

*Lemateca del DRAE: la vida de las palabras
en la lexicografía académica**
DRAE's Lemateca: the life of words in academic lexicography

M.^a ÁNGELES BLANCO IZQUIERDO

*Centro de Estudios de la RAE
bangeles@rae.es*

CRISTINA BUENAFUENTES

*Universitat Autònoma de Barcelona
Cristina.Buenafuentes@uab.cat*

GLORIA CLAVERÍA

*Universitat Autònoma de Barcelona
Gloria.Claveria@uab.cat*

ENRIQUE JIMÉNEZ RÍOS

*Universidad de Salamanca
enrique@usal.es*

NATALIA TERRÓN

*Universitat Autònoma de Barcelona
Natalia.Terron@uab.cat*

JOAN TORRUELLA

*ICREA-Universitat Autònoma de Barcelona
Joan.Turruella@uab.cat*

Resumen: *Lemateca* es una herramienta informática realizada en el marco del proyecto de investigación “Historia interna del *Diccionario de la lengua castellana* de la Real Academia Española en el siglo XIX”. A la vez que recoge y sistematiza la trayectoria de los lemas y los cambios, tanto formales como de lematización, que estos experimentan en las distintas ediciones de los diccionarios académicos, se constituye en un instrumento de gestión que permite explotar todos estos datos adecuadamente y desde distintos puntos de vista. Se intenta, así, tender un puente entre la investigación léxica y lexicográfica y las humanidades digitales.

Palabras clave: historia de la lexicografía; diccionario; Real Academia Española; humanidades digitales.

Abstract: *Lemateca* is a computerized tool that has been developed within the framework of the research project “Historia interna del *Diccionario de la lengua castellana* de la Real Academia Española en el siglo XIX”. This tool allows the users to track and systematize the course of the lemmas and their amendments, both formal and regarding lemmatisations, that they undergo in the different editions of the academic dictionaries. In addition, it serves as a management instrument that enables to exploit this data appropriately from different points of view. Thus, it is aimed to build a bridge between the research on lexicology and lexicography and the digital humanities.

Key words: history of lexicography; dictionary; Royal Spanish Academy; digital humanities.

Fecha de presentación: 12/04/2019 *Fecha de aceptación:* 02/11/2019

* Esta investigación ha sido posible gracias a la ayuda otorgada por el Ministerio de Economía y Competitividad (referencia PGC2018-094768-B-I00) y al apoyo de la Generalitat de Catalunya (SGR2017-1251).

1. LA HISTORIA INTERNA DEL *DICCIONARIO DE LA LENGUA CASTELLANA* DE LA REAL ACADEMIA ESPAÑOLA EN EL SIGLO XIX

El proyecto de investigación «Historia interna del *Diccionario de la lengua castellana* de la Real Academia Española en el siglo XIX» tiene como objetivo fundamental el establecimiento de la enmienda llevada a cabo en cada una de las ediciones del diccionario publicadas en el siglo XIX. En una primera fase, y tomando como punto de partida la cuarta edición del diccionario, se han sometido a análisis las seis ediciones del diccionario que se publicaron entre 1817 y 1852 (*DRAE*-1817, *DRAE*-1822, *DRAE*-1832, *DRAE*-1837, *DRAE*-1843 y *DRAE*-1852) por medio de la identificación de los cambios introducidos en cada una de ellas. En una segunda fase, se pretende ampliar la investigación a las tres últimas ediciones del siglo (*DRAE*-1869, *DRAE*-1884 y *DRAE*-1899), que implican una ampliación de la nomenclatura mayor que en las ediciones precedentes y también involucran notables innovaciones en la aplicación de la técnica lexicográfica¹.

El objetivo último de esta investigación consiste en averiguar la significación lexicográfica y lexicológica de estas transformaciones, y reconstruir de ese modo la historia interna del diccionario académico. Como ya se ha expuesto en otro lugar (Blanco, Clavería y Jiménez Ríos, 2018), el proyecto se desarrolla con un fuerte apoyo en las humanidades digitales, las cuales se encuentran en su misma concepción.

Uno de los resultados del proyecto es el repositorio de formas de los lemas que contienen los diccionarios estudiados y que estamos construyendo actualmente. Se trata de un instrumento que hemos llamado *Lemateca* y que, en cierta forma y tal como hemos recogido en el título de nuestra contribución, quiere reflejar la vida de las palabras en la lexicografía académica. El objetivo de este trabajo es la presentación de esta herramienta, acompañada de una reflexión sobre lo que su existencia representa en la investigación de la lexicografía académica.

2. *LEMATECA* DEL *DRAE*

Se sostiene en la historia de la ciencia que existen dos tipos de revoluciones científicas, unas resultan de nuevas herramientas y otras de nuevos conceptos (*cf.* Dyson, 1997: 49, *cf.* Rojo, 2009). No es necesario insistir en que, entre las nuevas herramientas, el ordenador y las nuevas posibilidades que brinda a la ciencia constituyen *per se* una revolución científica, y no es difícil adivinar que esta revolución instrumental puede desembocar en una revolución conceptual.

Como se ha señalado, la fuerte imbricación que existe en nuestro trabajo entre la investigación propiamente dicha y las humanidades digitales explica que nuestro punto de partida se encuentre en la explotación de instrumentos de este tipo (por ejemplo, el *NTLLE*) y que los resultados obtenidos se constituyan en aplicaciones informáticas que, a su vez, pueden ser la base de una nueva visión e interpretación de los datos ob-

¹ Adoptamos para la referencia de las distintas ediciones del diccionario la sigla *DRAE* y el año de edición, excepto en la última, para la que se usa *DLE*-2014. Remitimos para la citación de las primeras al *NTLLE*.

tenidos, con lo que también se puede, en palabras de Dyson (1997: 50), «explain old things in new ways».

Lemateca surge de los datos atesorados en nuestro proyecto de investigación y permite una observación dinámica de los avatares que experimentan las palabras en las distintas ediciones del *DRAE*². La herramienta intenta superar los problemas derivados de la lexicografía tradicional y artesanal de la que surgieron los diccionarios estudiados y que, como es natural, generó abundante variación y heterogeneidad en la evolución de la obra académica. De este modo, *Lemateca*, siendo una herramienta informática, permite establecer y recuperar la vida de las palabras en la lexicografía académica, es decir, su vida lexicográfica.

2.1. *La concepción*

El empleo de herramientas informáticas en los estudios filológicos tiene la finalidad de facilitar, desde una base científica, la comprensión de los principios y de las causas que rigen los elementos lingüísticos objeto de estudio.

La parte instrumental de *Lemateca* consiste en una base de datos y su gestor informático; ambos han sido creados *ad hoc* para poder tratar las ocurrencias del corpus —en este caso los lemas de las distintas ediciones del *DRAE*—, es decir, las palabras y sus variantes formales que en algún momento han figurado en la macroestructura de una, varias o todas las ediciones del diccionario académico. Esta herramienta debe facilitar el manejo e interpretación de todo el inmenso número de documentaciones de los lemas en las diferentes ediciones del diccionario. De esta manera, gracias a las nuevas posibilidades que ofrece la herramienta (parte instrumental), se dispondrá de una visión de los datos para describir y analizar fenómenos lingüísticos (parte conceptual) del todo novedosa y potente, y que permitirá unos análisis imposibles de llevar a cabo sin la mencionada herramienta. No olvidemos que, en ciencia, si se quieren obtener resultados nuevos, es preciso cambiar o la muestra que se estudia o el método de análisis que se utiliza. Por ello, la nueva herramienta hace posible un nuevo método de análisis que no solamente da acceso al total de ocurrencias del universo que se quiere investigar y a la facilidad de ordenarlas y combinarlas de distintas maneras, sino que hace posible un aspecto imprescindible en la investigación científica, que es la cuantificación de los elementos y de los procesos junto a su análisis cualitativo.

La investigación empírica en lingüística puede ser experimental u observacional. En el caso de *Lemateca*, al tratarse de material de carácter histórico, el método utilizado solamente puede ser observacional, pues se manejan únicamente los datos que ya existen y tal como existen, sin que se pueda intervenir en la modificación de las condiciones en las que se producen. El uso del método observacional, sin embargo, no implica pasividad en el investigador, sino que, para desarrollar una teoría o comprobar una hipótesis de trabajo, este puede y debe operar con *variables dependientes e independientes* referidas a las condiciones en que se producen los datos analizados, lo que hace que la investigación se sustente en una base científica.

² Por su concepción, objetivos y contenidos, se trata de un proyecto distinto al *Mapa de diccionarios* del Instituto de Investigación Rafael Lapesa de la Real Academia Española.

2.2. Principios instrumentales de Lemateca

Como ya se ha mencionado, *Lemateca* es una herramienta que permite el acercamiento a los datos de un modo que hasta ahora era imposible, no solo por el volumen de palabras que componen el universo estudiado, la *población* en términos estadísticos, sino, sobre todo, por la posibilidad de clasificarlas, ordenarlas y contabilizarlas. En este caso, el punto de partida es el corpus de alrededor de 64 000 formas que componen los diccionarios académicos del siglo XIX.

La estructura de la base de datos que se ha creado para facilitar la observación de los testimonios se fundamenta en la selección de aquellas *variables* que se han considerado oportunas para llevar a cabo investigaciones referidas a la trayectoria de los lemas que se recogen en las diversas ediciones del *DRAE*. Recordemos que «para poder trabajar con corpus siguiendo un método científico y poder aplicar las técnicas estadísticas a su análisis es necesario disponer de *variables* que hagan posible la experimentación, de manera que, cuando cambien alguna o algunas de las *variables*, se pueda observar cómo reaccionan otra u otras *variables*» (Torruella, 2017: 70). Así, por ejemplo, si se quiere observar la sustitución de la grafía *x* por la grafía *j* (variable dependiente) en las distintas ediciones del *DRAE* (como en *coxear* > *cojear*, *vid.* 2.3.1.), se deberá comprobar cómo evoluciona este aspecto cuando cambiamos de edición analizada (variable independiente).

El programa gestor de la base de datos permite localizar las voces que configuran el lemario del diccionario académico y las diferentes formas que han adoptado a lo largo de su trayectoria en el mismo. Cabe recordar que en las obras lexicográficas analizadas es posible hallar encabezando un artículo lemas simples (*canciller*, *cancilleresco*, *ca*), lemas múltiples (*viajecico*, *llo*, *to*) y lemas complejos (*sacro lumbar*), de ahí que un lema (*viajecico*, *llo*, *to*) pueda estar compuesto por varias formas (*viajecico*, *viajecillo*, *viajecito*). Asimismo, el programa gestor está pensado para dar cuenta de otras vicisitudes de las palabras analizadas, como los cambios formales que puedan experimentar (*relox* > *reloj*, *tenaza* > *tenazas*), unas veces motivados por condiciones lingüísticas y otras por decisiones de carácter lexicográfico. Finalmente, el gestor también hace posible vincular lemas que de algún modo están relacionados; de esta manera cuando se observa la trayectoria de una palabra se facilita el salto hipertextual a la historia lexicográfica de los lemas con los que establece algún tipo de ligazón, como pueden ser formas antiguas desaparecidas (*argullia* respecto a *orgullo*) o las distintas variantes formales que acoge el diccionario en una misma edición (*sustraer* y *substraer*).

Las consultas a *Lemateca* pueden ser de dos tipos: las que requieren una lista de palabras que cumplen unas determinadas condiciones en una edición concreta (por ejemplo, todas las palabras que experimentan la sustitución gráfica de *x* por *j* en el *DRAE*-1817) y las que proporcionan la trayectoria cronológica y variacionista de un vocablo a través de las distintas ediciones del diccionario (por ejemplo, *relox* sustituido por *reloj*). Entre los resultados de los dos tipos de consultas se pueden realizar saltos hipertextuales que relacionan los datos y, de este modo, es posible obtener el conocimiento de sus principios y de sus causas. Por ejemplo, la palabra *reloj* experimenta un cambio gráfico de *x* (*relox*) a *j* (*reloj*) que se produce en la edición de 1832. A partir de este

dato, es posible realizar un salto a la lista de palabras que muestran este mismo cambio, ya sea en la misma edición o en otras ediciones. Con ello, es posible indagar en las directrices que siguió la sustitución de *x* por *j* en el diccionario académico.

Con el fin de vincular las diversas formas que una misma palabra puede adoptar en el leuario de las diferentes ediciones del diccionario (por ejemplo, *bardaxe*, *bardage* y *bardaje*), ha sido necesario crear un campo invariante que permite relacionar todas las variantes de la palabra. De esta manera, y siguiendo con el ejemplo anterior, cuando el usuario busca por cualesquiera de las tres formas obtendrá, a través de la forma invariante, en primer lugar, la vigencia cronológica de la suma de las tres formas (1726, 1780-1817, 1899-2014)³ y, en segundo lugar, la vigencia cronológica de cada una de las formas (*bardaxe* [1726, 1803], *bardage* [1780-1791] y *bardaje* [1817, 1899-2014]).

BARDAJE

vigencia cronológica: 1726, 1780-1817, 1899-2014

formas:

<i>bardaxe</i>	1726, 1803
<i>bardage</i>	1780-1791
<i>bardaje</i>	1817, 1899-2014

Se ha acordado que la forma invariante que aglutina todas las formas variantes sea la que consta en la última edición del diccionario académico, aunque, para los casos de supresiones de lemas (*vid.* 2.3.4.), la invariante es la forma que tiene el lema en la última edición en la que se registra la voz.

Las informaciones que se recogen en la base de datos se pueden clasificar en tres tipos:

- las que hacen referencia a la trayectoria de las palabras en las distintas ediciones del diccionario académico;
- las que hacen referencia a las transformaciones que las palabras experimentan en su trayectoria en el diccionario;
- las que hacen referencia a los vínculos de una palabra con otras.

Las informaciones del primer grupo permiten recuperar la trayectoria de una sola palabra en las distintas ediciones del diccionario o un grupo de voces con unas características comunes, por ejemplo, las adiciones o supresiones de una determinada edición. Las informaciones del segundo grupo aportan información sobre los cambios que registra el lema (por ejemplo, *cuarteto* > *cuarteto*; *tranzon de tierra* > *tranzon*, *matafuegos* > *matafuego*). Las informaciones del último grupo permiten establecer relaciones entre lemas diferentes que sostienen algún tipo de relación (*oscuro/obscuro*).

2.3. La vida de las palabras a través de Lemateca: algunas biografías ilustrativas

La comparación entre los leuarios de las distintas ediciones de los diccionarios académicos permite trazar la trayectoria de cada palabra en la lexicografía académica, es decir, su biografía: su nacimiento o aparición en el diccionario, su vigencia y, en ocasiones, su desaparición en alguna de las ediciones estudiadas. Posibilita también

³ En la vigencia cronológica se señalan los intervalos con guion y la discontinuidad entre ediciones, con una coma.

establecer los avatares de su recorrido en las distintas ediciones con los cambios que afectan a su forma y a su tratamiento lexicográfico, de lo que se pueden extraer conocimientos tanto de la historia de las palabras como de la evolución de los diccionarios.

Exponemos a continuación algunos ejemplos ilustrativos como muestra de las posibilidades de la aplicación.

2.3.1. La ortografía como motor de cambios en la lematización

La importancia de los principios ortográficos en el diccionario puede observarse en el discurso sobre la ortografía de los preliminares del *Diccionario de autoridades*. Desde un punto de vista teórico y general, cabe citar el trabajo de N. Catach (1989: 501-508) sobre «L'orthographe dans le dictionnaire monolingüe», en el que se describen los diccionarios como «les principaux porteurs de norme». La historia de la relación entre la ortografía y el diccionario académicos está aún en buena parte por desentrañar: en ocasiones el diccionario aplica los cambios introducidos en la ortografía, mientras que en otras ocasiones es el propio diccionario el que introduce innovaciones en materia ortográfica (Alcoba, 2007, 2012; Quilis Merín, 2009). En cualquier caso, uno de los cambios fundamentales del registro de los lemas en el diccionario se encuentra en su forma gráfica: los cambios en el *DRAE-1817*, a la zaga de la octava edición de la ortografía (*ORAE*, 1815), son los mejores ejemplos (Blanco, 2018, Terrón, 2018); a ellos hay que añadir los cambios aislados registrados en un buen número de entradas en distintas ediciones.

Un ejemplo de cambio gráfico que afecta a todas las voces de una edición es la sustitución de *x* por *j* en la edición de 1817, según prescribe la ortografía de 1815 (*ORAE*, 1815: 55). Este es el caso de la voz *cojear*, registrada con *x* (*coxear*) desde el *Diccionario de autoridades* hasta la cuarta edición del diccionario (*DRAE-1803*). En la edición siguiente (*DRAE-1817*), la introducción del cambio sustituye la grafía *x* por *j* (*cojear*) y se mantiene así hasta la actualidad (*DLE-2014*). En *cojear*, como en otras muchas voces (*bajo*, *quijote*, *jilguero*, etc.), el cambio gráfico es definitivo en la forma normativa de la palabra y se mantiene hasta nuestros días sin otra mudanza.

COJEAR

vigencia cronológica: 1729-2014

formas:

<i>coxear</i>	1729-1803
<i>cojear</i>	1817-2014

La incorporación de las reformas ortográficas en el diccionario incide en la trayectoria vital de las palabras. Como muestran los ejemplos anteriores, y otros muchos, una regla propuesta en una determinada ortografía afecta a la lematización de todas las voces que se ven alteradas por ella y puede provocar abundantes cambios en la nomenclatura del diccionario, algo que es bien visible en la quinta edición (*DRAE-1817*).

Se registran, además, otros casos cuya variación gráfica experimenta mayores oscilaciones. *Almoradux*, por ejemplo, se encuentra desde el *Diccionario de autoridades* hasta la sexta edición (*DRAE-1822*) con la grafía *x* final; a partir de la séptima edición

sustituye esta grafía por *j* (*DRAE-1832⁴-DLE-2014*) siguiendo la norma general iniciada en la ortografía de 1815 y en la quinta edición del diccionario. La variante *almoradux*, sin embargo, es reintroducida en el *DRAE-1884* y permanece en el diccionario hasta la última edición (*DLE-2014*, cf. epígrafe 2.3.3).

Algunos cambios gráficos no son generales, sino que se producen de manera específica solo en algunas voces y en distintas ediciones, como, por ejemplo, el que afecta al vocablo *aljecería*. Desde el *Diccionario de autoridades* (1726) hasta el *DRAE-1803*, la voz se registra con las grafías *g* y *c*: *algecería*. Desde la quinta edición (*DRAE-1817*) hasta la undécima (*DRAE-1869*), se cambia la escritura del vocablo sustituyendo la *g* por *j* y la *c* por *z*: *aljezería*. En la siguiente edición, de 1884, se vuelve a modificar la grafía recuperando la forma antigua *algecería* (*DRAE-1884-1899*). Finalmente, desde la edición de 1914 hasta la actualidad, se produce un cambio a favor de la grafía *j* (*aljecería* 1914-2014) y se mantiene la *c*.

ALJECERÍA

vigencia cronológica: 1726-2014

formas:

<i>algecería</i>	1726-1803, 1884-1899
<i>aljezería</i>	1817-1869
<i>aljecería</i>	1914-2014

Pueden observarse en este ejemplo las vacilaciones en la norma gráfica que experimentan algunas palabras, que reflejan la variación que ha acompañado a su historia textual, algo especialmente importante en ciertos componentes genealógicos como los arabismos.

Son también destacables las modificaciones que registran los latinismos cuya forma contiene grupos consonánticos cultos. En múltiples casos se registran variaciones que son indicio de la norma cambiante con respecto a este tipo de grafías. Un ejemplo de ello es el término *septuagésimo*, registrado con el grupo culto desde el *Diccionario de autoridades* (1739) hasta el *DRAE-1803* y desde la edición de 1843 hasta la actualidad. En la edición de 1817, sin embargo, el grupo culto desaparece, dando lugar a la variante *setuagésimo*, una forma de la voz que tiene vigencia hasta la duodécima edición (*DRAE-1817-1884*). Por lo tanto, desde la novena hasta la duodécima edición (*DRAE-1843-1884*) ambas variantes, *setuagésimo* y *septuagésimo*, conviven en el diccionario, con preferencia desde 1869 por la forma con el grupo culto, puesto que es la portadora de la definición. La alternancia *septuagésimo/setuagésimo* desaparece en la duodécima edición con la supresión de esta última forma.

SEPTUAGÉSIMO

vigencia cronológica: 1739-2014

formas:

<i>septuagésimo</i>	1739-1803, 1843-2014
<i>setuagésimo</i>	1817-1884

⁴ El cambio de *x* a *j* en posición final de palabra se produce en la edición de 1832 (*boj*, *erraj*, *reloj*, etc.).

En algunas ocasiones una forma no sustituye a la otra en una determinada edición, sino que se incluye en el diccionario como variante gráfica del término ya existente en la nomenclatura. Este es el caso de *obsuro*, presente en todas las ediciones del diccionario académico, y *oscuro*, introducido en la edición de 1817 con vigencia hasta la actualidad. Esta vez, la forma preferida por la Academia desde la quinta edición es la variante con la simplificación, puesto que la que posee el grupo culto remite a ella.

OSCURO, OBSCURO

vigencia cronológica: 1737-2014

formas:

<i>obsuro</i>	1737-2014
<i>oscuro</i>	1817-2014

La incorporación de estas variantes gráficas en el diccionario es, pues, un reflejo de la alternancia de normas en el tratamiento de los grupos cultos y la conservación de las formas más latinizantes por el prestigio de su origen.

2.3.2. Los criterios de lematización

Se percibe en los cambios que verifican los lemas recogidos en *Lemateca* una tendencia progresiva a la homogeneización de los criterios de lematización que introduce abundantes variaciones en la forma de los lemas.

2.3.2.1. Lematización y morfología

Los criterios de lematización del diccionario implican, en general, que la forma canónica del lema aparezca en singular y en masculino. La terminación femenina, por su parte, se puede recoger a continuación de la masculina, formando parte del mismo lema (*ciego, ga; político, ca; lechero, ra*). Existe, sin embargo, en los diccionarios estudiados gran heterogeneidad en la lematización de las formas según el género y ello puede ocasionar abundantes mudanzas de lematización en la trayectoria de algunas palabras. Por ejemplo, el lema *papagayo, ya* se encuentra desde el *Diccionario de autoridades* con dos entradas distintas, una para el femenino («La hembra del papagayo») y otra para el masculino («Ave de color verde...» y otras acepciones). Esta estructuración se mantiene hasta la sexta edición del diccionario usual (*DRAE-1822*), mientras que en la edición siguiente los dos lemas conforman una sola entrada (*papagayo, ya, DRAE-1832-1869*); el origen de la reorganización se encuentra muy posiblemente en la tendencia a la «economía» verificada en el *DRAE-1832*. Habrá que estudiar si los cambios de lematización que comporta la duodécima edición, muy relacionados con la distinción de entradas homógrafas, ya sea por criterios etimológicos, ya por criterios categoriales, podrían haber motivado la vuelta a la situación inicial con dos lemas independientes (*papagaya, papagayo*), una configuración que se mantiene hasta la vigesimoprimera edición (*DRAE-1884-2001*). Finalmente, y como consecuencia de los cambios introducidos en la última edición del diccionario (*DLE-2014*) en lo que se refiere a lematización del género, se vuelve a fundir en una única entrada la forma del masculino y del femenino.

PAPAGAYO, YA

vigencia cronológica: 1737-2014

formas:

<i>papagayo</i>	1737-1822, 1884-2001
<i>papagaya</i>	1737-1822, 1884-2001
<i>papagayo, ya</i>	1832-1869, 2014

También en algunos casos, se registra variación en el lema provocada por la categoría gramatical de número. Por ejemplo, la forma singular *tenaza* se registra desde el *Diccionario de autoridades* hasta la tercera edición del *Diccionario usual* (DRAE-1791). En las ediciones de 1803 y 1817 el lema *tenaza* es sustituido por el plural *tenazas*, que en su significado de ‘instrumento de hierro...’ (DRAE-1803: s. v. *tenazas*) incorpora la indicación morfológica de sustantivo femenino plural (s. f. p.). A partir de la edición de 1822, y hasta la actualidad, *tenazas* con el significado de ‘instrumento de hierro...’ pasa a ser tratado como acepción de su correlato en singular. Desde la edición de 1884, el diccionario indica a través de marcas pospuestas a la acepción la posibilidad que tienen estos nombres de objetos dobles de alternar «el singular y el plural para designar un solo objeto» (NGLE, 2009: 3.8r).

TENAZA

vigencia cronológica: 1739-2014

formas:

<i>tenaza</i>	1739-1791, 1822-2014
<i>tenazas</i>	1803-1817

No ilustran estos ejemplos particularidades de la evolución de las voces, sino la heterogeneidad de los criterios de lematización aplicados, una variabilidad que se ve incrementada por el trabajo que involucra cada nueva edición del diccionario. Estas formas podían aparecer registradas en las primeras ediciones del diccionario en dos lemas independientes, y la tendencia ha consistido en unificarlas en un solo lema: en el caso de la moción de género, por ejemplo, se agrupa el femenino con el masculino, aunque la casuística es compleja, en especial cuando existe variación categorial (sustantivo y adjetivo como en *lechero, ra*; cf. Battaner y Lahuerta, 2002), siendo habituales los cambios de criterio (*papagayo, ya*). Además, se registra una tendencia a añadir el femenino en formas que inicialmente solo estaban lematizadas como masculinas: *túrgido* (DRAE-1843-1852) > *túrgido, da* (DRAE-1869-DLE-2014), *viagero/viajero* (DRAE-1803-1869) > *viajero, ra* (DRAE-1884-DLE-2014).

Algo parecido ocurre en la lematización de las formas verbales. En una primera fase del diccionario los usos pronominales y los transitivos e intransitivos aparecían desdoblados en dos entradas distintas. La tendencia ha consistido en unificar los diferentes usos de un mismo verbo en una sola entrada encabezada por la variante sin el pronombre, elegida como forma canónica representante de todos los usos. La revisión de este tipo de verbos se llevó a cabo en la cuarta edición del diccionario (DRAE-1803) y, de manera más sistemática, en la séptima (DRAE-1832) (cf. Terrón, 2017).

2.3.2.2. Los lemas múltiples

Desde el *Diccionario de autoridades* hasta la edición de 1884 no es raro encontrar lemas múltiples, es decir, formados por más de una palabra. Un ejemplo de ello se encuentra en la voz *nómada*, una adición del DRAE-1822. La octava edición (DRAE-

1837) incorpora, además, la variante *nómada* junto a la ya existente *nómade* formando el lema múltiple *nómade ó nómada*, que se conserva con esta configuración hasta el *DRAE*-1869. A partir de la duodécima edición y hasta la actualidad, cada forma del lema múltiple se convierte en un lema independiente, aunque *nómade* remite en todas estas ediciones a *nómada*.

Comportamiento similar se halla en el tratamiento de los apreciativos. *Rabanillo* es un diminutivo presente en el diccionario académico desde el *Diccionario de autoridades*. Este lema experimenta una serie de cambios en su configuración, al incorporar las diferentes variantes diminutivas generadas a partir de la intervención de los sufijos diminutivos *-ito* e *-ico*. De este modo, en la edición de 1817 se añade *rabanito* junto a *rabanillo* para constituir el lema múltiple *rabanillo, to*, y en el *DRAE*-1832 se introduce también *rabanico*, por lo que la forma de ese lema múltiple queda como *rabanico, llo, to*, lema que, como la mayoría de las formas apreciativas, se suprime en la edición de 1869. Aquellas formas diminutivas que experimentan una lexicalización reingresan en el diccionario: *flequillo* aparece como diminutivo de *fleco* desde 1791 hasta 1869; en 1925 se reincorpora con dos acepciones, la que se refiere al diminutivo ‘m. d. de Fleco’ y la del significado no composicional ‘porción de cabello recortado...’ (*DRAE*-1925).

Los lemas múltiples, pues, se componen de varias formas muy relacionadas desde el punto de vista formal⁵ y, aunque ahorran espacio, dificultan la consulta. Se ha tendido, en el desarrollo de la técnica lexicográfica, a eliminar estos maridajes en favor del establecimiento de un lema para cada forma y a relacionar esas formas por medio de las remisiones.

2.3.2.3. Los lemas complejos

Aunque las unidades fraseológicas aparecen lematizadas en alguna de las voces que las integran, no es raro encontrar en los diccionarios estudiados algunos lemas constituidos por una locución adverbial: por ejemplo, *por ce o por be* aparece registrada como lema complejo desde la edición de 1803 hasta la de 1822, ordenada alfabéticamente por la *c*. A partir de la edición de 1832 y hasta la actualidad, esta unidad se recoge en el artículo encabezado por el lema *ce*, lo que supone una supresión en la séptima edición del lema complejo *por ce o por be*.

En otros casos, sin embargo, el lema complejo cambia su configuración y se hace lema simple. Por ejemplo, *tranzón de tierra*, recogido como lema complejo desde 1817 hasta 1884, pasa a registrarse como *tranzón* a partir de 1899 (1899-2014).

Se verifica, pues, una tendencia a la eliminación de este tipo de lemas, aunque este proceso no afecta a las locuciones latinas, como *ipso facto*, que se recoge como lema complejo en la macroestructura de todas las ediciones del diccionario académico (1734-2014).

2.3.3. Supresiones y «reincorporaciones»

Parte de la reestructuración de la nomenclatura del diccionario en la primera mitad de siglo XIX se sustenta en la supresión de cierto tipo de entradas, acción con la que se

⁵ Hay, sin embargo, ejemplos de lemas múltiples del tipo *doblescudo* o *anteojo* (*DRAE*-1817-1869) que reflejan una relación semántica de sinonimia.

quería conseguir un diccionario más *manual* (Clavería, en prensa). Si el diccionario de la Academia es un diccionario de uso que refleja la lengua del momento de su publicación, es razonable pensar que se eliminan las voces y acepciones que han caído en desuso y, en consecuencia, las que tienen marca de arcaísmo. Sin embargo, si nos fijamos en las palabras con marca diacrónica, la realidad es que en ningún caso, en las ediciones estudiadas, la eliminación de voces arcaicas es mayor que la de voces no arcaicas⁶. Que se mantuvieran estas palabras encuentra una explicación en el interés de la Academia por recoger, primero, y conservar, después, el léxico arcaico en las ediciones de su diccionario (Álvarez de Miranda, 2000: 45-46, Ruhstaller, 2003: 242).

Se suprime, entonces, del diccionario léxico no necesariamente arcaico; y en esa supresión no destacan los lemas con remisión o definición (perifrástica o sinonímica), aspecto de interés también, pues el modo de explicar las voces puede informar del mayor o menor uso de una voz, de la existencia de variantes (Castillo Peña 1992: 537, Muñoz, 2018: 211) o de realidades anticuadas explicadas por medio de una definición (Haensch 1982: 491-492)⁷. En las seis ediciones estudiadas hasta el momento, se halla un aumento progresivo de supresiones en las tres primeras ediciones (*DRAE*-1817-1832), de forma que se puede establecer una distinción entre un primer grupo formado por estas ediciones y un segundo grupo formado por las tres ediciones siguientes (*DRAE*-1837-1852): de las 10 000 supresiones identificadas, la gran mayoría (96 %) se produce en las tres primeras y solo 4 % en las tres últimas. El mayor número de supresiones tiene lugar en la sexta y séptima ediciones: en la de 1822, se eliminan variantes gráficas, muchas de ellas arcaizantes, y en la de 1832, participios pasivos. Obedecen estas supresiones al deseo de hacer del diccionario una obra más manejable.

A estos cambios se unen otros que responden al modo de hacer el diccionario, a la concepción que se tiene de él y a la función que pretende cumplir. La Academia siempre ha mostrado interés por el léxico del pasado; lo prueba su presencia constante en el diccionario. Este interés explica la reincorporación en ediciones posteriores de algunos de los arcaísmos suprimidos. Si nos fijamos ahora en las tres primeras ediciones y en lo sucedido en la letra A, se observan casos de reincorporación, más abundantes en las ediciones con más supresiones:

Letra A	Supresión	Reincorporación
1817	30	2
1822	468	72
1832	372	25

Tabla 1. Supresión y reincorporación de arcaísmos

Los arcaísmos eliminados en la quinta edición, luego reincorporados, son *abaz* y *albañir*: uno, un arcaísmo léxico, en 1899, y otro, una variante anticuada, un arcaísmo

⁶ En las ediciones quinta, sexta y séptima se suprimen 877, 2373 y 6414 lemas y formas de lemas respectivamente; de ellos solo 367, 1728 y 763 en cada una de esas ediciones son arcaicos.

⁷ De las 877 supresiones de la quinta edición, 535 tienen remisión y 342 definición; en cambio, en la décima edición, de 173, 74 presentan remisión y 99 definición. Las diferencias entre un tipo de supresión y otro no son significativas, y las conclusiones que podrían obtenerse del examen de una edición no son confirmadas por la otra. Sin embargo, en la sexta edición, de 2373 supresiones, 1626 tienen remisión y marca de arcaísmo.

gráfico, en 1925. De los setenta y dos suprimidos en la sexta que vuelven posteriormente al diccionario, veintiocho experimentan cambios, bien de marca, bien de definición⁸; los cuarenta y cuatro restantes presentan el mismo tratamiento. De los veinticinco que lo hacen en la séptima, trece eliminan la marca de arcaísmo y cinco experimentan cambio en la definición⁹.

Las voces eliminadas que vuelven con el tiempo a estar presentes en el diccionario deben ser analizadas detenidamente para determinar si realmente son reincorporaciones que pueden implicar una vuelta al pasado o bien se trata de palabras que coinciden en la forma, aunque no en el contenido, por lo que no serían exactamente recuperaciones de lemas antiguos. Es esta una interpretación que resulta de poder comparar la información de las distintas ediciones.

Los ejemplos señalados a continuación prueban lo que se acaba de decir. La reincorporación de las supresiones o la adición de formas que ya habían estado en el diccionario pueden comportar cambios consistentes en la eliminación de la marca de arcaísmo y en la modificación de la definición. Si lo primero no supone lo segundo, lo segundo sí hace que se produzca lo primero, es decir, la modificación de la definición comporta cambios en la marcación. Así, por ejemplo, *absolvedor*, arcaísmo eliminado en 1822, se reincorpora en 1936 sin marca. En la quinta edición, de 1817, tiene la categoría gramatical de nombre; en 1936, la de adjetivo. En aquella edición se marcaba así porque mantenía la consideración presente en el *Diccionario de autoridades*, si bien con una definición simplificada:

ABSOLVEDOR. s. m. El que absuelve ò confiere la absolución de los pecádos, o censúras Eclesiásticas, en que uno huviere incurrido. Es verbal del verbo Absolver. No tiene mucho uso (*Diccionario de autoridades* 1726).

ABSOLVEDOR. s. m. ant. El que absuelve. *Absolvens* (*DRAE*-1817).

ABSOLVEDOR. adj. Que absuelve. Ú. t. c. s. (*DRAE*-1936).

Al aparecer de nuevo esta palabra en el diccionario, se marca como adjetivo usado también como sustantivo y ya no tiene ninguna conexión con la voz eliminada; por tanto, no sería un caso de reincorporación.

Con *abusador* acontece una situación similar. La voz aparece por primera vez recogida en 1780 marcada como arcaísmo y es eliminada en 1822:

ABUSADOR. s. m. ant. Lo mismo que ADIVINADOR (*DRAE*-1780S).

ABUSADOR, RA. adj. *Chile*. Abusón (*DRAE*-1956).

⁸ Se les elimina la marca a *abarrer*; *abastant*; *abondo*; *aciprés*; *acontentar*; *adolorado, da*; *ajabeba*; *alba-laero*; *alcarchofa*; *alcatara*; *almueza*; *anotomía*; *aprensionar*; *apócema*; *ardil*; *arritranca*; *asosiego*; *aviespa*. Hay cambio de definición en *absolvedor*; *abusador*; *ajabeba*; *alcarchofado, da*; *almojama*; *ambidos*; *Aparicio*; *ardil*; *aristin*; *arrollador*.

⁹ Se elimina la marca de arcaísmo en *abandar*; *acornado, da*; *adjurado, da*; *adverado, da*; *alcaria*; *alguinio*; *alienado, da*; *almajal*; *almojáter*; *almojat্রে*; *antuviado, da*; *atorgar*; *aventado, da*. Hay cambio en la definición en *adverado, da*; *afondado, da*; *alechigado, da*; *alongado, da*; *alumnado, da*, esta última con cambio de clase de palabra.

Al reincorporarse en 1956 se define con un sinónimo de la misma familia léxica, pero, al igual que en *absolvedor*, esta voz ya no tiene relación con la que se suprimió. Se ha eliminado una definición referida a una realidad anticuada, en este caso a los *adivinos*. El *Diccionario histórico* de 1933 permite aclarar su origen:

ABUSADOR. m. ant. Engatusador, adivino «Quando esta ave [el mochuelo] aúlla de noche, segund los adivinadores o abusadores, es señal de muerte.» Burgos, *De las prop. de las cosas*, lib. 12, cap. 6.

El *Corpus del Nuevo diccionario histórico (CDH)* ofrece 81 ocurrencias en 41 documentos¹⁰ de la palabra *abusador*; la primera es este texto de la traducción de Vicente de Burgos; la segunda está fechada ya en 1832. Al documentarse la voz por primera vez en el suplemento de la edición de 1780, desprovista de cualquier referencia textual, no es posible explicar el significado dado a la voz ni su consideración de arcaísmo. El testimonio, primero del *Diccionario histórico* de 1933 y luego de los corpus, permite aclarar su origen¹¹.

Caso distinto es el de la reintroducción de *almoradux* al que nos hemos referido anteriormente (*vid.* 2.3.1), que provoca que actualmente el diccionario albergue tres lemas distintos (*almoraduj*, *almoradux*, *moradux*) que son variantes sin que se establezca una relación clara entre ellas¹². Al margen del valor histórico-textual de la presencia de estas formas en el diccionario, no podemos evitar preguntarnos por la conveniencia de su aparición en la última edición, en la que atentan contra la eficiencia de una obra lexicográfica moderna (Gutiérrez Cuadrado, 2001-2002).

ALMORADUJ, ALMORADUX, MORADUX	
<i>almoradux</i>	1726-1822, 1884-2014
<i>almoraduj</i>	1832-2014
<i>moradux</i>	1925-2014

Las supresiones, sus datos cuantitativos, el modo como se recogen —con marca o sin marca, con remisión o definición— y la suerte en ediciones posteriores resultan de disponer de esta *Lemateca* del *DRAE*, de las bases de datos que la alimentan y de la información que la complementa (vigencia e historia lexicográfica, marcación, remisión, existencia de variantes, etc.). Lo ofrecido aquí sirve, además, para mostrar la utilidad y el aprovechamiento de esta nueva herramienta lexicográfica.

2.3.4. Las erratas

Las erratas de imprenta o, incluso, las correcciones inadecuadas introducidas en alguna de las ediciones del diccionario generan lemas aparentemente desordenados y, en las consultas a través del *NTLLE*, discontinuidades en la documentación de los lemas; hacen aparición de esta forma palabras fantasma y hápax (Clavería y Freixas, 2018: 131-132).

¹⁰ Son datos referidos a la forma *abusador*; se dejan de lado ahora las formas femenina y plural.

¹¹ Es un caso peculiar, por ser una voz documentada en un solo texto. Pero no es el único en que se recurrió a la traducción de esta enciclopedia medieval para incorporar alguna de sus palabras al diccionario (Jiménez Ríos 2012: 169, 2018: 422).

¹² *Almoraduj* y *moradux* remiten a *mejorana*, *almoradux* remite a *almoraduj* (*DLE*-2014). Casos paralelos se registran en otras palabras con *x/j* en posición final: *aj/ax*, *carcaj/carcax*, *gambuj/gambux*, *troj/trox*. A estos casos se pueden añadir *dix* (ant. *dije*) y *saxafrax* (*saxífraga*).

Un ejemplo de errata de imprenta se encuentra en el «hápx» *nardinc* de la edición de 1837 en lugar de *nardino*¹³, un lema que se encuentra en todas las ediciones del diccionario desde *Autoridades. Lemateca* permite establecer relación entre *nardino* y la errata *nardinc*.

NARDINO

vigencia cronológica: 1734-2014

formas:

nardino 1734-1832, 1843-2014

nardinc (errata) 1837

Las erratas pueden mostrar mayor o menor pervivencia en la historia textual del diccionario; así, el vocablo *arranzon*, presente desde el suplemento de la primera edición del *Diccionario usual (DRAE-1780)* hasta la edición de 1925, muestra una discontinuidad en su vigencia cronológica que abarca desde la novena (*DRAE-1843*) hasta la undécima edición (*DRAE-1869*). Esta interrupción en su historia lexicográfica se debe a que en esas ediciones el lema se introduce con la forma errónea *arrazon*, que no sigue aparentemente el orden alfabético.

ARRANZON

vigencia cronológica: 1780S-1925

formas:

arranzon 1780S-1837, 1884-1925

arrazon (errata) 1843-1869

Estos ejemplos ilustran el hecho de que la existencia de erratas, no siempre fáciles de detectar, dan una visión distorsionada de la trayectoria lexicográfica de las voces que las presentan, lo que intenta subsanarse en *Lemateca* en el caso de aquellas que se han podido identificar.

3. CONCLUSIÓN

Surge, pues, de los datos de *Lemateca* la reconstrucción de los distintos avatares de cada palabra en las páginas de los diccionarios académicos. Se relaciona esta con la propia historia de la voz, en especial con su forma ortográfica, y con su evolución. Todos los cambios observados contribuyen también a trazar la historia de la lexicografía académica.

Como se ha demostrado con la descripción de esta herramienta, es necesario crear puentes entre la investigación y las humanidades digitales. La elaboración de trabajos científicos basados en herramientas informáticas no solo permite la gestión y explotación de un conjunto de datos mucho más amplio, sino que también se constituye en semillero que brinda al investigador nuevas vías para la interpretación de los datos y los resultados, y también nuevos caminos que transitar a partir de las preguntas que dimanan de su investigación.

¹³ La errata aparece recogida en la fe de erratas de esa misma edición.

REFERENCIAS BIBLIOGRÁFICAS

- ALCOBA, Santiago (2007): «El debate de la reforma ortográfica y Andrés Bello», *Español Actual*, 88, pp. 127-172.
- ALCOBA, Santiago (2012): «El proceso de fijación ortográfica de las palabras en los DRAE», en Gloria Clavería, Margarita Freixas, Marta Prat y Joan Torruella, eds., *Historia del léxico: perspectivas de investigación*, Madrid-Frankfurt am Main, Iberoamericana-Vervuert, pp. 273-302.
- ÁLVAREZ DE MIRANDA, Pedro (2000): «La lexicografía académica de los siglos XVIII y XIX», en Ignacio Ahumada, ed., *Cinco siglos de lexicografía del español*, Jaén, Universidad de Jaén, pp. 35-61.
- Autoridades 1726* = REAL ACADEMIA ESPAÑOLA (1726-1739): *Diccionario de la lengua castellana, en que se explica el verdadero sentido de las voces, su naturaleza y calidad con las frases o modos de hablar, los proverbios o refranes, y otras cosas convenientes al uso de la lengua*, 6 vols., Madrid, Imprenta Francisco del Hierro.
- BATTANER, Paz y Javier LAHUERTA (2002): «Nombres femeninos en el diccionario: problemas de lematización», en José Ignacio Pérez Pascual y Mar Campos Souto, coords., *Cuestiones de lexicografía*, Lugo, Tris Tram, pp. 53-62.
- BLANCO IZQUIERDO, M.^a Ángeles (2018): «El contexto del diccionario: los códigos académicos vigentes en 1817», en Gloria Clavería y Margarita Freixas, coords., *El diccionario de la Academia en el siglo XIX: la quinta edición (1817) al microscopio*, Madrid, Arco/Libros, pp. 57-63.
- BLANCO IZQUIERDO, M.^a Ángeles, Gloria CLAVERÍA y Enrique JIMÉNEZ RÍOS (2018): «Fuentes lexicográficas y estudio del léxico: el *Diccionario de la lengua castellana* de la Real Academia Española (1817-1852)», en Dolores Corbella, Alejandro Fajardo y Jutta Langenbacher, eds., *Historia del léxico español y humanidades digitales*, Frankfurt am Main, Peter Lang, pp. 449-476.
- CASTILLO PEÑA, Carmen (1992): «La definición sinonímica y los círculos viciosos», *Boletín de la Real Academia Española*, LXXII, pp. 463-566.
- CATACH, Nina (1989): «L'orthographe dans le dictionnaire monolingue», en Franz Joseph Hausmann et al., eds., *Wörterbücher. Dictionaries. Dictionnaires. Ein Internationales Handbuch zur Lexikographie*, Berlín-Nueva York, W. de Gruyter, vol. I, pp. 501-508.
- CDH = INSTITUTO DE INVESTIGACIÓN RAFAEL LAPESA DE LA REAL ACADEMIA ESPAÑOLA (2013): *Corpus del Nuevo diccionario histórico (CDH)*. [Fecha de consulta 20/02/2019: <http://web.frl.es/CNDHE>].
- CLAVERÍA, Gloria (en prensa, 2020): «El *Diccionario de la lengua castellana* (1817-1832) y la propuesta decimonónica de diccionario manual», *Boletín de la Real Academia Española*, tomo C, cuaderno 322.
- CLAVERÍA, Gloria y Margarita FREIXAS (2018): «El *Nuevo tesoro lexicográfico de la lengua española*: un museo lexicográfico como base de datos», *Cuadernos del Instituto Historia de la Lengua*, 11, pp. 117-138.
- DLE-2014 = REAL ACADEMIA ESPAÑOLA Y ASOCIACIÓN DE ACADEMIAS DE LA LENGUA ESPAÑOLA (2014): *Diccionario de la lengua española*, vigesimotercera edición, Madrid, Espasa.
- DYSON, Freeman (1997): *Imagined Worlds*, Boston, Harvard University Press.

- GUTIÉRREZ CUADRADO, Juan (2001-2002): «El nuevo rumbo de la vigésima segunda edición (2001) del *Diccionario de la lengua española* de la Real Academia», *Revista de Lexicografía*, VIII, pp. 297-319.
- HAENSCH, Günther *et al.* (1982): *La lexicografía. De la lingüística teórica a la lexicografía práctica*, Madrid, Gredos.
- INSTITUTO DE INVESTIGACIÓN RAFAEL LAPESA DE LA REAL ACADEMIA ESPAÑOLA (2013): *Mapa de diccionarios* [en línea]. [Fecha de consulta 20/02/2019: <http://web.frl.es/ntllet>].
- JIMÉNEZ RÍOS, Enrique (2012): «Huellas de la traducción de Vicente de Burgos del *De Proprietatibus Rerum* en la lexicografía de la Real Academia Española», en José Antonio Bartol *et al.*, coords., *Estudios de filología española*, Salamanca, Luso Española de Ediciones, pp. 167-174.
- JIMÉNEZ RÍOS, Enrique (2018): «La revisión de arcaísmos en el *DRAE* en la primera mitad del siglo XIX», *Estudios humanísticos. Filología*, 40, pp. 407-431.
- MUÑOZ, Laura (2018): «La sinonimia y las remisiones», en Gloria Clavería y Margarita Freixas, coords., *El Diccionario de la Academia en el siglo XIX: la quinta edición (1817) al microscopio*, Madrid, Arco/Libros, pp. 203-222.
- NGLE: REAL ACADEMIA ESPAÑOLA y ASOCIACIÓN DE ACADEMIAS DE LA LENGUA ESPAÑOLA (2009-2011): *Nueva gramática de la lengua española*, 3 vols., Madrid, Espasa.
- NTLLE = REAL ACADEMIA ESPAÑOLA (2001): *Nuevo tesoro lexicográfico de la lengua española*, edición DVD, Madrid, Espasa-Calpe. [Fecha de consulta 20/02/2019: <http://ntlle.rae.es>].
- ORAE 1815 = REAL ACADEMIA ESPAÑOLA (1815): *Ortografía de la lengua castellana*, octava edición notablemente reformada y corregida, Madrid, Imprenta Real.
- QUILIS MERÍN, Mercedes (2009): «Diccionarios y normas ortográficas: panorama y aplicaciones en la lexicografía española de los siglos XVIII al XXI», en Elena de Miguel *et al.*, eds., *Fronteras de un diccionario: las palabras en movimiento*, San Millán de la Cogolla, Cilengua, pp.83-120.
- ROJO, Guillermo (2009): «Sobre la construcción de diccionarios basados en corpus», *Revista Tradumàtica. Traducció i Tecnologies de la Informació i la Comunicació*, 7, pp. 1-7.
- RUHSTALLER, Stefan (2003): «Las obras lexicográficas de la Academia», en Antonia M.^a Medina Guerra, coord., *Lexicografía española*, Barcelona, Ariel, pp. 235-261.
- TERRÓN, Natalia (2017): «Gramática y diccionario: aproximación al tratamiento de los usos pronominales de los verbos en la lexicografía académica», en Ignacio Sariego López, Juan Gutiérrez Cuadrado y Cecilio Garriga, eds., *El diccionario en la encrucijada: de la sintaxis y la cultura al desafío digital*, Santander, Escuela Universitaria de Turismo de Altamira-Asociación Española de Lexicografía Hispánica, pp. 917-936.
- TERRÓN, Natalia (2018): «La regularización ortográfica», en Gloria Clavería y Margarita Freixas, coords., *El diccionario de la Academia en el siglo XIX: la quinta edición (1817) al microscopio*, Madrid, Arco/Libros, pp. 67-91.
- TORRUELLA, Joan (2017): *Lingüística de corpus: génesis y bases metodológicas de los corpus (históricos) para la investigación en lingüística*, Frankfurt am Main, Peter Lang.