

***Portal de Léxico Hispánico: un recurso electrónico
para el estudio histórico del léxico****

Gloria Clavería

Universidad Autónoma de Barcelona

Carolina Julià

Universidad Autónoma de Barcelona

Mar Massanell

Universidad Autónoma de Barcelona

Joan Torruella

ICREA-Universidad Autónoma de Barcelona

Recibido: 19-11-2013

Aceptado: 21-11-2013

Resumen: El presente artículo constituye una descripción del *Portal de Léxico Hispánico*, una herramienta informática cuyo objetivo principal es ofrecer información para el estudio histórico del léxico iberorrománico a partir de la compilación de datos que actualmente se encuentran dispersos en multiplicidad de publicaciones. Este gran banco de datos se estructura en tres apartados que están interrelacionados y que recogen informaciones lingüísticas (*Banco de léxico*), bibliográficas (*Banco de bibliografía*) y documentales (*Banco de documentaciones*). Esta última base de datos, que es el principal objeto de descripción de este trabajo, actualmente contiene informaciones sobre el español y el catalán cuya procedencia es muy

* La investigación necesaria para llevar a cabo este artículo ha sido posible gracias a la ayuda del *Ministerio de Economía y Competitividad* para el proyecto «*Portal de Léxico Hispánico: documentación y morfología derivativa*» (n.º de referencia FFI2011-24183) y al apoyo del *Comissionat per Universitats i Recerca de la Generalitat de Catalunya* concedido al «Grupo de Lexicografía y Diacronía» (n.º de referencia SGR2009-1067).

diversa. Para el español, se parte esencialmente de las documentaciones del *DECH* y para el catalán, de datos documentales que atesora una obra lexicográfica inédita, el *Essai de dictionnaire historique de la langue catalane* de Julià-Bernat Alart.

Palabras clave: documentación, español, catalán, *DECH*, Julià-Bernat Alart.

Abstract: The present article constitutes a description of the *Portal de Lèxic Hispànic (Hispanic vocabulary Portal)*, an electronic tool whose main aim is to offer information for the historical study of Ibero-Romance vocabulary from the compilation of data which is currently scattered over a great variety of publications. This database is structured in three interrelated parts, gathering different kinds of information: linguistic (*Banco de lèxic*), bibliographic (*Banco de bibliografia*) and documental (*Banco de documentaciones*). The last, the main aim of description in this publication, currently has information about Spanish and Catalan of very different origin. Regarding Spanish, documentation from *DECH* is the starting point, whereas for Catalan the documentary data are taken from an unpublished lexicographical work: *Essai de dictionnaire historique de la langue catalane*, by Julià-Bernat Alart.

Keywords: Documentation, Spanish, Catalan, *DECH*, Julià-Bernat Alart.

1. PRESENTACIÓN

La investigación actual no puede concebirse fuera del trabajo en equipo y la aplicación de herramientas informáticas. Existen importantes grupos de investigación e instituciones que están desarrollando distintas líneas de análisis que, de una forma u otra, intentan contribuir a la generación de conocimiento suficiente que permita elaborar un diccionario histórico de la lengua española (Dworkin, 1994; Álvarez de Miranda, 2005 y 2006; Garcés, 2008; Pascual y García Pérez, 2007; Gómez y Carriazo, 2010; Clavería, 2012), un tipo de obra de cuya falta se resienten desde hace mucho tiempo tanto los estudios históricos como los sincrónicos. Además, un proyecto de tal envergadura no puede ser acometido actualmente sin tener en cuenta todas las aportaciones conseguidas en el estudio del léxico desde el punto de vista histórico y uno de los problemas a los que tiene que enfrentarse la investigación en esta materia a principios de siglo XXI es la gran dispersión y amplia tipología que presentan las publicaciones, que pueden ir desde glosarios dedicados a un componente léxico o a

una obra o a autor determinados hasta monografías sobre la historia de una palabra o grupo de palabras (Clavería, 2012). Esta fragmentación bibliográfica ya fue motivo de atención en un artículo de Pellen (2003: 74) en el que al evaluar el estado actual de los estudios diacrónicos se refería a los problemas de la «gestión de los conocimientos desparramados en un sinfín de revistas y libros, sin que la variedad creciente de los soportes consiga remediar esta fragmentación».

En este panorama, el *Portal de Léxico Hispánico (PLH)* pretende ser una herramienta que ayude a paliar esta dispersión y fragmentación de los estudios dedicados al léxico. Su objetivo es el de configurarse como un repositorio de datos en permanente ampliación y en acceso abierto que facilite la información de los investigadores en la materia. La plataforma ha sido desarrollada por el «Grupo de Lexicografía y Diacronía» del Departamento de Filología Española de la Universidad Autónoma de Barcelona, un grupo de investigación especializado en el desarrollo de herramientas informáticas aplicadas al estudio histórico, tal como se demuestra en el propio *PLH*, en la edición digital del *DECH* (Coromines y Pascual, 2012), en la publicación en CD-ROM del *Índex del Lèxic de l'obra de G. Colón* (Clavería *et al.*, 2005) o en la elaboración del *CICA (Corpus informatitzat del català antic*, Torruella, 2009a, 2009b, 2009c, 2012; Clavería y Torruella, 2012). El *PLH* pretende ser un centro de información en el que se albergan datos de las distintas variedades lingüísticas hispánicas (Clavería *et al.*, 2008).

Actualmente el *PLH* se está desarrollando mediante distintas ayudas institucionales para el proyecto «*Portal de Léxico Hispánico: documentación y morfología derivativa*» (n.º de referencia FFI2011-24183) con el que se acomete tanto el mantenimiento del *PLH* como el avance en la investigación de la documentación histórica del léxico y, en especial, de los mecanismos de creación de voces con reglas de formación de palabras. En este estudio se expondrán los trabajos desarrollados en la primera línea de investigación: la documentación del léxico (§ 3).

2. ESTRUCTURA Y CONTENIDO DEL PORTAL DE LÉXICO HISPÁNICO

El *PLH* se organiza en tres bancos de datos con contenido y características diferentes que son complementarios y se encuentran interrelacionados y conectados mediante saltos hipertextuales. Estos son el *Banco de léxico*, el *Banco de bibliografía* y el *Banco de documentaciones*.

El *Banco de léxico* incluye información de dos tipos: de carácter lingüístico-filológico (lema, forma, categoría gramatical, significado, lengua y tipo de in-

formación) y de localización (referencia bibliográfica en la que se encuentra). Se compone, actualmente, de las voces contenidas en el *Diccionario crítico etimológico castellano e hispánico* (DECH) de J. Corominas y J. A. Pascual según la versión en CD-ROM (2012) e integra también los lemas aparecidos en los estudios y monografías del *Banco de bibliografía* con el que se halla relacionado.

El *Banco de bibliografía* está constituido fundamentalmente por referencias bibliográficas de estudios publicados en homenajes, actas de congresos, libros misceláneos y bibliografías de autor. Hasta estos momentos se han introducido en él las monografías léxicas de especialistas en el estudio histórico del léxico con amplia y dispersa producción como Germán Colón (Clavería *et al.*, 2005), Margherita Morreale (Rivarola y Pérez, 2006) y Yakov Malkiel (Duggan y Faulhaber, 1988). Este banco también contiene todos los estudios que son adiciones al DECH (Blecuá y Clavería, 1999; Ferrer, 2008).

El *Banco de documentaciones* tiene como finalidad primordial recoger documentaciones del léxico hispánico. Actualmente se está ampliando este banco de datos en dos lenguas distintas: el español y el catalán. Para el español, el objetivo fundamental es la complementación de las dataciones del DECH (§ 3) mientras que para el catalán se está trabajando con los materiales de las documentaciones que contiene el *Essai de dictionnaire historique de la langue catalane* de J.-B. Alart (§ 4).

3. BANCO DE DOCUMENTACIONES

El *Banco de documentaciones* del PLH almacena y gestiona informaciones documentales del léxico iberorrománico. En estos momentos, este banco de datos contiene más de 48.000 registros con información de carácter documental (autor, fecha, obra, texto, etc.) del léxico español.

Estas informaciones documentales poseen una doble procedencia. Por un lado, el PLH incluye todas las documentaciones del *Diccionario crítico etimológico castellano e hispánico* de Joan Corominas y José Antonio Pascual (1981-1990) a partir de la edición en CD-ROM (Corominas y Pascual, 2012) y, por otro lado, contiene registros que incluyen documentaciones adicionales al DECH. Estas últimas tienen distinto origen y siempre toman como punto de partida la información documental proporcionada por este diccionario.

A continuación se describe el funcionamiento general del banco de datos en relación con las otras bases de datos que conforman el PLH (§ 3.1.)

y, posteriormente, se exponen las características de los registros según la fuente de información de la que proceden (§ 3.2.).

3.1. Organización y funcionamiento del *Banco de documentaciones*

El *Banco de documentaciones* está directamente relacionado con los otros dos bancos de datos del *PLH*, el de léxico y el de bibliografía (vid. § 1), pues los lemas que contiene pueden también presentar registros en el *Banco de léxico* y todas las obras de las que proceden las informaciones documentales se encuentran incluidas en el *Banco de bibliografía*. Los datos que permite recuperar son de carácter léxico, documental y bibliográfico.

Desde el punto de vista del léxico se pueden obtener las documentaciones de los lemas que se consultan. Así, por ejemplo, se pueden recuperar todos los registros documentales que incluye el *PLH* de la palabra *brocal*. Se trata de 4 fichas, una de las cuales pertenece al artículo «Palabras del siglo XVI» de Esther Hernández (1996) y se corresponde con una documentación de Fray Alonso de Molina (*Vocabulario de lengua mexicana y castellana*, 1571) y las otras tres proceden del *DECH* (1356 y 1373, *Inventarios aragoneses de los siglos XIV y XV* y Elio Antonio de Nebrija, 1581).

Desde la búsqueda por documentaciones, el *PLH* permite recuperar los registros relativos a un autor u obra concretos o a una fecha determinada. Si se realiza una búsqueda por *Miguel de Cervantes Saavedra*, el banco de datos proporciona 459 documentaciones pertenecientes de 443 lemas¹. Asimismo, si se busca por obra *Quijote (de M. de Cervantes)*², se hallan 255 documentaciones relativas a 253 lemas. La consulta de las documentaciones también se puede realizar a partir de una fecha determinada —p. e. de 1492 se recogen 111 documentaciones de 111 lemas distintos³— o de un período cronológico concreto (p. e. siglo XI). En este caso, la búsqueda realizada para el intervalo (1000-1099) da como resul-

¹ Cabe señalar que algunos lemas presentan más de una documentación, de ahí la diferencia numérica entre el número de lemas y el de documentaciones. Así, por ejemplo, el lema *chantar* presenta dos documentaciones del mismo autor (*Quijote*, 1615 y *Ocho comedias y ocho entremeses nuevos nunca representados*, 1615) y otros, como *madrugón* (*Novelas Ejemplares*, 1613) y *magia* (*Quijote*, 1615), contienen un solo registro.

² Para discriminar este *Quijote* del de Alonso Fernández de Avellaneda, el *PLH* distingue las obras incluyendo el nombre del autor tras el título (*Quijote (de M. de Cervantes)* y *Quijote (de A. Fernández de Avellaneda)*). Para más información sobre la normalización de la nomenclatura de las obras que aparecen citadas en el *DECH*, véase Torruella (2006: 86).

³ Véanse algunos casos como *adherir* (Alonso Fernández de Palencia, *Crónica de Enrique IV*), *anteojo* (Elio Antonio de Nebrija) y *rubarbo* (*Lexicon ex sermone Latino in Hispaniense*, Cristóbal Colón (Cartas C. Colón)).

tado 11 registros, entre ellos, *enfermo* (Anónimo Sevillano, *Glosario de voces romances*), *escanda* (*Glosario Latino-Arábigo de Leyden*) o *mama*.

Además, el PLH proporciona la información sobre la fuente bibliográfica de la que proceden las documentaciones incluidas en el banco de datos. Se trata, en este caso, de los estudios en los que se aportan las documentaciones. Así, se pueden recuperar las documentaciones que un investigador ofrece en su(s) trabajo(s). Por ejemplo, el *Banco de documentaciones* incluye 1784 registros procedentes de la tesis doctoral de Laura Muñoz Armijo (*La historia de los derivados en -ismo e -ista en español moderno*, 2010) y estos se pueden recuperar buscando o bien por el nombre de la autora o bien por el título del estudio.

El PLH permite filtrar las informaciones obtenidas a partir de la combinación de los distintos parámetros de las búsquedas simples mencionadas. Por ejemplo, se puede obtener la lista de todos los lemas que contienen una determinada secuencia de caracteres, documentados en una fecha concreta o en un autor u obra que se encuentren reseñados en uno de los artículos que se halle en el *Banco de bibliografía*. Por ejemplo, este tipo de búsqueda proporcionaría las documentaciones de las palabras acabadas en *-itis*, documentadas en el siglo XVIII, que tienen en el campo autor «Academia Española». En este caso, aparecen documentaciones de 3 lemas con sus correspondientes registros: *capnitis*, *raquitis* (con dos variantes formales: *rachitis* y *rhachitis*) y *sideritis*.

3.2. Datos que contiene el *Banco de documentaciones*

Actualmente, este banco de datos contiene un corpus de 48.000 documentaciones textuales y lexicográficas del léxico del español. Los datos almacenados en él proceden o bien de las documentaciones del *Diccionario crítico etimológico castellano e hispánico* (§ 3.2.1.) o bien de las documentaciones de otras fuentes de datos (§ 3.2.2.).

3.2.1. Documentaciones procedentes del DECH

Los registros de las documentaciones del DECH que contiene el PLH constituyen un corpus integrado por 41.813 dataciones que se incluyen en este diccionario. Se trata de un grupo de registros obtenido a través del análisis crítico de los datos documentales de la obra (Blecua y Clavería, 1999; Clavería, 1999, 2007, 2008; Clavería y Torruella, 2005; Torruella, 2006) llevados a cabo en el marco de los trabajos de digitalización del diccionario que se publicó en formato CD-ROM en el 2012.

Un profundo estudio de las documentaciones del diccionario etimológico ha permitido ofrecer de forma homogénea en el *PLH* las informaciones que contiene. En primer lugar, esta herramienta proporciona la primera documentación discriminada (la primera documentación se marca con un *Sí* en el campo correspondiente y el resto, con un *No*), lo que es de especial relevancia, puesto que, como han comentado diversos investigadores, muchos de los lemas del *DECH* pueden llegar a recoger un nada despreciable número de documentaciones (Torruella, 2006: 78). Véase, por ejemplo, el caso de *abdarramía* o *abderramía*, voz antigua que «parece ser un compuesto de *ave*» (*DECH*, s. v. *ave*) sobre la que se dan seis documentaciones distintas (Don Juan Manuel, 1348; Johan de Sant Fagunt; Pedro López de Ayala, 1400; Gonzalo Fernández de Oviedo, 1526; Lope Félix de Vega Carpio, 1635; Padre Bernabé Cobo, 1653) o el de *abacero* que, como refleja la ficha que se halla en el *PLH* sobre este lema (imagen 1), presenta tres documentaciones en el *DECH*:

DECH
(s. v. *abacero*)

1.ª doc.: *fabacero*, S. XIII («las *fabaceras* e los *f.* non comprén pescado reziente para revender», *Fueros Leoneses*, publ. p. Castro y Onís, 135.5); *abacero*, h. 1400 (Glos. de Palacio), h. 1475 (*Gaya* de Segovia, 89)

PLH

Diccionario crítico etimológico castellano e hispánico	
LEMA: abacero, ra	
Sub voce:	abacero
Forma:	
Forma compleja	
Autor:	
Obra:	Fueros leoneses
Fecha:	1200_1299
Ejemplo:	fabaceras: las fabaceras e los f. non comprén pescado reziente para revender
Primera documentación:	Sí
Datos bibliográficos:	
Referencia DECH:	I: publ. p. Castro y Onís, 135.5.
Sub voce:	abacero
Forma:	
Forma compleja	
Autor:	
Obra:	Glosarios Latino-Españoles de la Edad Media
Fecha:	1400
Ejemplo:	
Primera documentación:	No
Datos bibliográficos:	
Referencia DECH:	D
Sub voce:	abacero
Forma:	
Forma compleja	
Autor:	Guillén de Segovia, Pedro
Obra:	La Gaya Ciencia
Fecha:	1475
Ejemplo:	
Primera documentación:	No
Datos bibliográficos:	

Imagen 1. Registros documentales del lema *abacero, ra* procedentes del *DECH*

En segundo lugar, el *PLH* presenta la homogeneización completa de las numerosas variantes denominativas de las obras y de los autores que constituyen las fuentes documentales del *DECH*. Así, por ejemplo, el caso de *Sor María de Jesús de Ágreda*, autora sobre la que el *PLH* recoge documentaciones de 26 lemas —para 21 de los cuales constituye la primera documentación—, cuyo nombre aparece en el diccionario con 9 formas designativas distintas (*Ágreda*, *Sor María de Jesús de Ágreda*, *M. Ágreda*, *M. de Ágreda*, *M. de Ágreda*, *M.^a de Ágreda*, *Ma. de Ágreda*, *Madre Ágreda*, *Sor María de Ágreda*). Asimismo, puede citarse el ejemplo de *Juan Alonso y de los Ruyzes de Fontecha*, autor unificado con este nombre, que aparece en el *DECH* con tres nombres distintos: *Ruices de Fontecha (chala-za)*, *R. Fontecha (fárfara II)*, *Fontecha (aroma)*.

Finalmente, los registros del *PLH* presentan también de forma unificada la información relativa a la fecha de documentación, un dato que se recoge de manera bastante asistemática en las páginas del diccionario etimológico, pues se pueden hallar desde períodos inferiores a un siglo (*berilo*: 1.^a mitad del siglo XV; *útero*: segundo cuarto del siglo XV) o períodos de varios siglos (*bravonet*: ss. XVI-XVII) hasta fechas concretas (*paladio*: 1914 o 1899), o las más variadas aproximaciones cronológicas (*acapillar*: algo antes de 1550; *parecer*: orígenes del idioma; *sollozo*: h. 1400) y fechas hipotéticas (*zurcir*: ¿h. 1300?). En todos los casos se han reducido las apreciaciones cronológicas aproximadas a datos numéricos para facilitar las búsquedas. Así, por ejemplo, para *andamio*, cuya primera documentación en el *DECH* es la «2.^a mitad del siglo X» (*Glosas de Silos*), se ha homogeneizado la datación como 950_999⁴. Esto permite generar búsquedas por períodos cronológicos bien delimitados a pesar de que el diccionario ofrezca, en ocasiones, datos aproximados y no una fecha determinada.

Cabe señalar que la homogeneización de las informaciones documentales no ha implicado cambios en los datos que proporciona el diccionario, es decir, el *PLH* ofrece las fichas de documentación tal cual se recogen en el diccionario, pues se ha respetado el original sin ningún tipo de corrección. Así, en el caso de las documentaciones de la Academia, el *DECH* men-

⁴ El guión bajo implica que se trata de una fecha no real, es decir, de una equivalencia de una fecha aproximada en el *DECH*. Por ejemplo, el intervalo 1249-1251 indica que se trata de este período de tres años, en cambio, 1249_1251 refleja que constituye una fecha aproximada, es decir, mediados del siglo XIII. Para más información sobre la normalización de las fechas y los símbolos empleados para ello, véase Torruella (2006: 86-89).

ciona que para el lema *paralizar* la voz aparece en «Acad. 1884, no 1843» y así lo refleja la documentación del *Banco de documentaciones* a pesar de que en el *NTLLE* se documenta desde 1852 (10.^a ed.).

3.2.2. Documentaciones adicionales al DECH

El *Banco de documentaciones* complementarias o adicionales al *DECH* contiene en la actualidad unos 6.500 registros. Esta información posee distintas procedencias ((a) documentaciones provenientes de fuentes lexicográficas y (b) documentaciones procedentes de estudios y reseñas sobre el *DECH*) y está vinculada a los trabajos que actualmente se están llevando a cabo por parte del grupo de investigación y siempre hallan como punto de partida la información documental del *DECH*, de la que se configuran como datos complementarios.

(a) Fuentes lexicográficas

En el léxico más moderno del español Joan Coromines utilizó las distintas ediciones del *DRAE* como una de las fuentes documentales básicas (Blecuá y Clavería, 1999). La primera documentación de una palabra en el *DRAE* tiene un valor relativo puesto que, aunque la primera datación en el diccionario no significa que se trate de su primer uso en la lengua, es un dato importante desde el punto de vista de la historia de la palabra por cuanto marca su admisión en el léxico oficial según los parámetros de recepción del léxico en el *DRAE* en cada momento de su historia. En este sentido, las veintidós ediciones del *Diccionario de la lengua española* de la Real Academia Española (1780-2001) junto al *Diccionario de Autoridades* (1726-1739) forman un corpus lexicográfico y documental extraordinario para la documentación y evolución de la recepción del léxico más moderno del español. Uno de los objetivos prioritarios de los estudios del grupo de investigación es el de reconstruir la historia de la ampliación del léxico en las distintas ediciones del *DRAE* y estudiar su significado (Clavería, 2001, 2004, 2005, 2007; Julià, 2008, 2012, en prensa a y b; Freixas, 2012; Muñoz Armijo, 2008, 2010a, 2010b, 2011, 2012a, 2012b y 2012c, 2013; Clavería y Freixas, en prensa; Huertas, en prensa a).

De hecho, el *DECH* usó abundantemente los diccionarios de la Academia aunque con ciertas limitaciones e inexactitudes (Clavería y Morales, 1999-2000). Por ejemplo, en el caso de *faldeta* el *DECH* proporciona la documentación aproximada «Acad. S. XX», mientras que al consultar el *NTLLE* se advierte que el término *faldeta*, como diminutivo de *falda*, se encuentra ya en el *DAut.* y que no es hasta la 12.^a ed. (1884)

cuando se incluye la acepción especializada referida a «en la maquinaria teatral, lienzo con que se encubre lo que ha de aparecer a su tiempo».

Actualmente, se está desarrollando una investigación detallada de la edición de 1817 (Clavería y Freixas, en prensa) y, de manera más general, de todos los diccionarios del siglo XIX (Clavería, en prensa). Asimismo, los estudios llevados a cabo en el marco de una de las principales líneas de investigación desarrolladas en el proyecto, la de la morfología léxica histórica, también toman como base documental las diferentes ediciones del diccionario académico (Clavería *et al.*, 2013; Muñoz Armijo, 2008, 2010a, 2010b, 2011, 2012a, 2012b y 2012c, 2013, en prensa a, b y c; Julià, 2012, en prensa a y b) —que habitualmente se complementan también con documentaciones procedentes de la lexicografía no académica (Muñoz Armijo, 2007; Huertas, en prensa b y c)—. El mejor ejemplo de ello lo proporcionan los 1784 registros procedentes de la tesis doctoral de Laura Muñoz Armijo (2010) que se corresponden con las primeras documentaciones sobre los derivados en *-ismo* e *-ista* en la historia del *DRAE*. Asimismo, el *Banco de documentaciones* también ofrece las primeras documentaciones académicas de voces derivadas en *-ado* (Clavería *et al.*, 2013), en *-itis* (Julià, 2012 y en prensa a) y en *-osis* (Julià, en prensa b). En estos momentos, diversos miembros del grupo están desarrollando investigaciones sobre otros elementos derivativos y compositivos cuya información documental relativa a la lexicografía académica se irá incorporando al *PLH* progresivamente (*-ato*, *-áceo*, *-algia*, *-ense*, *-és/-esa*, *-ivo*, *-logo*, *-logía*, *-ina*, *-ístico*, *-oma*, *-triz/-triz*, *-udo*, *-ido*).

Estos registros complementan la información que aparece en el *DECH* de distinta forma. El *PLH* puede incluir voces que no se encuentran en el diccionario etimológico, como es el caso de *abandonismo*, *artrosis*, *celulitis* y *conformatriz*. Asimismo, contiene dataciones complementarias al *DECH* tanto en aquellos casos en los que las indicaciones son inexactas como en los casos en los que el *DCEH* no proporciona información documental alguna, algo que es muy frecuente en las voces incluidas en los apartados de derivados y compuestos; así, por ejemplo, proporciona información documental de voces como *alarmista*, *faringitis* y *carbuncosis*. Puede incluir documentaciones anteriores a las del *DECH*, como sucede en el caso de *matriz*, cuya primera documentación del *DECH* pertenece 1604 (Florencia, Fr. Jerónimo) y en el *PLH*, según la investigación de Huertas (en prensa a y b), se incluye la primera documentación procedente del *CORDE* que pertenece a 1437 (Fernández de Madrigal, Alonso de (El Tostado)).

Además del *DRAE*, el *PLH* también recoge primeras documentaciones de otras fuentes lexicográficas de innegable interés histórico para el estudio del léxico español. Así ocurre con la inclusión de las documentaciones del *Nuevo diccionario francés-español* de A. de Capmany (1805) como reflejo de las investigaciones sobre este diccionario llevadas a cabo por Clavería, Freixas y Torruella (2010) y Clavería, Julià y Torruella (2013). El *PLH* cuenta actualmente con más de 350 registros procedentes de esta fuente lexicográfica que se caracterizan por no estar incluidas en el *DECH* (*acefalía*) o por presentar una fecha de documentación anterior a la que recoge este diccionario. *Afonía*, por ejemplo, cuya primera documentación en el *DECH* es de 1882 (Pedro Antonio de Alarcón), ya aparece en la obra de Capmany.

(b) Documentaciones procedentes de estudios sobre el *DECH*

El *Banco de documentaciones* también incorpora algunos de los datos procedentes de las investigaciones que constituyen adiciones documentales al *DECH*. Se trata de trabajos recogidos actualmente en los bancos de datos de bibliografía y de léxico. En el *Banco de bibliografía*, el usuario puede encontrar marcados con la etiqueta «+ DECH» todos los registros de este tipo. Alrededor de 115 registros aparecen distinguidos con esta marca.

Las informaciones procedentes de los artículos que se han considerado adiciones al *DECH* suelen aportar nuevas dataciones o informaciones documentales que el diccionario etimológico no menciona; así pues, cuando el *PLH* recoge estos datos permite poner en contraste, únicamente haciendo un clic, los datos del *DECH* con los de investigadores dedicados al estudio del diccionario. Véase el caso de *acial*, voz para la que el *DECH* proporciona la fecha de 1564 sin más aclaración, mientras que el trabajo de Esther Hernández (1996) recoge una documentación algo anterior del vocabulario de Fray Alonso de Molina, según se muestra en la imagen 2:



Imagen 2. Visualización del registro de la voz *acial* en el Banco de documentaciones

4. INFORMACIONES DOCUMENTALES DEL CATALÁN: EL *ESSAI DE DICTIONNAIRE HISTORIQUE DE LA LANGUE CATALANE* DE JULIÀ-BERNAT ALART

La ampliación del *Banco de documentaciones* del catalán se desarrolla con la incorporación del *Essai de dictionnaire historique de la langue catalane* de Julià-Bernat Alart, un material inédito de gran interés para la filología románica. A continuación, se describen los aspectos esenciales de la obra y su proceso de informatización.

4.1. Julià-Bernat Alart

Julià-Bernat Alart nació en 1824 en la población de Vinçà, en la comarca catalana septentrional del Conflent, que pertenece, desde el punto de vista administrativo, al departamento francés de los Pirineos Orientales, y murió en esta misma población en 1880, a los 56 años de edad. Alart, archivero de profesión, desarrolló a lo largo de su vida una ingente tarea en el ámbito de la historia y la filología, que pudo ver la luz solo parcialmente. Rasico (2002: 375) considera que Alart es «el erudito rosellonés más ilustre del siglo XIX desde la perspectiva del valor científico de su obra así como de la envergadura de esta, tanto la publicada como la inédita».

ditas».⁵ Entre la obra inédita e inacabada de Alart figuran los materiales para un diccionario histórico de la lengua catalana, que su discípulo Pere Vidal tituló *Essai de dictionnaire historique de la langue catalane* (Vidal, 1911: 142), y que cabe entender dentro de la corriente de revalorización de esta lengua que supuso el movimiento cultural de la *Renaixença*, que también tuvo acogida entre los eruditos roselloneses (Gallart, 2006). Para dimensionar la figura de Alart en su contexto y conocer más detalles de su vida y su obra, remitimos al estudio de Catafau y Escudero (2002).

4.2. El *Essai de dictionnaire historique de la langue catalane*

El *Essai de dictionnaire historique de la langue catalane*, que Coromines en el *DECat* cita como *Inventari de la Llengua Catalana* y que actualmente es conocido también con la denominación más breve de *Inventari Alart*, es una colección de 14.639 cédulas lexicográficas manuscritas de puño y letra de Alart, cada una de ellas destinada a recoger documentaciones antiguas de un lema de la lengua catalana. Estas fichas fueron organizadas de forma póstuma por su hermano Pere Alart, quien las ordenó alfabéticamente (aunque no de forma estricta) y las fijó sobre hojas de 326 x 224 mm, hoy reunidas en catorce volúmenes conservados en la *Médiathèque de Perpignan* bajo la signatura Ms 116. Se calcula que el número total de documentaciones contenidas en este elenco es del orden de 170.000 (Melchor y Casals, 2003: 175).

Hasta el día de hoy, el *Essai* ha permanecido inédito, a pesar de que diversos lingüistas se han lamentado por ello. Así, Rico y Solà (1995: 146), en su *Gramàtica i lexicografia catalanes*, califican el *Essai* como un elenco «muy valioso que algún día se tendrá que editar» y Colón y Sobe-ranas (1986: 210), en su *Panorama de la lexicografia catalana*, expresan que «el hecho que el diccionario histórico de Alart haya permanecido inédito es desolador». Esta circunstancia es la causa por la cual hasta la fecha estos materiales han sido escasamente aprovechados, aunque en estudios léxicos puntuales se ha visto su utilidad para localizar primeras documentaciones, ampliar el alcance dialectal de variantes o deshacer hápax (Massanell, 2002: 386-387).

Una muestra de la escasa repercusión que ha tenido hasta la actualidad el valiosísimo *Essai*, al haber quedado inconcluso e inédito, es el aprovechamiento que hicieron de él los dos grandes diccionarios históricos de la lengua catalana publicados durante el siglo XX: el *Diccionari ca-*

⁵ Nota de los autores: La traducción de todas las citaciones de obras catalanas es nuestra.

talà-valencià-balear (DCVB), de Antoni M. Alcover y Francesc de B. Moll, integrado por diez volúmenes aparecidos entre 1930 y 1968, y el *Diccionari etimològic i complementari de la llengua catalana* (DECat), de Joan Coromines, compuesto a su vez por nueve volúmenes publicados entre 1980 y 1991, a los cuales más tardíamente se sumaría el décimo, de índices.

Según concluyen Casals y Massanell (2003), el aprovechamiento que hace el DCVB del *Essai* de Alart es ínfimo y por vía indirecta, ya que no va más allá de las 77 fichas de la letra A que el año 1911 un discípulo de Alart a quien ya hemos mencionado, Vidal, publicó, previamente reelaboradas, en el primer volumen de la revista *Ruscino. Revue d'Histoire et d'Archéologie du Roussillon et des autres pays catalans* (1911-1913), movido por la «Letra de convit» con que Alcover exponía el proyecto de elaboración de su diccionario y convocaba a todos a participar en él. El DCVB recoge una tercera parte de las documentaciones procedentes de estas cédulas, que frecuentemente somete a un proceso de reducción del contexto y de unificación del sistema de referencia, que pierde precisión («doc[umento]» sustituye a «declaración testimonial», «inventario de bienes» o «ordinaciones» y «ros[ellonés]» reemplaza a Elna, Jóc o Sant Llorenç de la Salanca). Cabe señalar que las 21 documentaciones escogidas constituyen aportaciones relevantes: son la única atestiguación del lema o de la acepción, suponen la datación más antigua de la palabra o aportan variantes formales nuevas.

Es distinto el caso del *DECat*. Coromines consultó directamente el cedulario inédito de Alart para incorporar datos a su magna obra lexicográfica, bajo la indicación *ImLC*. Ya lo avanzó Gulsoy (1979) con anterioridad al inicio de la publicación del *DECat*, en su estudio sobre la obra filológica de Alart, que sucedía, después de un largo lapso temporal, a la biobibliografía con que Vidal (1896) había honrado la memoria de su predecesor. Y también dejó constancia de ello el mismo Coromines en el prefacio del *DECat*, donde además pueden leerse grandes elogios tanto hacia la persona de Alart, a quien llama «el gran archivero, erudito y filólogo rosellonés», como hacia su obra, que considera «sometida al control de una ciencia cartística, paleográfica y lingüística de perfecto rigor, que encontraría bien pocos iguales en la erudición europea de su tiempo» (*DECat*: «Prefaci», x). Coromines declara a continuación que dedicó bastantes meses a hacer copiosos extractos del *Essai*, «cuando la preparación del libro estaba ya muy adelantada, y sabía yo bien los puntos en los que serían útiles sus datos para la documentación histórica de nuestro libro». Colón y Soberanas (1986: 210) postulan que, a pesar de esta afirmación, la parte recogida por Coromines no debió de ser muy grande y

revelan la omisión de diversas documentaciones tempranas. Como hacen notar Casals y Massanell (2002: 65), la declaración de Coromines implica que no hubo un vaciado sistemático del *Essai*, sino que la obra de Alart fue usada para cubrir lagunas en la documentación histórica de ciertas entradas del *DECat*, lo cual explica que una breve cata permita descubrir materiales no aprovechados e incluso que, en las entradas con la referencia *InvLC*, pueden no reproducirse las dataciones que la han motivado, o bien haberse seleccionado citas que no son las más antiguas o las más adecuadas entre las que proporciona Alart.

4.2.1. Macroestructura

El *Essai* consta de 14.369 cédulas lexicográficas manuscritas y cada una de ellas está dedicada a documentar una palabra de la lengua catalana, incluidos topónimos y antropónimos. Estas fichas están ordenadas alfabéticamente y se observa la ausencia de las letras K, W, Y, Z. En conjunto el *Essai* recoge unas 170.000 documentaciones procedentes mayoritariamente de fondos de archivos, pero también de textos literarios y de obras lexicográficas.

4.2.2. Microestructura

Cada cédula lexicográfica elaborada por Alart se estructura en los siguientes apartados:

1. Lema, situado normalmente en la parte superior central.
2. Variantes gráficas, fonéticas o morfológicas, o bien formas del paradigma, solo en algunas ocasiones y repartidas a ambos lados del lema; su disposición puede no dejar claro qué forma es la escogida como lema.
3. Traducción al francés, únicamente para algunos lemas, ubicada sistemáticamente en la esquina superior derecha y a veces acompañada de un interrogante de duda; muy raramente puede ser sustituida por una paráfrasis similar a una definición, o incluso por la etimología.
4. Documentaciones, no muy extensas, en número variable según la cédula: desde una sola hasta varias decenas.
5. Fecha de cada documentación, sea precediéndola o siguiéndola, que revela que éstas frecuentemente se presentan en orden cronológico, de la más antigua a la más reciente, aunque no se cumple de forma estricta.

6. Referencia de la fuente de procedencia de cada datación, al final de ésta, presentada casi siempre de forma muy abreviada, y que las más de las veces resulta críptica para el investigador actual que se acerca al *Essai*, por lo menos de momento.

Olera
bordure
bord
 1. cobutoir de diap d'or rialat per les
 horeres 1370.2.208
 1393. que lo dit ayxant vagon
 horera del val, fins lo lork hon.
 R. 146. en longear la fosse -
 1400. e age a handie les aygues
 qui van ~~en~~ per la dita agulla
 e aquelles mesme entorn la
 sua vinya es la olera qui
 afronta ab Na Morra. S. 350.

Documentacions O	
Id	2054
Lema	orera
Citació	I- cobertor de drap d'or rialat per les horeres
Any	1370
Observacions	208] 308, corr.
Codi	1370-Q-208
Font	

Documentacions O	
Id	2055
Lema	orera
Citació	que fo dit eyxauch vaga horera del val, fins lo loch hon
Any	1393
Observacions	Aclariment en francès després del codi: "en longeant le fossé".
Codi	R-144 en longeant le fossé
Font	

Documentacions O	
Id	2056
Lema	orera
Citació	e aga a pendre les aygues qui venran per la dita agulla e aquelles menar entorn la sua vinya per la orera qui afronta ab Na Morera
Any	1400
Observacions	venran] vendran, corr.
Codi	S-350
Font	

Imagen 3. Reproducción de la cédula lexicográfica del *Essai* de Alart correspondiente al lema *orera* 'borde, orilla' y registros de la base de datos de documentaciones

Véase como ejemplo la reproducción de la cédula del *Essai* de Alart correspondiente al lema *orera* ‘borde, orilla’ (Imagen 3). Esta ficha lexicográfica contiene tres dataciones de la palabra, con las variantes *orera* y *borera*, de 1370, 1393 y 1400, todas de origen documental, las cuales han sido introducidas en la base de datos de documentaciones (Imagen 3) y aportan información relevante respecto del *DCVB* (que carece de la entrada *orera*) y del *DECat* (cuya primera documentación para este término es de 1458 y de un origen geográfico muy distante, puesto que procede de Castelló), según establecen Cucurell y Massanell (en preparación).

4.2.3. Fuentes

Las fuentes del *Essai* son de tres tipos: obras lexicográficas, obras literarias y fondos de archivos documentales.

Las fuentes lexicográficas usadas en el *Essai* parece ser que son dos. Por un lado, la adaptación al catalán que el año 1507 Busa hizo del *Vocabularium* de Nebrija (Gabriel Busa, *Diccionario latín-catalán y catalán-latín*, Barcelona, 1507), y por el otro, el *Lexique roman* de Raynouard (François Raynouard, *Lexique roman ou Dictionnaire de la langue des troubadours, comparée avec les autres langues de l'Europe latine*, 6 vol., París, 1838-1844).

Las obras literarias de las cuales Alart obtuvo material para el *Essai* abarcan del siglo XIII al XIX (hay citas, por ejemplo, del *Cançoner* de Milà i Fontanals, de 1853), aunque la inmensa mayoría son medievales. Entre las citadas se encuentran, por poner unos ejemplos: obras en verso y en prosa de Ramon Llull; las cuatro grandes crónicas historiográficas, de Jaume I, Desclot, Muntaner y Pere el Cerimoniós; creaciones de poetas valencianos como Fenollar, Escrivà y Vilaespina, procedentes del primer libro impreso en catalán, las *Trobes en labor de la Verge Maria*, de 1474; y textos literarios e históricos editados por Prósper de Bofarull, Antoni de Bofarull y Antoni Bergnes de las Casas.

Las fuentes archivísticas del *Essai* provienen principalmente de la recopilación conocida como *Cartulaire rousillonais* (llamado también *Cartulari Alart*), conservado asimismo en la *Médiathèque de Perpignan*. Nuevamente se trata de una obra vastísima, compuesta por 64 volúmenes (más tres de índices), escrita de puño y letra de Alart, que permanece inédita. En ella se reproducen, íntegramente o de forma parcial, innumerables documentos originales custodiados en los *Archives Départementales des Pyrénées-Orientales*, que es un depósito documental formado por el conjunto de archivos históricos y administrativos del Rosselló y la Cerdanya, que fueron reunidos a partir de la Revolución francesa. Están integrados por

diversos archivos municipales, notariales, hospitalarios y por fondos de obispados, prioratos y abadías de las comarcas catalanas septentrionales.

4.2.4. Delimitación diacrónica, diatópica y diastrática

En el *Essai* se recogen documentaciones de fondos archivísticos fechados entre el siglo IX, en los orígenes mismos de la lengua catalana, y el siglo XVII, y más concretamente, el año 1660, fecha límite que no resulta casual (recordemos que el 7 de noviembre de 1659 se firmó el Tratado de los Pirineos, por el cual eran cedidos a Francia el condado del Rosselló y una parte del condado de la Cerdanya). A pesar de esta amplitud cronológica, unas tres cuartas partes de los documentos son de los siglos XIII y XIV. Por lo que concierne a las obras literarias, estas siguen criterios cronológicos distintos y abarcan del siglo XIII al XIX, aunque la mayoría son obras medievales consultadas a partir de ediciones de la época de la *Reinaixença*, el punto de partida de la cual se considera que es la publicación, en 1833, de la oda *La Pàtria* de Bonaventura Carles Aribau.

Por lo que respecta al origen geográfico de las documentaciones, se sabe que proceden del dialecto septentrional del catalán, aunque el sistema de referencias de Alart dificulta muchas veces la identificación de la localidad concreta. A partir del estudio de Casals y Massanell (2003) sobre la reelaboración que Vidal hizo de un puñado de cédulas de la letra A (las que luego se incorporarían parcialmente al *DCVB*), se puede tener una muestra de las localidades de procedencia de los documentos, ya que Vidal en ocasiones lo esclarece: Bajoles, Cornellà del Bèrcol, Cotlliure, Elna, Illa, Jóc, Molig, Montesquiu d'Albera, Nefiac, Palau del Vidre, Perpinyà, Sant Llorenç de la Salanca, Sant Nazari, Santa Maria la Mar y el mismo Vinçà, de donde Alart era hijo.

La reelaboración de Vidal también permite hacerse una idea del tipo de documentos que proporcionaron datos para el *Essai*, ya que el discípulo de Alart esclarece también este punto oculto bajo los códigos de citación y hace aparecer ante los ojos del lector declaraciones testimoniales, inventarios de bienes, ordinationes, cabreos, cartas y minutas notariales.

4.3. Informatización del *Essai*

Rasico (2002: 380) lamenta que los materiales inéditos de Alart conservados en la biblioteca municipal de Perpiñán «no parecen haber llamado la atención de muchos investigadores hasta hoy». En realidad, como exponen Melchor y Casals (2003), en el seno del Departamento de

Filología Catalana de la Universidad Autónoma de Barcelona se constituyó un proyecto de investigación con la vista fijada en estos materiales, gracias al cual fue posible conseguir, no sin arduas negociaciones, una reproducción completa de los copiosos materiales del *Essai*, y transcribir una décima parte (17.068 registros) de las documentaciones en él contenidas.

Finalizado el proyecto, se formó en el seno del mismo departamento el Grupo de Investigación de Historia de la Lengua Catalana de la Época Contemporánea, que custodia los materiales y ha podido hacer avanzar su transcripción con la incorporación de una estudiante de licenciatura mediante el programa de prácticas de la titulación. Su dedicación ha hecho posible que estén transcritos todos los lemas con la excepción de los que corresponden a la letra C, lo que supone un total de 13.620 lemas introducidos en una base de datos.

Pero el impulso al progreso del trabajo sobre el *Essai* ha venido con su incorporación al *PLH*. Esto ha hecho posible la transcripción e incorporación a una base de datos informatizada de 37.399 documentaciones, provenientes de 5.664 lemas, que son más de un tercio de los que contiene el *Essai*, y que han dado pie a retomar los estudios sobre esta obra (Cucurell y Massanell, en preparación).

Obviamente, todos estos materiales, y los que vayan incorporándose, deberán ser sometidos a la revisión de las transcripciones, además de pasar por un proceso de sistematización de la información y de clarificación de las fuentes. La tarea pendiente es mucha, pero el volumen de labor ya realizada, juntamente con la inscripción de una tesis doctoral que toma estos materiales como corpus de estudio, hacen ser optimistas sobre las posibilidades de conclusión de la edición informatizada del *Essai*, largamente esperada por la comunidad científica por los muchos beneficios que puede reportar para el estudio histórico de la lengua catalana.

Referencias bibliográficas

- ÁLVAREZ DE MIRANDA, Pedro (2005): «Estar de vuelta sin haber ido. Sobre la situación de los estudios léxicos en la lingüística histórica española», *La Corónica*, 34.1, 131-135.
- (2006): «Problemas y estado actual de los estudios sobre historia del léxico español», en José Jesús de Bustos y José Luis Girón (eds.), *Actas del VI Congreso Internacional de Historia de la Lengua Española*, Arco Libros, Madrid, vol. 2: 1229-1240.

- BLECUA PERDICES, José M. y Gloria CLAVERÍA NADAL (1999): «La lexicografía castellana, antes y después de Coromines», en Joan Solà (ed.), *L'obra de Joan Coromines*, Fundació La Caixa, Sabadell, pp. 29-43.
- CAPMANY, Antonio de (1805): Nuevo diccionario francés-español. En este van enmendados, corregidos, mejorados, y enriquecidos considerablemente los de Gattel y Cormon, Madrid, Imprenta de Sancha.
- CASALS, Daniel y Mar MASSANELL (2002): «El *Essai de dictionnaire historique de la langue catalane* como precedente de dos obras lexicográficas catalanas: el *DCVB* de Alcover y Moll y el *DECLC* de Coromines», en Mar Campos y José Ignacio Pérez (ed.), *De historia de la lexicografía*, Toxosoutos, La Coruña, pp. 51-66.
- (2003): «Aprofitament de l'*Essai de dictionnaire historique de la langue catalane*, de Julià-Bernat Alart, en el *Diccionari català-valencià-balear*, d'Alcover i Moll», en Jaume Guiscafrè y Antoni Picornell (ed.), *Actes del Congrés Internacional Antoni M. Alcover*, Publicacions de l'Abadia de Montserrat, Barcelona, pp. 239-260.
- CATAFAU, Aymat y Jean-Paul ESCUDERO (2002): «Julià Bernat Alart (1824-1880), un català de França: de la història del Rosselló i Cerdanya a la filologia catalana», *Revue d'Études Catalanes*, n° 5, pp. 15-25.
- CLAVERÍA, Gloria (1999): «La documentación en el diccionario etimológico», en José Manuel Blecua, Gloria Clavería, Carlos Sánchez y Joan Torruella (eds.), *Filología e informàtica. Nuevas tecnologías en los estudios filológicos*, Lleida, Editorial Milenio / Universitat Autònoma de Barcelona, pp. 259-283.
- (2001): «El léxico especializado en la lexicografía de finales del siglo XIX: la decimotercera edición (1899) del *Diccionario de la Lengua Castellana de la Academia*», en Jenny Brumme (ed.), *La historia de los lenguajes iberorrománicos de especialidad: la divulgación de la ciencia*, UPF-Vervuert-Iberoamericana, Barcelona-Frankfurt-Madrid, pp. 207-222.
- (2004): «La Real Academia Española a finales del siglo XIX: notas sobre las *voces americanas* en la decimotercera edición del *Diccionario de la Lengua Castellana*», en Milka Villayandre (ed.): *Actas del V Congreso de Lingüística General (León, 5-8 de marzo de 2002)*, Arco/Libros, Madrid, vol. I, pp. 621-633.
- (2005): «La Real Academia Española a finales del siglo XIX: el *Diccionario de la lengua castellana* de 1899 (13.^a edición)», *Boletín de la Real Academia Española*, LXXXIII, pp. 255-336.
- (2007): «Historia del léxico en los diccionarios: la deuda del *Diccionario de la Lengua Castellana* de la Real Academia Española con los diccionarios de M. Núñez de Taboada», *Revista de Historia de la Lengua Española*, 2, pp. 3-27.

- (2008): «Joan Coromines i la filologia espanyola: el *Diccionario Crítico etimológico*», en Teresa Cabré, Marta Prat y Joan Torruella (eds.): *Joan Coromines i la filologia románica*, Servei de Publicacions de la Universitat de Barcelona, Barcelona, pp. 129-143.
- (2012): «Nuevas perspectivas en torno al estudio de la evolución del léxico», en Gloria Clavería, Margarita Freixas, Marta Prat y Joan Torruella (eds.): *Historia del léxico: perspectivas de investigación*, Iberoamericana/Vervuert, Madrid/Frankfurt, pp. 13-90.
- (en prensa): El neologismo en la lexicografía académica del siglo XIX.
- CLAVERÍA, Gloria y Cristina BUENAFUENTES (eds.). (2005): *Germà Colon: les llengües romàniques juntes i contrastades*, Universitat Autònoma de Barcelona, Bellaterra.
- CLAVERÍA, Gloria y Margarita FREIXAS (en prensa): «La quinta edición del *Diccionario de la lengua castellana* de la Real Academia Española (1817): el aumento de voces», en *Actas del IX Congreso Internacional de Historia de la Lengua Española* (Cádiz, Universidad de Cádiz, 10-14 de septiembre de 2012).
- CLAVERÍA, Gloria y Carmen MORALES (1999-2000): «El DCECH y los diccionarios de la Academia», *Revista de lexicografía*, VI, pp. 37-55.
- CLAVERÍA, Gloria y Joan TORRUELLA (2005): «Base de datos para un corpus de documentaciones léxicas», en Claus D. Pusch, Johannes Kabatek und Wolfgang Raible (eds.), *Romanistische Korpuslinguistik II/Romance Corpus Linguistics II. Corpus und diachrone Sprachwissenschaft/Corpora and Diachronic Linguistics*, Gunter Narr, Tübingen, pp. 215-228.
- (2012): «El Corpus informatizado del catalán antiguo (CICA) y su herramienta de explotación (Estación de análisis documentales, EAD)», en María Jesús Torrens y Pedro Sánchez-Prieto (eds.), *Nuevas perspectivas para la edición y el estudio de documentos hispánicos antiguos*, Peter Lang, Bern, pp. 405-425.
- CLAVERÍA, Gloria, Margarita FREIXAS, Marta PRAT y Joan TORRUELLA (2012): (eds.): *Historia del léxico: perspectivas de investigación*, Iberoamericana / Vervuert, Madrid / Frankfurt.
- CLAVERÍA, Gloria, Margarita FREIXAS y Joan TORRUELLA (2010): «La traducción de términos científico-técnicos del francés al español en el *Nuevo diccionario francés-español* (1805) de Antonio de Capmany», *Cuadernos del Instituto de Historia de la Lengua*, 4, pp. 27-53.
- CLAVERÍA, Gloria, Carolina JULIÀ y Joan TORRUELLA (2012): «El léxico científico en la lexicografía de principios de siglo XIX: el *Diccionario de la lengua Castellana* de la RAE (4.^a ed., 1803) y el *Nuevo diccionario francés-español* de A. de Capmany (1805)», en Graça RioTorto (ed.), *Léxico de la Ciencia: tradición y modernidad*, Lincom Europa, München, pp. 97-122.

- CLAVERÍA, Gloria, Laura MUÑOZ, Marta PRAT y Joan TORRUELLA (2005): *Índex del lèxic estudiat en l'obra de Germà Colòn* [CD-ROM], en G. Clavería y C. Buenafuentes (eds.) (2005).
- CLAVERÍA, Gloria, Cristina BUENAFUENTES, Margarita FREIXAS, Carolina JULIÀ, Mar MASSANELL, Laura MUÑOZ, Marta PRAT, Joan TORRUELLA y Sonia VARELA (2008): «Portal de léxico hispánico: una herramienta para el estudio del léxico», en Elisenda Bernal y Janet de Cesaris Ward (eds.), *Proceedings of the XIII Euralex International Congress (15-19 July 2008)*, Institut Universitari de Lingüística Aplicada/Universitat Pompeu Fabra, Barcelona, pp. 655-662 [Edición en CD-ROM].
- CLAVERÍA, Gloria, Cristina BUENAFUENTES, Margarita FREIXAS, Carolina JULIÀ, Laura MUÑOZ, Marta PRAT y Joan TORRUELLA (2013): «*Morfología derivativa especializada: el sufijo -ado y su recepción en el DRAE*», en Gloria Clavería, Cecilio Garriga, Carolina Julià, Francesc Rodríguez y Joan Torruella (eds.): *Historia, lengua y ciencia: una red de relaciones*, Peter Lang, Frankfurt am Main, pp. 51-72.
- COLÓN, Germà y Amadeu-J. SOBERANAS (1986): *Panorama de la lexicografía catalana. De les gloses medievals a Pompeu Fabra*, Enciclopèdia Catalana, Barcelona.
- COROMINES, Joan (1980): «Prefaci», en Joan Coromines, *Diccionari etimològic i complementari de la llengua catalana*, Curial, Barcelona, vol. I, pp. ix-xii.
- CUCURELL, Ares y Mar MASSANELL (en preparación): «Nous lemes i primeres documentacions en l'*Inventari Alaró*».
- DECH = Corominas, Joan y José A. Pascual (2012 [1980-1991]): *Diccionario crítico etimológico castellano e hispánico*, Gredos, Madrid, [versión en CD-ROM].
- DUGGAN, Joseph J. y Charles B. Faulhaber (1988): *Yakov Malkiel: A Tentative Autobiography*, University of California Press Berkeley, (*Romance Philology*, 42, special issue).
- DWORKIN, Steven N. (1994): «Progress in Medieval Spanish Lexicography», *Romance Philology*, XLVII, 4, pp. 406-425.
- FERRER, Joan (2008): «Les múltiples cares de les piràmides», en Teresa Cabré, Marta Prat y Joan Torruella (eds.), *Joan Coromines i la filologia romànica*, Universitat Autònoma, Bellaterra, pp. 99-109.
- FREIXAS, Margarita (2012): «El método lexicográfico de la Real Academia Española entre los años 1732-1770: las *Plantas* de la segunda edición del *Diccionario de Autoridades* (tomo I, A-B)», en Antoni Nomdedeu, Esther Forgas y Maria Bargalló, *Avances de lexicografía hispánica*, Universitat Rovira i Virgili, Tarragona, vol. I, pp. 301-315.

- GALLART, Josep (2006): «Els erudits rossellonesos del segle XIX i Jacint Verdaguer. Codis disponibles: estudis, gramàtiques i diccionaris», *Anuari Verdaguer*, n° 14, pp. 291-309.
- GULSOY, Joseph (1979): «L'obra filològica de Julià Bernat Alart», en *Miscel·lània Aramon i Serra*, Curial, Barcelona, vol. I, pp. 243-253.
- GARCÉS, M.^a Pilar (ed.) (2008): *Diccionario histórico: nuevas perspectivas lingüísticas*, Iberoamericana-Vervuert, Madrid-Frankfurt.
- GÓMEZ, Marta y José Ramón CARRIAZO (2010): *La marcación en lexicografía histórica*, Cilengua, San Millán de la Cogolla.
- HERNÁNDEZ HERNÁNDEZ, Esther (1996): «Palabras del siglo XVI», *Revista de Filología Española*, 76, pp. 171-175.
- HUERTAS, Sheila (en prensa a): «Suplementos del DRAE (1817-1832): la incorporación de voces de la historia natural», en *Actas del V Congreso Internacional de Lexicografía Hispánica* (Universidad Carlos III, 25-27 de junio de 2012).
- (en prensa b): «El sufijo *-triz* / *-triz* en la lexicografía preacadémica», en *Actas del XIII Congreso Internacional de la AIIHLE* (Universidad de Salamanca, del 10- 12 de abril de 2013).
- (en prensa c): «Los derivados agentivos en *-triz* / *-triz* en la lexicografía preacadémica», en *Actas del IX Encuentro Internacional de Morfólogos* (Universidad de Cádiz 16-18 de mayo de 2013). [Póster]
- JULIÀ, Carolina (2008): «El léxico de la metrología en la lexicografía académica de los siglos XVIII y XIX: las unidades de capacidad tradicionales», en Dolores Azorín (dir.): *El diccionario como puente entre las culturas y lenguas del mundo. Actas del II Congreso Internacional de Lexicografía Hispánica (Alicante, 19-23 de septiembre de 2006)*, Biblioteca Virtual Miguel de Cervantes, Alicante, pp. 706-714. [Edición en CD-ROM y en línea: <<http://www.cervantesvirtual.com/FichaObra.html?Ref=30799>>].
- (2012): «La recepción del léxico científico en la lexicografía académica: las voces derivadas en *-itis*», *Revista de Lexicografía*, 18, pp. 77-102.
- (en prensa a): «Los patrones derivativos del sufijo *-itis* en español moderno: análisis de algunos ejemplos», en Elisenda Bernal (ed.), *Los afijos: variación, rivalidad y representación. VIII Encuentro de Morfólogos* (Universitat Pompeu Fabra, 10-12 de mayo de 2012).
- (en prensa b): «La recepción de los derivados en *-osis* en la lexicografía académica», en *Actas del V Congreso Internacional de Lexicografía Hispánica* (Universidad Carlos III, 25-27 de junio de 2012).
- MASSANELL, Mar (2002): «El mot valencià i nord-occidental *picaport*: aportació a la divisió català oriental / català occidental», en Emili Casanova, Joaquim

- Martí y Abelard Saragossà (ed.), *Estudis del valencià d'ara*, Denes, València, pp. 375-402.
- MELCHOR, Vicent de y Daniel CASALS (2003): «InvAlart: presentació del projecte d'edició de l'Essai de Dictionnaire Historique de la Langue Catalane o Inventari Alart», en *Actes del Dotzè Col·loqui Internacional de Llengua i Literatura Catalanes*, vol. III, Publicacions de l'Abadia de Montserrat, Barcelona, pp. 173-188.
- MUÑOZ ARMIJO, Laura (2007): «La recepción de los derivados en *-ismo* e *-ista* en la lexicografía española no académica de la primera mitad del siglo XIX», *Revista de Lexicografía*, XIII, pp. 75-104.
- (2008): «La terminología lingüística en la historia del DRAE: los derivados en *-ismo* e *-ista*», en Dolores Azorín Fernández (dir.): *El diccionario como puente entre las lenguas y culturas del mundo. Actas del II Congreso Internacional de lexicografía hispánica (Alicante, 19-23 de septiembre de 2006)*, Biblioteca Virtual Miguel de Cervantes, Alicante, pp. 753-763 [Edición en CD-ROM y en línea
<<http://www.cervantesvirtual.com/obra/la-terminologa-lingstica-en-la-historia-del-drae--los-derivados-en-ismo-e-ista-0/>>].
- (2010 a): «Los derivados en *-ismo* e *-ista* en las ediciones del DRAE de la segunda mitad del siglo XX», en: *Ars longa. Diez años de AJIHLE (Asociación de Jóvenes Investigadores de Historiografía e Historia de la Lengua Española)*, Voces del Sur, Buenos Aires, pp. 335- 354.
- (2010 b): *La historia de los derivados en -ismo e -ista en el español moderno*, Universitat Autònoma de Barcelona, Bellaterra, en línea:
<<https://www.educacion.gob.es/teseo/imprimirFicheroTesis.do?fichero=16528>>.
- (2011): «Aspectos de la recepción del léxico en *-ismo* e *-ista* en la lexicografía académica española y en la hispanoamericana: las actitudes en los siglos XVIII, XIX y XX», *Cuadernos del Instituto Historia de la Lengua – Cilengua*, vol. 6, pp. 44-72.
- (2012a): «La ideología en la tradición lexicográfica académica: el esquema morfológico de los nomina actionis, nomina qualitatis y nomina agentis», *Mediazioni* 12, <<http://mediazioni.sitlec.unibo.it>>, [Disponible en:
<http://www.mediazioni.sitlec.unibo.it/images/stories/PDF_folder/document-pdf/2012/muoz%20armijo.pdf>].
- (2012b): *La historia de los sufijos -ismo e -ista: evolución morfológica y semántica en la tradición lexicográfica académica española*, Cilengua, San Millán de la Cogolla.

- (2012c): «Procesos semánticos en la formación de creaciones en *-ismo* e *-ista* del español moderno», en Antoni Nomdedeu *et al.* (eds.), *Avances de lexicografía hispánica*, vol. I, Universitat Rovira i Virgili, Tarragona, pp. 427-440.
- (2013): «Patrones lexicogenésicos de los sufijos *-ismo* e *-ista* en su transmisión a la lengua española», en Isabel Pujol Payet (ed.), *Formación de palabras y diacronía*, *Anexos Revista de Lexicografía*, Servicio de Publicaciones Universidade da Coruña, A Coruña.
- (en prensa a): «La representación lexicográfica de los sufijos de agente en los diccionarios españoles: propuestas de mejora», en Elisenda Bernal (ed.), *Los afijos: variación, rivalidad y representación. Actas del VIII Encuentro de Morfólogos Españoles, Universitat Pompeu Fabra, 10-12 de maig de 2012*.
- (en prensa b): «La difusión de la nomenclatura química en la lexicografía académica española: el sistema derivativo para la formulación de sales y compuestos», en *Actas del V Congreso Internacional de Lexicografía Hispánica (Universidad Carlos III, 25-27 de junio de 2012)*.
- (en prensa c): «Morfología léxica del sufijo *-ina*. análisis de los derivados en la lexicografía académica española del siglo XVIII», en *Hispanic Studies in the 21st century*, University of Belgrade Press.
- NITLLE = Real Academia Española (2001): *Nuevo Tesoro Lexicográfico de la Lengua Española*, Madrid, Espasa Calpe [edición en DVD].
- PASCUAL, José A. y Rafael GARCÍA PÉREZ (2007): *Límites y horizontes en un diccionario histórico*, Diputación de Salamanca, Salamanca.
- PELLEN, René (2003): «Diacronía y descripción del cambio lingüístico», *La Corónica*, 31.2, 73-82.
- PLH = *Portal de léxico hispánico*: <<http://www.portaldelexico.es>>.
- RASICO, Philip D. (2002): «Julià Bernat Alart i la toponímia de la Catalunya Nord», en Emili Casanova y Vicenç M. Rosselló (ed.), *Congrés Internacional de Toponímia i Onomàstica Catalanes*, Denes y Universitat de València, València, pp. 375-386.
- RICO, Albert y Joan SOLÀ (1995): *Gramàtica i lexicografia catalanes. Síntesi històrica*, Universitat de València, València.
- RIVAROLA, José Luis y José PÉREZ (eds.) (2006): *Escritos escogidos de lengua y literatura española*, Gredos, Madrid.
- TORRUELLA, Joan (2006): «La documentación en el DCECH» en Gloria Clavería y M.^a Jesús Mancho (eds.): *Estudio del léxico y bases de datos*, Universitat Autònoma de Barcelona, Bellaterra, pp. 77-104.
- (2009a): «Los ejes principales en el diseño de un corpus diacrónico: el caso del cic», en Pascual Cantos Gómez, Aquilino Sánchez Pérez (eds.),

- A Survey on Corpus-based Research / Panorama de investigaciones basadas en corpus*, Asociación Española de Lingüística del Corpus.
 <<http://www.um.es/lacell/aelinco /contenido/pdf/2.pdf>>.
- (2009b): «Bases científicas en la investigación a partir de corpus: el caso del Corpus Informatitzat del Català Antic» en *Diacronía de las lenguas iberorrománicas: nuevas aportaciones desde la lingüística de corpus*, Iberoamericana / Vervuert, Madrid/Frankfurt, pp. 95-115.
- (2009c): «Computerized Corpus of Old Catalan (CICA)», en The fifth Corpus Linguistics Conference (University of Liverpool, 20-23 de juliol de 2009) [Póster].
- (2012): «Computerised Corpus of Old Catalan», en Guadalupe Aguado de Cea, Mari Carmen Suárez-Figueroa, Raúl García-Castro y Elena Montiel-Ponsoda (eds.), *Proceedings of the 10th Terminology and Knowledge Engineering Conference: New frontiers in the constructive symbiosis of terminology and knowledge engineering*. Madrid, pp. 323-325.
- VIDAL, Pere (1896): «Notice sur la vie et les travaux de Julien-Bernard Alart, ancien archiviste des Pyrénées-Orientales», *Bulletin de la Société Agricole, Scientifique et Littéraire des Pyrénées-Orientales*, n° XXXVII, pp. 87-203.
- (1911): «Contribution au Dictionari de la llengua catalana», *Ruscino*, n° I, pp. 142-150, 298-306, 613-622.

