 2016* [Vídeo]. YouTube. https://www.youtube.com/watch?v=O6oyfd-qc7_M [Fecha de consulta: 03 de febrero de 2026].
- Carcamusas (13 de julio de 2015). En *Wikipedia*. <https://es.wikipedia.org/w/index.php?title=Carcamusa&oldid=83757039> [Fecha de consulta: 03 de febrero de 2026].

- Carcamusas. (16 de enero de 2026) En *Wikipedia*. https://es.wikipedia.org/w/index.php?title=Carcamusa&oldid=171545694#Enlaces_externos [Fecha de consulta: 03 de febrero de 2026].
- Corominas, J. & Pascual, J. A. (1985). *Diccionario crítico etimológico castellano e hispánico. Tomo IV*. Gredos.
- Eberenz, R. (2014). El léxico español de la alimentación y la culinaria en su historia: fuentes y líneas de investigación. En V. Álvarez Vives, E. Díez del Corral Areta & N. Reynaud Oudot (eds.), *Dándole cuerda al reloj. Ampliando perspectivas en lingüística histórica de la lengua española* (pp. 23-46). Tirant Humanidades.
- El Clamor Público*. (24 de abril de 1855). *Sección política* (N.º 5299). <https://hemerotecadigital.bne.es/hd/es/viewer?id=75dfb8ae-99af-48f3-a201-b9caf3e81aad>
- El Vapor*. (22 de junio de 1852). *La comisión especial nombrada por la Sociedad Económica Matritense...* (N.º 220). Madrid. <https://hemerotecadigital.bne.es/hd/es/viewer?id=8c5d2529-07a7-43b4-94e3-d98e1d8fb12a&page=2>
- Fauchet, C. (1611). *Les antiquités et histoires gauloises et françoises*. Paul Marceau. https://www.google.es/books/edition/Les_antiquités_et_histoires_gauloises_e/9MwWAAAAQAAJ
- Fernández González, M. (1866). *El collar del diablo: Memorias de un resucitado*. Imprenta y librería de Miguel Guijarro. https://www.google.es/books/edition/El_collar_del_diablo/3qt-AAAAIAAJ
- Fresne, C. (1710). *Glossarium ad scriptores mediae et infimae latinitatis*. Oficina Zuneriana de Johan Adam Juncio. https://www.google.es/books/edition/Glossarium/y_DAB-mYOozwC?hl=es&gbpv=1
- García Manga, M.^a C. (2006). La Etimología Popular en su contexto: en torno a los factores que favorecen su aparición. En M.^a L. Calero Vaquera, F. Osuna García & A. Zamorano Aguilar (eds.), *Studia lingüística et philologica* (pp. 101-122). Servicio de Publicaciones Universidad de Córdoba.
- Guindilla*. (1 de septiembre de 1842). *El glorioso* (N.º 14). <https://www.google.es/books/edition/Guindilla/VHHnIPXP7N0C?pg=PA214>
- Hernández Alonso, J. (Koketo). (05 de septiembre de 2018). *Carcamusa: La leyenda del encuentro entre viejos y doncellas de Toledo*. Koketo.es. <https://koketo.es/carcamusa> [Fecha de consulta: 03 de febrero de 2026].
- James, C. (1805). *A new and enlarged military dictionary*. T. Egerton. https://www.google.es/books/edition/A_New_and_Enlarged_Military_Dictionary/tOypDj9DoZMC
- La Prensa Alcarreña*. (12 de febrero de 1984). *Gastronomía y vinos* (N.º 705). <https://prensahistorica.mcu.es/es/consulta/registro.do?id=10005090861>
- Lipsio, J. (1602). *Analecta sive observationes reliquae ad militiam et bosce libros*. Ioannem Moretum. https://www.google.es/books/edition/Iusti_Lipsii_Poliorceticon_siue_De_machi/KaxlFsk07hAC

- Marín, J. M. (26 de mayo de 2017). En *Discusión: Carcamusas*. Wikipedia. <https://es.wikipedia.org/w/index.php?title=Discusi%C3%B3n:Carcamusa&oldid=99405137> [Fecha de consulta: 03 de febrero de 2026].
- Mélida Ramón, J. (1884). *El demonio con faldas: Memorias de un gato*. Ricardo Fé. <https://bnedigital.bne.es/bd/card?oid=0000191763>
- Mélida Ramón, J. (1887). *A orillas del Guadarza*. Biblioteca, Arte y Letras. https://www.google.es/books/edition/%C3%81_orillas_del_Guadarza/d20vAAAAYAAJ
- Moratilla, P. (2 de febrero de 2022). *Las carcamusas de Toledo*. Agua, Luz y Viento. <https://agualuzyviento.wordpress.com/2022/02/02/las-carcamusas-de-toledo/> [Fecha de consulta: 03 de febrero de 2026].
- Moretti, F. (1828). *Diccionario militar español-francés*. Imprenta Real. https://www.google.es/books/edition/Diccionario_militar_espa%C3%B1ol_franc%C3%A9s/c1CEPbC4l9kC
- Pithou, P. (1588). *Annalium et historiae Francorum*. París: Claudium Chappelet. https://www.google.es/books/edition/Annalium_et_historiae_francorum/AQPhBgT2D9sC
- Real Academia Española. (1729). *Diccionario de autoridades. Tomo segundo*. Imprenta de Francisco del Hierro. <https://apps2.rae.es/ntlle/SrvltGUILoginNtlle> [Fecha de consulta: 03 de febrero de 2026].
- Real Academia Española. (2001). *Nuevo tesoro lexicográfico de la lengua española (NTLLE)*. <http://ntlle.rae.es> [Fecha de consulta: 03 de febrero de 2026].
- Real Academia Española. (2009). *Nueva gramática de la Lengua Española*. Espasa. <https://www.rae.es/gram%C3%A1tica/>
- Real Academia Española. (2013-). *Corpus del Diccionario histórico de la lengua española (CDH)*. <https://apps.rae.es/CNDHE> [Fecha de consulta: 03 de febrero de 2026].
- Real Academia Española. (s. f.). *Corpus Diacrónico del Español (CORDE)*. <https://corpus.rae.es/cordenet.html> [Fecha de consulta: 03 de febrero de 2026].
- Real Academia Española. (s. f.). *Corpus de Referencia del Español (CREA)*. <https://corpus.rae.es/creanet.html> [Fecha de consulta: 03 de febrero de 2026].
- Real Academia Española. (s. f.). *Corpus del español del siglo XXI (CORPES XXI)*. <https://www.rae.es/corpes/> [Fecha de consulta: 03 de febrero de 2026].
- Real Academia Española. (s. f.). *Diccionario de la lengua española* (23.^a ed., versión 23.8). <https://dle.rae.es> [Fecha de consulta: 03 de febrero de 2026].
- Rincón, G. F. (1805). *Colección de poesías póstumas de D. Manuel Rincón*. Gómez Fuentenebro y Compañía. https://www.google.es/books/edition/Coleccion_de_poesias_posthumas/pDJEAAAACAAJ
- Rodríguez-Navas, M. (1906). *Diccionario completo de la lengua española*. Saturnino Calleja. <https://archive.org/details/diccionariocompl00rodruft/page/286/mode/2up?q=cancamusa>

- Rost Bagudanch, A. (2009). La evolución de Nj: consideraciones diacrónicas en el marco de la Fonología Evolutiva. En J. R. Cufí (ed.), *Diachronic Linguistics* (pp. 165-190). Documenta Universitaria.
- Rost Bagudanch, A. (2016). El extraño caso de la rótica infiltrada. En A. M.^a Fernández Planas (ed.), *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística* (pp. 109-117). Universidad de Barcelona.
- Rost Bagudanch, A. (2024). Rotacismo de las consonantes coronales /n, l, d, s/ en español. *Onomázein. Revista de lingüística, filología y traducción*, 64(1), 112-136. <https://doi.org/10.7764/onomazein.64.06>
- RTVE. (14 de julio de 2023). *Receta de carcamusas: El guiso de carne que se deshace en la boca*. RTVE.es. <https://www.rtve.es/television/20230714/receta-carcamusas-guiso-toledo-carne-guisantes-facil-sabroso/2452019.shtml> [Fecha de consulta: 03 de febrero de 2026].
- Seco, M., Andrés, O. & Ramos, G. (2023). *Diccionario del español actual* (3.^a ed., en línea). Fundación BBVA. <https://www.fbbva.es/diccionario/info/guia-usuario/> [Fecha de consulta: 03 de febrero de 2026].
- Torres Martínez, M. (2022). *Léxico culinario en el Libro de cuentas de cocina y repostería de la Casa de Arcos (1750)*. *Cuadernos AISPI*, 20, 114-140. <https://doi.org/10.14672/2.2022.2043>
- Varios autores. (1603). *Hispaniae illustratae* (t. II). Frankfurt: Claudium Marniuu y herederos de Iohannis Aubrij. https://www.google.es/books/edition/Hispaniae_illvstrat%C3%A6_sev_rervm_vrbivmq/iY3KiMLG8VcC



ORIENTACIONES PARA INSTRUIR AL PROFESORADO DE E/L2 EN LA ENSEÑANZA DE VERBOS FRECUENTES: EL CASO DE 'TENER'

GUIDELINES FOR INSTRUCTING SFL TEACHERS ON THE HIGHG FREQUENCY VERBS: THE CASE OF 'TO HAVE'

Natalia Alarcón Mosquera
Universidad de Castilla – La Mancha
Natalia.Alarcon@uclm.es

RESUMEN

El objetivo general de este artículo es orientar la instrucción del profesorado de Español como Segunda Lengua o Lengua Extranjera (E/L2) sobre un tema poco habitual en los programas de másteres y cursos de actualización pedagógica: la enseñanza de verbos frecuentes (*dar, hacer, tener, poner*). Para su consecución, el trabajo alberga dos objetivos específicos. El primer objetivo estriba en la descripción de unas pautas para el tratamiento de los verbos que no aparecen en el diccionario combinatorio *Redes* (Bosque, 2004), como *tener, hacer* o *poner*. El segundo objetivo consiste en ilustrar el procedimiento anterior mediante un análisis del verbo *tener* basado en la tesis doctoral *El verbo tener en el Corpus Oral de Aprendientes de Español* (Alarcón, 2026). Así, tras la definición del concepto de *verbo frecuente* (VF), la parte central del trabajo se divide en las tres fases en las que se distribuye el proceso del primer objetivo: reducción del valor polisémico del VF, definición simplificada e integradora del VF que integre los sentidos prototípico y esquemático y, por último, análisis de errores frecuentes (**hacer sentido* en lugar de *tener sentido*, por ejemplo). Adicionalmente, cada uno de los apartados o fases se completa con el ejemplo del verbo *tener*. El estudio pretende contribuir al desarrollo de la incipiente área de estudio dedicada a la enseñanza de verbos frecuentes (Jiménez Calderón y Rufat, 2019; Rufat, 2019, 2023; Jiménez Calderón y Sanz, 2023). En concreto, en este trabajo se dirige el foco hacia la necesaria formación del profesorado como andamiaje imprescindible para acortar la distancia entre el conocimiento teórico de estos VF y la explotación didáctica en el aula de E/L2.

Palabras clave: E/L2, verbos frecuentes, verbo *tener*.

ABSTRACT

The aim of this article is to provide guidance for Spanish as a Foreign Language teachers on an issue that it is not often addressed in master's programs or training teachers courses: the instruction of high-frequency verbs (*dar, hacer, tener, poner*). To achieve this, the paper has two specific objectives. The first objective lies in describing guidelines for the treatment of verbs that do not appear in the combinatorial dictionary *Redes* (Bosque, 2004), such as *tener, hacer*, or *poner*. The second objective is to illustrate the above procedure through an examination of the verb *tener* based on the doctoral dissertation 'The verb TO HAVE in the Oral Corpus of Spanish Learners (CORADE)' (Alarcón, 2026). Thus, after defining the concept of a high-frequency verb, the core part of the paper is divided into the three phases in which the process of the first objective is distributed: reduction of the polysemic value of the high-frequent verb, simplified and integrative definition of the high-frequency verb that integrates both prototypical and systemic meanings, and finally, analysis of commons errors (e.g., **hacer sentido* instead of *tener sentido*). Furthermore, each section or phase is completed by an example using the verb *tener*. The study intends to foster the growth of the emerging field focused on the teaching of high-frequent verbs (Jiménez Calderón & Rufat, 2019; Rufat, 2019, 2023; Jiménez Calderón & Sanz, 2023). More precisely, this paper directs its focus toward the crucial development of teacher training as an essential scaffolding to shorten the distance between the theoretical knowledge of high-frequency verbs and their practical application in the SFL classrooms.

Keywords: SFL, high-frequency verbs, to have.

Recibido: 01-05-2026
Aceptado: 10-06-2026

DOI: <https://doi.org/10.17561/rilex.9.2.10444>



1. VERBOS FRECUENTES: DELIMITACIÓN DEL TÉRMINO E INTERÉS DIDÁCTICO EN EL AULA DE E/L2¹

Los verbos frecuentes (VF) han sido objeto de estudio en un número razonable de trabajos teóricos y aplicados sobre Adquisición de Segundas Lenguas como, por ejemplo, los de Granger (2002) o Guo (2006) sobre las formas *to make*, *to take* y *to keep* o los de Li (2022), Sawaguchi y Mizumoto (2022) o López Pérez y Benali Taouis (2019) sobre *to take*, *to make* y *to do*, respectivamente. En el ámbito del español, destaca el trabajo de Rufat (2019) sobre el verbo *dar*, junto con los de Molina (2013) sobre *dar*, *tener* y *hacer* en hablantes de italiano o los estudios comparativos entre español e inglés de Tsutahara de, entre otros, *ir* y *to go* (2020). Si bien, hasta donde se nos alcanza, no existe una definición del término *verbo frecuente*, en la bibliografía es habitual recurrir, en primer lugar, a trabajos empíricos que avalen de manera fiable la frecuencia de uso de los posibles verbos que componen el repertorio. Así, del trabajo de Jiménez Calderón y Rufat (2019, p. 134) se desprende que un rasgo para pertenecer a la categoría de VF consiste en ocupar una de las cincuenta primeras posiciones de la clasificación correspondiente del diccionario de frecuencias de Davies y Davies (2018) (Tabla 1).

1) ser	11) saber	21) seguir	31) conocer	41) empezar
2) haber	12) querer	22) quedar	32) tomar	42) buscar
3) tener	13) pasar	23) llevar	33) sentir	43) existir
4) estar	14) deber	24) encontrar	34) tratar	44) perder
5) hacer	15) llegar	25) pensar	35) vivir	45) escribir
6) decir	16) creer	26) volver	36) acercar	46) realizar
7) poder	17) dejar	27) salir	37) esperar	47) entrar
8) ir	18) parecer	28) venir	38) gustar	48) leer
9) ver	19) hablar	29) llamar	39) contar	49) recordar
10) dar	20) poner	30) mirar	40) trabajar	50) morir

TABLA 1: Verbos más frecuentes en español según Davies y Davies (2018)

¹ En el presente trabajo, el concepto E/L2 alude al español no nativo tanto en contexto de inmersión o L2 como en el de lengua extranjera o LE.

Además de la frecuencia estadística, el segundo rasgo de estas piezas es su polisemia. Se trata de una característica estrechamente relacionada con la anterior, dada la extensa evidencia de que, como afirma Pinker (2007, p. 163), “cuanto más frecuente es una palabra, más polisémica es, y vice-versa”. Prueba de ello es el elevado número de acepciones que aparecen en las entradas de diccionarios de carácter general sobre las formas de la Tabla 1. Para sustentar lo anterior, en la segunda columna de la Tabla 2 aportamos el número de acepciones que aparecen en las entradas del diccionario de la RAE de los diez verbos más frecuentes según Davies y Davies (2018).

Verbo	Número de acepciones en el <i>DRAE</i>
ser	16
haber	11
tener	25
estar	28
hacer	58
decir	13
poder	6
ir	42
ver	22
dar	54

TABLA 2: *Número de acepciones de los 10 verbos más frecuentes de Davies y Davies (2018)*

En tercer lugar, es preciso subrayar que, a pesar de que aparecen con mucha frecuencia fuera y dentro del aula de E/L2, el conocimiento productivo o capacidad para usar con corrección determinados usos de los VF constituye un área de dificultad para los hablantes no nativos. Como señalan Altenberg y Granger (2001, p. 174), esto se debe a que “despite fundamental semantic similarities across languages, high-frequency verbs have developed many language-specific differences, which make them treacherous for foreign language learners”. Por ejemplo, la especificidad de *tener* podría explicar que, durante los primeros estadios de aprendizaje, estudiantes anglófonos cometan errores

en la expresión de la edad como el de (1) motivados por la transferencia del verbo *to be* empleado en inglés para esta función.

(1) *Soy 20 años [Tengo].

En definitiva, podemos definir los verbos frecuentes como piezas verbales de alta frecuencia en el discurso de hablantes nativos —ver Rufat (2019) y Alarcón (2026) para las frecuencias de *dar* y *tener* en hablantes no nativos, respectivamente— que se caracterizan por su capacidad para albergar numerosos significados y por actualizar determinados usos que constituyen un área de dificultad para aprendices de lenguas adicionales. A continuación, describimos los dos rasgos relevantes para la enseñanza-aprendizaje de estas formas verbales en el aula de segundas lenguas: su polisemia y sus restricciones combinatorias.

1.1. TRATAMIENTO DE LA POLISEMIA EN E/L2

La enseñanza comunicativa en general ha priorizado tanto el acceso contextualizado al léxico como su correspondiente análisis situado. Según este planteamiento, un uso figurado de *congelar* como *congelar el salario* podría adquirirse de manera incidental al integrarse en un texto sobre el empleo. De acuerdo con los enfoques léxicos —ver Pérez Serrano (2017, pp. 29-40) para una revisión de las distintas corrientes—, además, es recomendable que la exposición anterior se combine con actividades dirigidas y orientadas a que el alumnado consolide las diferentes combinaciones frecuentes del texto —por ejemplo, *congelar el salario*, *quedarse en el paro*, *contrato indefinido*—. En definitiva, frente al aprendizaje descontextualizado de corrientes metodológicas anteriores, este enfoque incorpora una rentabilidad clara para el aprendizaje del vocabulario en general en términos comunicativos.

Sin embargo, también resulta evidente que ese —necesario— comunicativismo limita la posibilidad de que los aprendientes perciban la coherencia interna de las palabras polisémicas (Vicente Balenciaga & Goethals, 2026). En ese sentido, cabe preguntarse si para comprender y saber usar *congelar el*

salario es necesario que el aprendiente adquiera de manera intencionada la relación entre este uso figurado de *congelar* —predecible debido a su transparencia— y el uso literal de, por ejemplo, *congelar la carne*. A tenor de los numerosos estudios sobre el valor pedagógico de presentar explícitamente la motivación semántica de las palabras polisémicas en inglés (Boers & Lindstromberg, 2008; Littlemore, 2023); los análisis sobre competencia metafórica y verbos de percepción y movimiento en español (Masid Blanco, 2022; Ibarretxe-Antuñano, 2006; Hijazo-Gascón, 2021, respectivamente), así como de los trabajos que apuntan hacia la dificultad de los aprendientes para integrar en su conocimiento productivo los usos menos prototípicos de la palabra polisémica en cuestión (Laufer, 1997; Qu, 2023), no podemos desdeñar la posible rentabilidad de capacitar al alumnado para que detecte el vínculo semántico que conecta las distintas acepciones de las palabras. Del mismo modo, tampoco es posible asegurar que el profesorado posea este conocimiento en lo relativo a los VF. En otras palabras, los docentes previsiblemente dispondrán —o podrán desarrollar— los recursos y la intuición suficientes para abordar de manera explícita los usos de *congelar* mencionados. No obstante, parece razonable asumir que el análisis de palabras polisémicas que activan significados menos transparentes, como sucede con los VF (por ejemplo, *dar un paseo* o *tener una pelea*), requiere de la instrucción o el estudio.

1.2. LAS RESTRICCIONES COMBINATORIAS DE LOS VF: LUCES Y SOMBRAS DE SU APLICACIÓN EN LA ENSEÑANZA DE E/L2

En muchos de los usos de VF menos transparentes, los verbos funcionan como verbos de apoyo (VA). En oposición a los verbos plenos (VP), se denominan *verbos de apoyo*, *verbos ligeros* o *verbos soporte* a aquellas formas que en apariencia se limitan a aportar información flexiva, tal y como se ha propuesto en buena parte de la literatura. En (2) se incluyen algunos ejemplos de construcciones en las que VF participan en construcciones con verbos plenos (CVP), mientras que en (3) aparecen construcciones con verbos de apoyo (CVA) con las mismas formas verbales. Entre corchetes incluimos la

traducción en inglés para ilustrar cómo, desde la perspectiva interlingüística, los usos plenos tienden a coincidir en las distintas lenguas, mientras que en las CVA se concentra un mayor número de diferencias.

(2)

- a. Eva tiene una casa. [*Eva has a house*]
- b. Luis me dio un libro. [*Luis gave me a book*]
- c. Ana puso el libro en la mesa. [*Ana put the book on the table*]

(3)

- a. Jorge tiene cuidado cuando cruza la calle. [*Jorge is careful when he crosses the street*]
- b. Sara da un paseo. [*Sara took a walk*]
- c. Ana puso una excusa para no ir a la fiesta. [*Ana made an excuse not to go to the party*]

Las construcciones con verbo de apoyo (CVA) han sido analizadas desde diferentes enfoques como la fraseología (Alonso Ramos, 2004, 2010; Koike, 2001; Corpas Pastor, 1996), la teoría Sentido-Texto (Barrios, 2015) o la gramática generativa (Bosque, 2001; De Miguel, 2008). Por nuestra parte, el presente trabajo se encuentra alineado con los planteamientos generales de Bosque y De Miguel en cuanto que asumimos que los VA, lejos de perder su significado cuando participan en CVA (De Miguel, 2008), especifican su interpretación mediante un proceso de concordancia de rasgos léxicos con los complementos a los que acompañan. Se trata de una operación similar a la que encontramos en la aceptabilidad de *cometer un error*, donde tanto el verbo como el sustantivo incorporan el rasgo [+negativo] frente a la falta de concordancia con sustantivos de polaridad positiva como sucede, por ejemplo, en **cometer un acierto*. A propósito de lo anterior, conviene subrayar que esta concordancia o redundancia entre las dos piezas no representa anomalía alguna, antes bien se trata de un mecanismo altamente frecuente en la lengua común. Sirvan como ejemplo de lo anterior los ejemplos del tipo

imponer unilateralmente, chirrido estridente o insistir repetidamente, etc. aportados por Bosque (2004, pp. CXXXVIII-CXL) en el prólogo del diccionario combinatorio *Redes*, así como la redundancia morfológica de *interponerse entre* o la de *si* en la interrogación indirecta en *dilucidar si*.

La tesis de la selección léxica anterior ha sido ampliamente aceptada en la literatura de las últimas décadas. De hecho, en trabajos fundamentales en el ámbito de la fraseología como el estudio sobre CVA de Alonso Ramos (2004) se admite la existencia de concordancia, aunque se dude de la relevancia de esta información:

No excluyo la posibilidad de que descomponiendo semánticamente los nombres de las CVA y estableciendo la jerarquía de sus componentes genéricos, podamos llegar a algún tipo de predicciones sobre la afinidad de tal grupo de nombres con tal verbo de apoyo. Ahora bien, la descripción sería tan compleja y tan hiperspecificada que no sería útil ni para el aprendiz de una segunda lengua ni para un sistema de tratamiento automático (Alonso Ramos, 2004, p. 191).

Respecto a la utilidad de examinar determinados procesos de selección léxica, trabajos teóricos y aplicados recientes como los de Rufat (2019, 2023) sobre el verbo *dar* o el de Jiménez Calderón y Sanz (2023) sobre *llevar* ponen de manifiesto que el hecho de que el alumnado conozca al menos una parte de la información relativa a las restricciones combinatorias de los VF puede favorecer el aprendizaje de E/L2. Dicho de otro modo, trabajos recientes abogan no solo por dirigir la atención del aprendiz hacia errores habituales como el de (4), sino por fomentar la comprensión de las causas que lo han motivado.

④ ¿Vamos a *tomar un paseo? [dar]

Para contextualizar el alcance del proceso de selección léxica que, por ejemplo, aprueba *dar un paseo* y censura **dar una ruta senderista* o **dar un viaje*, es necesario distinguir entre restricciones combinatorias extralingüísticas o basadas en el conocimiento del mundo, e intralingüísticas o basadas en las reglas del sistema lingüístico. Las primeras se rigen por nociones de naturaleza

física; por consiguiente, es posible construir su paradigma extensional de acuerdo con el conocimiento que tenemos del mundo, como en *dar un regalo* (Bosque, 2001; Rufat & Jiménez Calderón, 2019). En concreto, tal como detalla Rufat (2019, pp. 47-49), la restricción de tipo extralingüístico que impone *dar* en su sentido básico *de entregar, hacer llegar, proporcionar* exige en su interpretación literal la coaparición con sustantivos que designan entidades materiales. Dicho de otra manera, en las combinaciones de tipo extralingüístico, el conocimiento enciclopédico es suficiente para deducir con qué otros nombres podría coaparecer el verbo. Por el contrario, este tipo de saber es insuficiente para explicar por qué la construcción *dar un paseo* es aceptable en el español europeo. La respuesta se encuentra en las reglas internas del sistema, según las cuales, *dar* se combina con *paseo* porque el verbo selecciona, entre otros, el rasgo subléxico [+desplazamiento breve]. Habida cuenta de que este rasgo también forma parte de las propiedades de *paseo*, se establece el proceso de concordancia o redundancia mencionado supra y, por consiguiente, la construcción resulta aceptable —a diferencia de la coaparición de *dar* con otros sustantivos que implican desplazamientos largos, como *ruta* o *viaje*—.

Así, en respuesta a las dudas sobre la rentabilidad de instruir al alumnado de E/L2 en restricciones de *dar* como las descritas, consideramos razonable afirmar que este conocimiento puede, por un lado, contribuir a un aprendizaje más significativo del verbo y, por otro, facilitar la identificación de los vínculos semánticos que conectan la red de sentidos que albergan las palabras polisémicas. Además, si consideramos que buena parte de esta información se encuentra en el diccionario combinatorio *Redes*, la tarea resulta, además, relativamente abarcable. Ahora bien, el gran inconveniente estriba en que, tal como han observado Jiménez Calderón y Rufat (2019, p. 142), solo siete de los cincuenta verbos más frecuentes (*dar, seguir, llevar, tomar, sentir, perder, entrar*) aparecen en *Redes*.

En suma, de lo anterior se desprende que en el caso de los verbos que no aparecen en *Redes* resulta complicado satisfacer lo expuesto en 1.1 y

1.2: dotar al profesorado de los conocimientos necesarios para abordar de manera explícita en el aula de E/L2 (i) las conexiones semánticas entre las distintas acepciones de los VF y (ii) las restricciones combinatorias de aquellos usos que presenten mayor dificultad para al alumnado. Para contribuir a superar estas limitaciones, en las páginas siguientes describimos un marco de actuación enfocado en instruir al profesorado en el abordaje de estos verbos. A modo de ilustración del procedimiento, presentamos el análisis del VF *tener* expuesto en Alarcón (2026).

2. VERBOS QUE NO APARECEN EN *REDES*

Para facilitar el análisis de los VF que no aparecen en *Redes*, en primer lugar, es pertinente preguntarse por la causa de su ausencia. A ese respecto, entendemos que el hecho de que verbos como *tener*, *poner* o *hacer* no cuenten con una entrada en el diccionario combinatorio se debe a la dificultad de caracterizar su extensión mediante una propiedad compartida por un conjunto de entidades, al igual que *interesante* (Bosque, 2004, pp. LXXXIX-XCI). Con todo, la justificada ausencia en *Redes* de estas palabras de gran extensión —y, por tanto, poca intensión—no contradice el interés didáctico de conocer las exigencias combinatorias de, por ejemplo, *tener* cuando no actualiza su sentido prototípico de *posesión*, como en *tener cuidado* en (3a). De hecho, esta limitación inicial promueve análisis complementarios al de la selección de clases léxicas que permiten detectar otras áreas específicas de interés para la enseñanza de algunos de estos VF. Es el caso de, por ejemplo, la diferencia de foco entre el valor posesivo de *tener* en (5) frente al existencial de *haber* en (6).

(5) La casa tiene dos baños.

(6) En la casa hay dos baños.

El mapa de redes combinatorias que encontraríamos en una hipotética entrada de *tener* en *Redes* no aclararía la distinción entre (5) y (6). En otras palabras, saber

que el verbo *tener* se combina con la clase de sustantivos que denotan partes de un lugar no facilitaría la labor docente en el tratamiento de errores habituales en sinohablantes de niveles iniciales como el de (7). La desviación de (7) se debe a que en chino se selecciona el mismo verbo (yǒu) para (5) y (6).

(7) La casa *hay dos baños [tiene].

Para evitar que determinados usos de los VF de interés para el aula de E/L2 como los de (5) y (6) resulten excluidos de la propuesta, el marco de actuación para la formación del profesorado ha de concebirse como un instrumento concreto, pero también flexible. A ese respecto, consideramos que una eventual secuencia didáctica sobre VF dirigida a docentes ha de constar de las definiciones simplificadas de los verbos en cuestión. Además, en estas definiciones debe aparecer (i) el significado unitario presente en todos los usos (ii) junto con tres o cuatro sentidos generales que engloben las múltiples acepciones atribuibles a estos verbos polisémicos. Con ello, el futuro docente de E/L2 será capaz comprender en profundidad la red semántica de estos verbos y ofrecer explicaciones más precisas que trasciendan el análisis interlingüístico. Por ejemplo, el docente ofrecerá una mejor retroalimentación ante errores como el de (8) si previamente ha interiorizado que, como veremos en el apartado 2.3, este uso se adscribe a uno de los valores más generales de *tener*: el de la experimentación.

8) ITEX1. El viaje es la única cosa que te permite *hacer experiencias² [*tener*].

En síntesis, como marco general para la instrucción a docentes sobre la enseñanza de los VF que no aparecen en *Redes* proponemos un procedimiento distribuido en tres fases y sustentado en la propia experiencia docente tanto en la enseñanza de E/LE como en la formación de profesorado. El objetivo de

² Todos los ejemplos precedidos de un código alfanumérico proceden del Corpus Oral de Aprendientes de Español (CORADE) (Alarcón, 2026).

la primera fase consiste en reducir el número de acepciones de los VF seleccionados; por su parte, el propósito de la segunda fase radica en formular una definición unitaria que integre todos los usos de cada una de las formas meta; por último, la tercera fase se dedica a dirigir la atención del profesorado hacia las construcciones problemáticas. A continuación, describimos cada una de las fases a partir del análisis de *tener*.

2.1. FASE 1: REDUCCIÓN DE LA POLISEMIA

Uno de los procedimientos para identificar los tres o cuatro sentidos que actúan como hiperónimos de la extensa red de significados del verbo estriba en tratar de agrupar las distintas acepciones presentes en los diccionarios de carácter general estableciendo jerarquías entre las distintas interpretaciones. Gracias a los recursos tecnológicos actuales y a su rápido avance, el análisis de corpus constituye una manera complementaria o alternativa de obtener esta reducción semántica. En el caso de *tener*, nos hemos decantado por esta última opción. En concreto, la simplificación del verbo posesivo se articula en torno a los trabajos de Aguilar Caro (2007), basado en un corpus de hablantes nativos, y de Alarcón (2026), enfocado en un corpus de aprendientes de E/L2. Antes de describir estos estudios, nos detenemos a caracterizar someramente al verbo *tener*:

Tener es el verbo canónico para la expresión de relaciones posesivas en español, de ahí su capacidad para actualizar diferentes tipos de relaciones que, en última instancia, se acogen a la noción de *posesión* en el sentido más amplio del término. Esta interpretación flexibiliza el sentido restringido de Taylor (1996, p. 340), según el cual el poseedor es un ser humano concreto o, entre otras características, lo poseído es una entidad inanimada. De este modo, el verbo *tener*, además de aludir a la propiedad de algo (a) en su uso idiosincrásico, expresa relaciones de parentesco (b), cualidad (c), parte/todo (d), estados físicos (e) y mentales (f), entre otras.

- a. Ana tiene una casa.
- b. Pedro tiene un hermano.

- c. Eva tiene paciencia.
- d. La casa tiene garaje.
- e. Luis tiene sed.
- f. Pedro tiene miedo.

Asimismo, *tener* cumple las características de la definición inicial de VF: es un verbo altamente polisémico —las entradas dedicadas al verbo *tener* en diccionarios de carácter general presentan más de veinte acepciones— que alberga algunos usos de cierta dificultad para hablantes no nativos cuando no encuentra correspondencia plena en otras lenguas, como el expuesto a propósito de *haber* y *tener* en chino, así como en determinadas de CVA, como se vio en (2) a propósito de *tener cuidado*.

En relación con el primer trabajo referido, Aguilar Caro (2007) se ocupa de analizar las construcciones V + N en las que participa *tener* atendiendo al rasgo de inherencia del sustantivo. Tras evaluar las ocurrencias del *Corpus del Español Mexicano Contemporáneo* (CEMC), la autora concluye que la propiedad de inherencia no es discreta, sino que se manifiesta en un *continuum* en el que los poseídos se adscriben a una de estas tres categorías: atributos, entes del entorno y, por último, acciones o eventos. Como resultado adicional de la investigación, la autora identifica los siguientes tres grandes valores de *tener*: la caracterización, cuando se combina con lo que considera *atributos*; la pertenencia, en compañía de los denominados *entes del entorno*, y la experimentación, junto con lo que Aguilar Caro define como *acciones y eventos*.

Los *atributos* cumplen la función de describir al poseedor mediante la calificación y representan el mayor grado de inherencia respecto al poseedor de toda la escala (Aguilar Caro, 2007, p. 56). La autora distingue entre atributos concretos y abstractos. Los primeros engloban entidades que pueden ser percibidas con los sentidos físicos (9); por su parte, los segundos aluden a entidades cuya percepción requiere de una elaboración mental que trasciende el campo de acción de los sentidos, lo que podría considerarse como abstracto (109).

- 9) Jaime tiene lunares.
- 10) Los niños bilingües tienen ventajas de todo tipo.

En segundo lugar, los predicados de la categoría *entes del entorno* comparten el vínculo con el entorno físico (11), social (12) o económico (13) del poseedor, por lo que el repertorio de sustantivos es más heterogéneo que en el grupo anterior.

- 11) La ciudad tiene un templo.
- 12) Ramona tiene razón.
- 13) Luis no tiene mucho dinero.

Según Aguilar Caro (2007, p. 67), los miembros de este grupo presentan menor fijeza en su grado de inherencia, dado que pueden entrar o salir del dominio posesivo con la rapidez o frecuencia que su naturaleza les confiere; es decir, su estabilidad está ligada a las circunstancias particulares de la relación posesiva. La autora ilustra esta variabilidad con ejemplos como los anteriores, en cuanto que la probabilidad de perder la razón es mayor que la de que una ciudad deje de tener un templo.

Por último, en la categoría denominada *acciones y eventos* se encuentran los objetos con menor inherencia del *continuum*, caracterizados por su transitoriedad, así como por referir a las experiencias que el sujeto experimenta. En palabras de Aguilar Caro (2007, p. 91), los objetos poseídos de este grupo describen el ámbito experiencial del sujeto poseedor, porque nos dicen qué es y cómo es lo que se encuentra viviendo. La presencia de dos subgrupos responde a que la autora distingue entre acciones realizadas por los sujetos (14) y eventos que suceden a los sujetos (15).

- (14) Hoy tenemos fiesta en la escuela.
- (15) Mi equipo siempre tiene suerte en los sorteos.

En cuanto a la clasificación general, es necesario indicar que los criterios para distinguir los tres subtipos de posesiones en ocasiones pueden resultar

controvertidos. De hecho, Aguilar Caro (2007, pp. 70-71) admite que casos como el de (10) pueden considerarse atributos, aunque descarta esta opción aduciendo que el poseedor no los exige como indispensables. En otras palabras, constar de un templo no es condición necesaria para que un lugar sea considerado una ciudad, por tanto, *templo* no constituye un atributo. Encontramos una interpretación similar en las relaciones entre personas que se sitúan fuera de la red familiar. Así, también clasifica en el grupo de entes del entorno complementos como el de (16) al entender que se trata de relaciones que el poseedor establece libremente o en las que, al menos, cuenta con la posibilidad de decidir cuándo terminar. Esto no ocurre en las relaciones de parentesco, identificadas por la autora como atributos debido a su indisolubilidad, pues independientemente de que se mantenga el vínculo entre las personas implicadas, uno no puede dejar de ser hijo de su madre.

(16) Sara tiene dos amigas.

De lo anterior se desprende que separar los tipos de posesiones expresadas por *tener* en construcciones V + N atendiendo a criterios basados en rasgos posesivos como la inherencia o el control puede resultar complicado. Además, un análisis de estas características se alejaría del objetivo de mejora de la práctica que nos ocupa. No obstante, como anunciamos, el trabajo de Aguilar Caro presenta una aplicación de gran valor para el tratamiento didáctico de *tener* en cuanto que permite reducir considerablemente el extenso número de acepciones de la forma meta. En concreto, en su trabajo todos los significados atribuibles a *tener* se engloban en uno de los siguientes tres grandes sentidos: *poseer*, *caracterizar* y *experimentar*.

Respecto al segundo estudio sobre *tener* basado en el corpus mencionado, Alarcón (2026) analiza la frecuencia de uso y el grado de acierto del verbo meta por parte de aprendientes de español autoidentificados con una de las siguientes lenguas maternas: italiano, francés, alemán, polaco, ruso y chino. El trabajo se basa en el de Corpus Oral de Aprendientes de Español

(CORADE), un corpus de aprendientes de español de nivel avanzado que incluye una muestra de control de hablantes nativos de español.

De interés para el presente trabajo resulta la revalidación de la reducción semántica de Aguilar Caro (2007) efectuada durante la caracterización de *tener*. Para ello, en Alarcón (2026, pp. 94-95) se comprueba la viabilidad de asignar a todas las construcciones V + N con *tener* del CORADE uno de los tres hiperónimos aportados por Aguilar Caro —*poseer*, *caracterizar* y *experimentar*—. El resultado es favorable en todos los casos, excepto en aquellos predicados como el de (17) en los que la forma meta actualiza un sentido cercano al de la perífrasis modal *tener que* + infinitivo.

(17) Tengo dentista: no puedo ir a la reunión.

En definitiva, la proximidad semántica de las construcciones del tipo *tener dentista* con la perífrasis nos permite considerar esta interpretación como una categoría complementaria más vinculada a *tener que* + infinitivo que a *tener* y, por consiguiente, aprobar la reducción de significados de *tener* a los tres valores identificados por Aguilar Caro. Con todo, dada la alta polisemia de *tener*, no es posible descartar que futuros estudios empíricos arrojen otras categorías o presenten otros casos fronterizos similares a los de (17). Por otro lado, el hecho de que el análisis se efectúe en un corpus de hablantes no nativos contribuye a optimizar la aplicación didáctica del estudio en cuanto que los usos desviados de *tener* presentes en el corpus proporcionan evidencia fiable sobre posibles áreas de dificultad para los aprendientes, como se verá en 2.3.

En conclusión, el acceso contextualizado a datos que proporciona el análisis de corpus puede facilitar la necesaria tarea de reducir el gran número de acepciones de los verbos frecuentes. En el caso de corpus de aprendientes de español, la evaluación permite simultáneamente acotar el área de interés para el ámbito de la ASL mediante el análisis de aciertos y errores en el uso de los VF. Por consiguiente, a pesar de que “no está claro que algo que se produce

correctamente en un momento dado se haya adquirido” (Liceras, 1992, p. 14), el análisis de la producción no nativa facilita la detección de posibles áreas de dificultad para el aprendiz. A esta ventaja se suma el relativo fácil acceso a corpus de aprendientes. Respecto a esto último, nombramos algunas de las compilaciones de textos escritos y orales de hablantes no nativos recomendadas en la bibliografía. Entre los corpus escritos, destacan por su extensión y prestigio el Corpus Escrito del Español como L2 (CEDEL2), desarrollado por la Universidad Autónoma de Madrid y la Universidad de Madrid (Lozano, 2022) y el Corpus de aprendices del español (CAES), de la Universidad de Santiago de Compostela (Rojo y Palacios, 2016). En lo relativo a corpus orales, no nos consta la existencia de ningún trabajo equivalente al CAES o al CEDEL2. A ese respecto, el proyecto más destacable por el momento es el Spanish Learner Language Oral Corpora (SPLLOC), de la Universidad de Southampton (Mitchell, Domínguez, Arche, Myles & Marsden, 2008), junto con otros de menor tamaño como el Corpus Oral de Español como Lengua Extranjera (CORELE) de Campillos Llanos (2014) o el ya mencionado CORADE (Alarcón, 2026, pp. 114-139).

2.2. FASE 2: FORMULACIÓN DE UNA DEFINICIÓN SIMPLIFICADA Y OPERATIVA

Para manejar con coherencia y operatividad los sentidos obtenidos en el proceso de reducción semántica —*poseer*, *caracterizar* y *experimentar* en el caso de *tener*— resulta adecuado empaquetar la información en una definición integradora y cohesionada del VF. A ese respecto, es preciso no perder de vista dos ideas clave en el tratamiento de la polisemia que, a su vez, se desprenden de la fase de análisis anterior. La primera idea alude a dos tipos de significado de las palabras polisémicas: un significado idiosincrásico o básico y otro esquemático o inherente. El idiosincrásico no es compartido por todos los usos de la forma meta, sino solo por aquellos más cercanos al prototipo de Taylor (1996), mientras que el esquemático contiene la parte de significado presente en todos los usos de la palabra en cuestión. En línea con la distinción de Rufat (2019), el segundo tipo de significado entronca con

la noción de *punte semántico* empleada en el marco de la fraseología por Alonso Ramos (1997). La segunda idea clave consiste en el hecho de que el sentido esquemático no necesariamente coincide con la interpretación más ampliamente conocida. Por ejemplo, en *tener una pelea* el verbo no contiene la noción prototípica de *posesión* asociada a *tener*. Esto se traduce en que la definición simplificada que aspiramos a formular debe articularse en torno al sentido esquemático y, además, incluir al prototipo: en esa dirección, y no de lo idiosincrásico a lo esquemático. La aceptación de esta direccionalidad —de lo esquemático a lo prototípico— previene el intento de forzar dudosas interpretaciones semánticas como la que resultaría de afirmar que en *tener una pelea* se expresa una posesión —sentido idiosincrásico—, en lugar de una localización —sentido esquemático—.

A tenor de lo expuesto, podemos afirmar que la identificación del sentido esquemático de una palabra polisémica requiere de un elevado ejercicio de introspección, así como de la consulta de bibliografía especializada sobre el VF en cuestión. Por ejemplo, el estudio de *dar* exige comprender el funcionamiento de, entre otros, los verbos de transferencia, del mismo modo que para conocer el verbo *llevar* es preciso profundizar en la expresión del movimiento (Jiménez Calderón y Sanz, 2023), así como los análisis de *haber* o *quedar* entroncan con la existencia. Por su parte, en la investigación sobre verbos posesivos como *tener* resulta esencial indagar sobre el componente locativo.

A propósito de lo anterior, conviene subrayar que la idea de que tras la posesión subyace la locación está plenamente aceptada en la bibliografía (Lyons, 1968; Givón, 1984; Stassen, 2009, entre otros). En esta línea, Cifuentes Honrubia (2010, p. 123) destaca los trabajos de Lyons (1980, pp. 651-657) en los que se evidencia que en muchas lenguas las construcciones existenciales y posesivas derivan, tanto sincrónica como diacrónicamente, de construcciones locativas. Tal como subraya Cifuentes, la naturaleza de esta relación se encuentra en la mayor accesibilidad y relevancia de lo locativo para la cognición.

Las expresiones espaciales son más básicas, gramatical y semánticamente, que los diversos tipos de expresiones no espaciales, sirviendo de plantilla estructural para otras expresiones, ya que la organización espacial tiene una importancia máxima en el conocimiento humano (2010, pp. 123-125).

Así, del análisis de Moreno Cabrera (2003, p. 67) se desprende que las construcciones V + N en las que participa *tener* se clasifican como localizaciones de continente, en oposición a las de contenido. En las localizaciones de contenido (18) se sitúa a la entidad localizada (*Pedro*) en un lugar (*casa*), mientras que en las de continente (19) se parte del lugar (*la casa*) para ubicar a la entidad localizada (*patio*).

(18) Pedro está en casa.

(19) La casa tiene patio.

El planteamiento locativo coincide con la definición infraespecificada que propone De Miguel (2008, p. 575) en el marco de la Teoría del Lexicón Generativo, según la cual el verbo *tener* equivale a “ser el lugar donde está algo [un árbol en un jardín o una duda en una persona]”. Por consiguiente, el esquema *A tiene B* según este enfoque se define como sigue: *A es el lugar donde está B*.

Sin embargo, aludir al valor estrictamente locativo como el sentido esquemático de *tener* en construcciones V + N resulta insuficiente para explicar diferencias con los otros dos grandes verbos que expresan ubicación: *estar* y *haber*. Dicho de otro modo, manejar la definición de De Miguel (2008) no aclara la diferencia de significado entre (20), (21) y (22), por ejemplo.

(20) La buhardilla tiene cuadros.

(21) Los cuadros están en la buhardilla.

(22) Hay cuadros en la buhardilla.

Los tres enunciados informan sobre el lugar en el que se encuentran los cuadros. No obstante, la selección de *tener* en (20) permite saber que, además,

los cuadros forman parte de la decoración de la estancia, mientras que en (21) y (22) la ambigüedad sobre si los cuadros se encuentran en una caja o en la pared solo se resuelve mediante el contexto. Para incorporar la información específica que distingue a *tener* de los valores locativo y existencial de los verbos *estar* y *haber*, respectivamente, sin forzar la interpretación del término *posesión*, es necesario acudir a la habilidad del punto de referencia en los términos de Langacker (2000). En concreto, el autor define esta habilidad como la capacidad del poseedor para servir de punto de referencia para el establecimiento de contacto mental con otra —el poseído—. Por ejemplo, a propósito de las relaciones parte/todo como la de 18, Langacker (2000, p. 178) observa que una parte solo puede concebirse en relación con el todo, que funciona como un punto de referencia natural para su configuración. Así, *patio*, presenta en su base la concepción del todo del que se sirve para perfilar una entidad cuya naturaleza y estatus de subparte depende, a su vez, de su función dentro de la configuración general —*la casa*—. Para ilustrar la incapacidad de reconocer las partes autónoma e individualmente, el autor alude a contextos muy específicos como los únicos en los que sí resultaría posible concebir la parte sin el todo, como podría ser la capacidad de una fisioterapeuta para visualizar el tríceps sin necesidad de insertarlo en el cuerpo de un individuo concreto o general.

A tenor de lo expuesto, una definición simplificada y operativa de *tener* cuando actualiza el esquema *A tiene B* ha de integrar en su sentido esquemático tanto el componente locativo como la habilidad del punto de referencia. He aquí la propuesta de Alarcón (2026, p. 280).

A es (i) la ubicación donde está B, y (ii) el punto de referencia a partir del cual se conceptualiza B.

Ahora bien, en aras de un mejor manejo de la definición y de su posible trasposición didáctica proponemos una versión adaptada en la que la alusión a la habilidad del punto de referencia se sustituye por los tres sentidos en los que A sirve para la conceptualización a B, es decir, B puede interpretarse como una posesión de A, una característica de A o algo que A experimenta:

A *tiene* B: A es el lugar donde está B, siendo B algo que pertenece a A, lo caracteriza o es experimentado por él.

De este modo, el análisis teórico de los VF que no aparecen en *Redes* iniciado con la reducción del número de acepciones de la primera fase de la instrucción, se completa en esta segunda fase mediante dos acciones: la identificación de los sentidos esquemático y prototípico, por un lado, y, por otro, la formulación de una definición del verbo que integre lo anterior con los hiperónimos resultantes de la Fase 1.

2.3. FASE 3: IDENTIFICACIÓN DE LAS ÁREAS DE INTERÉS PARA EL AULA DE E/L2

Adquiridos los conocimientos teóricos anteriores, en la tercera y última fase de formación del profesorado se analizan los usos de los VF seleccionados susceptibles de suponer alguna dificultad para el aprendiz. A ese respecto, como avanzamos en 2.2, trabajar con corpus de aprendientes durante las fases anteriores facilita la identificación de posibles focos de errores. En concreto, el análisis de Alarcón (2026) descrito en el apartado anterior arroja tres datos de interés. En primer lugar, destacamos que las ocurrencias en las que el verbo actualiza un significado equivalente a la noción más idiosincrásica de posesión de Taylor (1996) son las que contienen el menor número de errores de la muestra de hablantes no nativos en su conjunto. Por consiguiente, sería posible inferir que el dominio productivo del valor prototípico de *tener* resulta más accesible para los participantes del CORADE —hablantes de italiano, francés, alemán, polaco, ruso o chino como L1—. En concreto, del total de usos del verbo con el sentido de *poseer*, el 14 % son inaceptables, frente al 26 % y al 30 % de usos erróneos en las ocurrencias en las que el verbo *tener* actualiza los valores de *caracterizar* y *experimentar*, respectivamente.

En segundo lugar, en el análisis de errores se han registrado desviaciones recurrentes en aquellas construcciones en las que *tener* activa el sentido de *caracterizar*. Se trata del uso inadecuado de *haber* (23) en sustitución de *hacer*. Como comentamos al principio, se trata de una desviación habitual en

sinohablantes con una competencia en español inferior al Nivel Umbral o B1. Lo relevante de este asunto estriba en que, a diferencia de otros errores transitorios propios de interlenguas en estadios iniciales que se anulan a medida que el aprendiente avanza en su dominio de la lengua, la confusión entre *haber* y *tener* en hablantes de chino muestra cierta resistencia a desaparecer (Alarcón, 2026, p. 261). En otras palabras, el hecho de que en la producción oral de aprendientes de nivel B2 de origen chino persistan errores como el de (22) podría indicar que se trata de un error fosilizable.

(23) CHIT1 Las clases en línea por internet hay muchas ventajas.

Por último, en el empleo de *tener* con el sentido de *experimentar* también se ha detectado otro posible candidato a la fosilización relativo a la sustitución de la forma meta por *hacer* (8), así como otros errores interlingüísticos del tipo *tener un buen tiempo* (24) motivados posiblemente por la interferencia del inglés como lengua adicional (*to have a good time*).

(8) ITEX1 El viaje es la única cosa que te permite *hacer experiencias [tener].

(24) PON9. Es importante tener la sensación de que has *tenido un tiempo muy bueno [pasado].

En lo relativo a los errores, entra dentro de lo esperable que el profesorado en formación albergue cierta intuición sobre las áreas de dificultad de *tener* o de otros VF debido a su propia experiencia como aprendiz de segundas lenguas, o al contacto con hablantes no nativos de español. Asimismo, resulta inviable tratar de anticipar todos los posibles errores recurrentes en el uso de los VF teniendo en cuenta la diversidad de lenguas adicionales del alumnado. Por ello, el objetivo de esta última fase estriba en dotar a los futuros docentes del conocimiento teórico y aplicado necesario para que, en última instancia, contribuyan a desarrollar la capacidad de los aprendices para observar, reflexionar, argumentar y, en definitiva, comprender el sistema lingüístico (Rufat, 2023, p. 154). En particular, asumimos que un conocimiento profundo

sobre la motivación de la polisemia inherente a los VF por parte del profesorado favorece una mejor instrucción en el aula de E/L2. En otras palabras, la práctica docente sensible a la transmisión de conocimiento explícito sobre el funcionamiento de los VF es crucial para la aplicación de dos de los principios que han de regir la enseñanza de VF, según Jiménez Calderón y Rufat (2019, p. 131): combinar los aprendizajes incidental e intencional, por un lado, y, por otro, atender a los factores que condicionan la retención del léxico.

3. CONCLUSIÓN

El presente trabajo proporciona orientaciones para el análisis de los VF desde una perspectiva basada en la formación del profesorado. Con motivo de la escasa o nula presencia de los VF en los programas de másteres y cursos de formación docente, en este trabajo hemos tratado de dirigir la atención hacia la necesidad de que el profesorado de E/L2 adquiriera unos conocimientos sólidos —aunque no necesariamente exhaustivos— sobre la naturaleza de estos VF. Trabajos como los de Rufat (2023) o Jiménez Calderón y Sanz (2023) sobre *dar* y *llevar*, respectivamente, constituyen un modelo imprescindible para el análisis de VF que aparecen en *Redes*. Por nuestra parte, en las páginas anteriores hemos descrito un marco de acción general para el estudio de aquellos VF que no figuran en el diccionario combinatorio. A pesar de las notables diferencias que separan a estos últimos, consideramos que el procedimiento de tres fases descrito —reducción semántica, definición simplificada y reconocimiento de las áreas de interés para el alumnado—, es lo suficiente abierto como para que acoger a verbos tan dispares como el existencial *quedar* o el posesivo *tener*. Asimismo, la concepción del mecanismo de la polisemia como eje vertebrador de las tres fases otorga consistencia al vínculo existente entre los miembros de esta suerte de categoría con la que el docente ha de familiarizarse: la categoría formada por los VF que no aparecen en *Redes*. No obstante, la evaluación de esta propuesta metodológica deberá validarse empíricamente en futuras investigaciones.

REFERENCIAS BIBLIOGRÁFICAS

- Aguilar Caro, N. M. (2007). *El verbo tener y las relaciones de posesión*. [Tesis de maestría. Universidad Nacional Autónoma de México]. Repositorio Universitario ENALLT: <https://ru.enallt.unam.mx/jspui/handle/ENALLT.UNAM/449>
- Alarcón, N. (2026). *El verbo tener en el Corpus Oral de Aprendientes de Español (CORADE)*. [Tesis doctoral. Universidad de Castilla – La Mancha]. Repositorio Institucional. <https://ruidera.uclm.es/items/f7405f11-758e-4711-ba58-4da4edf7f392>
- Alonso Ramos, M. (1997). Coocurrencia léxica y descripción lexicográfica del verbo DAR: hacia un tratamiento de los verbos soportes. *Zeitschrift für Romanische Philologie*, 113(3), pp. 380-417.
- Alonso Ramos, M. (2004). *Las construcciones con verbo de apoyo*. Visor.
- Alonso Ramos, M. (2010). No importa si la llamas o no colocación, descríbela. En C. Mellado (ed.), *La fraseografía del S. XXI: Nuevas propuestas para el español y el alemán* (pp. 55-80). Frankfurt & Timme. [http://www.dicesp.com/app/webroot/files/file/Alonso%202010\(1\).pdf](http://www.dicesp.com/app/webroot/files/file/Alonso%202010(1).pdf)
- Altenberg, B. & Silviane, G. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22(2), 173-195. <https://doi.org/10.1093/applin/22.2.173>
- Barrios, M. A. (2015). *Las colocaciones del español*. Arco/Libros.
- Boers, F. & Lindstromberg, S. (2008). Opening chapter: How cognitive linguistics can foster effective vocabulary teaching. En F. Boers & S. Lindstromberg (eds.), *Cognitive linguistic approaches to teaching vocabulary and phraseology* (pp. 1-61). De Gruyter Mouton. <https://doi.org/10.1017/S0272263109990179>
- Bosque, I. (2001). Sobre el concepto de colocación y sus límites. *Lingüística española actual*, XXIII(1), 9-40.
- Bosque, I. (dir.) (2004). *Redes. Diccionario combinatorio del español contemporáneo*. SM.
- Campillos Llanos, L. (2014). A Spanish Oral Learner Corpus for Computer-Aided Error Analysis. *Corpora*, 9(2), 207-238. <https://doi.org/10.3366/cor.2014.0058>
- Cifuentes Honrubia, J. L. (2010). *Clases semánticas y construcciones sintácticas: alternancias locales en español*. Axac.
- Corpas Pastor, G. (1996). *Manual de fraseología española*. Gredos.
- Davies, M. & Davies, K. H. (2018). *A Frequency Dictionary of Spanish. Core Vocabulary for Learners*. Routledge. <https://doi.org/10.4324/9781315542638>
- De Miguel, E. (2008). Construcciones con verbo de apoyo en español. De cómo entran los nombres en la órbita de los verbos. En I. Olza Moreno, M. Casado Velarde & R. González Ruiz (eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística* (pp. 567-578). Servicio de Publicaciones de la Universidad de Granada. <https://core.ac.uk/download/pdf/83572046.pdf>
- Givón, T. (1984) *Syntax: A Functional-Typological Introduction*. John Benjamins. <https://doi.org/10.1075/z.17>

- Granger, S. (2002). A Bird's-eye view of learner corpus research. En S. Granger, J. Hung & S. Petch-Tyson (eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 3–33). John Benjamins. <https://doi.org/10.1075/llt.6.04gra>
- Guo, X. (2006). *Verbs in the Written English of Chinese Learners. A Corpus-based Comparison between Non-native Speakers and Native Speakers*. [Tesis doctoral. Universidad de Birmingham]. Repositorio UBIRA Etheses: <https://etheses.bham.ac.uk/id/eprint/871/>
- Hijazo-Gascón, A. (2021). *Moving Across Languages: Motion Events in Spanish as a Second Language*. De Gruyter Mouton. <https://doi.org/10.1515/9783110721072>
- Ibarretxe-Antuñano, I. (2006). Cross-linguistic polysemy in tactile verbs. En J. Luchjenbroers (ed.), *Human cognitive processing* (pp. 235–253). John Benjamins Publishing Company. <https://doi.org/10.1075/hcp.15.16iba>
- Jiménez Calderón, F. & Rufat, A. (2019). The teaching of frequent verbs through lexical approaches. *Verba Hispanica*, 27(1), 131-151. <https://doi.org/10.4312/vh.27.1.131-151>
- Jiménez Calderón, F. & Sanz, M. (2023). El trabajo con ditransitivas en ELE: un estudio del verbo *llevar*. En F. Jiménez Calderón (ed.), *Los verbos ditransitivos y su enseñanza: transferencia, locación y movimiento* (pp. 157-198). Peter Lang. <https://doi.org/10.3726/b21275>
- Koike, K. (2001). *Colocaciones léxicas en el español actual: estudio formal y léxico semántico*. Universidad de Alcalá.
- Langacker, R. W. (2000). *Grammar and Conceptualization*. Mouton de Gruyter. <https://doi.org/10.1515/9783110800524>
- Laufer, B. (1997). What's in a word that makes it hard or easy? Intralexical Factors Affecting the Difficulty of Vocabulary Acquisition. En N. Schmitt & M. McCarthy (eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 1140–1155). Cambridge University Press.
- Li, Y. (2022). Corpus-Based Error Analysis of Chinese Learners' Use of High-Frequency Verb Take. *English Language Teaching*, 15(2), 21-30. <https://doi.org/10.5539/elt.v15n2p21>
- Liceras, J. (1992). *La adquisición de lenguas extranjeras: hacia un modelo de análisis de la interlengua*. Visor.
- Littlemore, J. (2023). *Applying cognitive linguistics to second language learning and teaching* (2.^a ed.). Cham: Palgrave. Macmillan. <https://doi.org/10.1007/978-3-031-22107-9>
- López Pérez, S. & Benali Taouis, H. (2019). Analysis of noun (direct object) collocations with the high-frequency verb DO by Spanish students in an online learner corpus. *Complutense Journal of English Studies*, 27, 99-120. <https://doi.org/10.5209/cjes.63337>
- Lozano, C. (2022). CEDEL2: Design, compilation and web interface of an online corpus for L2 Spanish acquisition research. *Second Language Research*, 38(4), 965-983. <https://doi.org/10.1177/02676583211050522>

- Lyons, J. (1968). *Introduction on theoretical linguistics*. Cambridge University Press. Versión española: *Introducción en la lingüística teórica*. (1977). Semantics. Cambridge University Press. Versión española: *Semántica*. Teide, 1980.
- Lyons, J. (1980). *Semántica*. Teide.
- Masid Blanco, O. (2022). La enseñanza de la metáfora lingüística en ELE. En M. L. Regueiro Rodríguez (coord.), *Del léxico y la semántica a la pragmática en ELE* (pp. 175-198).
- Mitchell, R., Domínguez, L., Arche, M., Myles, F. & Marsden, E. (2008). *SPLLOC: A New Corpus for Spanish Second Language Acquisition*. EUROSLA Yearbook. <https://doi.org/10.1075/eurosla.8.15smit>
- Molina, V. (2013). *Enseñanza de las colocaciones con verbo soporte: dar, tener y hacer. Estudio contrastivo español-italiano*. Memoria de Máster. Universidad Complutense de Madrid.
- Moreno Caberra, J. C. (2003). *Semántica y gramática. Sucesos, papeles semánticos y relaciones sintácticas*. Antonio Machado
- Pérez Serrano, M. (2017). *La enseñanza-aprendizaje de vocabulario en ELE desde los enfoques léxicos*. Arco/Libros.
- Pinker, S. (2007): *El mundo de las palabras. Una introducción a la naturaleza humana*. Paidós.
- Qu, C. (2023). The correlated influence of semantic types, L1 translation equivalents and language proficiency on EFL learners' production of polysemous words. *Frontiers in Psychology*, 14, 1-10. <https://doi.org/10.3389/fpsyg.2023.1162008>
- Rojó, G. & Palacios, I. (2016). Learner Spanish on computer: The CAES 'Corpus de Apprentices de Español' Project. En M. Alonso Ramos (ed.), *Spanish learner corpus research: Current trends and future perspectives* (pp. 55-87). John Benjamins. <https://doi.org/10.1075/scl.78.03roj>
- Rufat, A. (2019). *La investigación de corpus de aprendientes basada en el análisis contrastivo de la interlengua. El caso de DAR*. Universidad de Jaén. UJA Editorial.
- Rufat, A. (2023). Los verbos de transferencia desde la gramática cognitiva y las restricciones basadas en la concordancia de los rasgos léxicos. En F. Jiménez Calderón (ed.), *Los verbos ditransitivos y su enseñanza: transferencia, locación y movimiento* (pp. 113-156). Peter Lang.
- Rufat, A. & Jiménez Calderón, F. (2019). Vocabulario. En J. Muñoz-Basols, E. Gironzetti & M. Lacorte (eds.), *The Routledge Handbook of Spanish Language Teaching: metodologías, contextos y recursos para la enseñanza del español L2* (pp. 229-242). Routledge. <https://doi.org/10.4324/9781315646169-17>
- Sawaguchi, R. & Mizumoto, A. (2022). Exploring the use of make + noun collocations by Japanese EFL learners through a bilingual essay corpus. *Corpora*, 17(SI), 61-77. <https://doi.org/10.3366/cor.2022.0247>
- Stassen, L. (2009). *Predicative possession*. Oxford University Press. <https://doi.org/10.1093/oso/9780199211654.001.0001>

- Taylor, J. R. (1996). *Possessives in English: An Exploration in Cognitive Grammar*. Oxford University Press. <https://doi.org/10.1093/oso/9780198235866.001.0001>
- Tsutahara, R. (2020). Un estudio comparativo de ir y go como verbos semicopulativos. *Hispanica*, 64, 51–77. <https://doi.org/10.4994/hispanica.2020.51>
- Vicente Balenciaga, S. & Goethals, P. (2026). La polisemia léxica como oportunidad pedagógica: un enfoque cognitivo del currículum de vocabulario en ELE. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 20(40), 169–196. <https://doi.org/10.26378/rnlael2040664>