

EL ESPAÑOL A TRAVÉS DEL TIEMPO
ESTUDIOS OFRECIDOS
A RAFAEL CANO AGUILAR

VOLUMEN I

ARACELI LÓPEZ SERENA
ANTONIO NARBONA JIMÉNEZ
SANTIAGO DEL REY QUESADA
Directores



Sevilla 2016

Colección Lingüística

Núm.: 50

COMITÉ EDITORIAL:

José Beltrán Fortes

(Director de la Editorial Universidad de Sevilla)

Manuel Espejo y Lerdo de Tejada

Juan José Iglesias Rodríguez

Juan Jiménez-Castellanos Ballesteros

Isabel López Calderón

Juan Montero Delgado

Lourdes Munduate Jaca

Jaime Navarro Casas

M^a del Pópulo Pablo-Romero Gil-Delgado

Adoración Rueda Rueda

Rosario Villegas Sánchez

COMITÉ ORGANIZADOR DE ESTA EDICIÓN:

DIRECTORES: Araceli López Serena, Antonio Narbona Jiménez y

Santiago del Rey Quesada

COORDINADORA DE EDICIÓN: Marta Fernández Alcaide

COLABORADORES DE EDICIÓN: Elena Leal Abad, Olga Julián Mariscal,
Elena Carmona Yanes y María Méndez Orense

Reservados todos los derechos. Ni la totalidad ni parte de este libro puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico, incluyendo fotocopia, grabación magnética o cualquier almacenamiento de información y sistema de recuperación, sin permiso escrito de la Editorial Universidad de Sevilla.

Diseño de cubierta: José Antonio Molina Infante

© Editorial Universidad de Sevilla 2016

C/ Porvenir, 27 - 41013 Sevilla.

Tlfs.: 954 487 447; 954 487 451; Fax: 954 487 443

Correo electrónico: eus4@us.es

Web: <<http://www.editorial.us.es>>

© Araceli López Serena, Antonio Narbona Jiménez y
Santiago del Rey Quesada, Directores 2016

© Por los textos, los autores 2016

© José Cano Méndez, retrato de Rafael Cano Aguilar 2016

Impreso en papel ecológico

Impreso en España-Printed in Spain

ISBN. Obra completa: 978-84-472-1834-9

ISBN. Volumen I: 978-84-472-1870-7

ISBN. Volumen II: 978-84-472-1871-4

Depósito Legal: SE 1390-2016

Maquetación: Fernando Fernández. ed-Libros

Impresión: Imprenta Kadmos

ANÁLISIS CUANTITATIVO DE LAS CITAS DE OBRAS EN EL *DICCIONARIO DE AUTORIDADES*

Guillermo Rojo

Universidad de Santiago de Compostela / Real Academia Española

RESUMEN

El análisis de las citas textuales contenidas en el *Diccionario de Autoridades* (DA) ha tropezado tradicionalmente con los problemas esperables por el volumen de la obra, que ha limitado las investigaciones al trabajo con autores concretos, el léxico de un cierto tipo o muestras de diferentes tamaños. La preparación de una versión digitalizada y su puesta a disposición pública en la página electrónica de la Real Academia Española <<http://web.frl.es/DA.html>> facilita ya la automatización de la parte más tediosa de la extracción de los textos y permite centrarse en las tareas de identificación, cuantificación y distribución de las citas. El presente trabajo describe el procedimiento utilizado para obtener las listas de obras y su frecuencia, muestra la relación de las más utilizadas y esboza algunas de las cuestiones fundamentales que se derivan de su análisis.

PALABRAS CLAVE: *Diccionario de Autoridades*, historia de la lexicografía, recuperación de información.

SUMMARY

The analysis of the textual quotations contained in the *Diccionario de Autoridades* (DA) has traditionally encountered the predictable problems given the scale of the work, which has limited research to the analysis of specific authors, a certain type of lexicon, or samples of different sizes. The preparation of a digitized version, and its public availability on the website of the Real Academia Española <<http://web.frl.es/DA.html>>, has facilitated the automation of the most tedious part of the extraction of the texts, and has thus made it possible to focus on the tasks of identification, quantification and distribution of the quotations.

The current study describes the procedure used to obtain lists of works and their frequency, shows the relationship of those which are most used, and outlines some of the key issues arising from their analysis.

KEYWORDS: *Diccionario de Autoridades*, history of lexicography, information retrieval.

1. ANTECEDENTES

Como es bien sabido, lo más destacable, valioso y actual del conocido como *Diccionario de Autoridades* (DA) consiste en la intención deliberada que tuvieron sus redactores de producir un diccionario en el que la mayor parte de las palabras y acepciones incluidas estén autorizadas por textos en los que aparezca precisamente con ese significado. El propósito, inicialmente normativo, de demostración de que el español de comienzos del XVIII constituye una lengua culta, elegante, depurada de incorrecciones y de extranjerismos se manifiesta en la voluntad de tomar como punto de partida un conjunto de autores y textos muy seleccionados, que son los que, inicialmente a propuesta del propio marqués de Villena, figuran en la “lista de los autores elegidos por la Real Academia Española, para el uso de las voces y modos de hablar...”. Pero, afortunadamente, esta orientación se ve desde el mismo arranque del proyecto complementada y superada por el deseo de hacer un diccionario muy amplio (“el más copioso que pudiere hacerse”, Estatutos, cap. I), lo cual implica incorporar no solo las voces y acepciones que se encuentran en esa selección, sino las que figuran en textos de los más diversos tipos, incluidas las “voces peculiares y propias, que se usan frecuentemente en algunas provincias y reinos de España” (DA: I, v) y, con especial atención, las voces de germanía¹.

El análisis de las citas del DA tiene, pues, un enorme interés desde varios puntos de vista. En primer lugar, permite conocer con detalle el grado en que los primeros académicos llevaron a la práctica su intención inicial de apoyar las entradas con citas extraídas de textos reales, a pesar de las bien conocidas dificultades que implica un proceso de este tipo con los medios propios de la época². En segundo término, analizar de forma exhaustiva la relación de autores y obras utilizadas en las citas, el número en que aparece cada uno de ellos y su distribución a lo largo del diccionario puede proporcionar una imagen fidedigna de las características del DA. El problema es, por supuesto, que el volumen del DA hace que las aproximaciones a este tema hayan sido forzosamente

1. Para los aspectos fundamentales de la génesis, el desarrollo y el contenido de la obra, cf., entre muchos otros trabajos, Lázaro Carreter (1972), Ruhstaller (2001), Freixas (2003 y 2010), Álvarez de Miranda (2005).

2. Realizaron un esfuerzo enorme y, como es lógico, aluden a ello en la introducción del tomo primero, en el que dicen que adoptaron y mantuvieron la decisión de documentar los usos “sin embargo de la gran dificultad que esta resolución incluía: porque hallar en un libro una voz es fortuna que ofrece el acaso, y muchas veces no consigue el mas aplicado estúdio”. Por ello, a pesar del esfuerzo colectivo realizado, “[a]un con toda esta providencia obliga la necesidad à que una, ù otra vez falten autoridades para algunas voces: singularmente para aquellas, cuya vulgaridad las excluye de escritos serios, y no ha logrado el cuidado encontrarlas en los de asunto jocoso; sin que por esto sea culpable la Academia, que se reconoce obligada à todo lo que es estúdio; pero no à lo que en tanta parte pende de la contingencia (DA: I, xviii). Para la cuantificación de este esfuerzo, cf. *infra*.

parciales. Ha habido trabajos centrados en algunos autores, como los de James (1996) para las citas de Góngora, Bravo (1998) con las de Villegas, Florit (2004) para las de Tirso de Molina, Candelas (2004) con respecto a las de Quevedo en el primer tomo del DA o Álvarez de Miranda (2004) con Quevedo y Cervantes en una muestra de 200 páginas de cada uno de los tomos. Otros han estudiado el grado de utilización de algunas obras, como ha hecho Prieto García-Seco (2015) con *La pícaro Justina*. También ha habido análisis de cierto tipo de léxico, como han hecho Gutiérrez Rodilla (1994) con el relacionado con la medicina, Salvador (1985) con las palabras que llevan marca geográfica, Aliaga (1994) para las características de Aragón, Campos y Pérez Pascual (2012) para las relacionadas con Galicia, Ruhstaller (2004 y 2006) y Prieto García-Seco (2010) para las ‘jocosas e inventadas’ y Breuer (2007) ha localizado todas las que llevan una marca. Por fin, Freixas (2003, 2010: 267 y sigs.) ha trabajado con la totalidad de las citas y referencias de dos muestras distintas, que suponen, en conjunto, alrededor del 5% de la obra.

Evidentemente, el progreso en el conocimiento del DA pasa por la elaboración de una versión digitalizada de la obra, que es la tarea que acometió la Real Academia Española en paralelo a la construcción del *Corpus diacrónico del español (CORDE)* y con el objetivo general de preparar materiales para el *Nuevo diccionario histórico del español (NDHE)*. El resultado de ese proyecto es un texto que está a disposición de todas las personas interesadas en la página electrónica de la Academia <<http://web.frl.es/DA.html>>, con una aplicación de consulta desarrollada en el Instituto de investigación Rafael Lapesa.

Aunque ese texto digital está todavía insuficientemente codificado y contiene algunos errores de transcripción, disponer de él me ha permitido (Rojo 2014) aplicar algunas técnicas básicas de recuperación de información y obtener resultados generales que presentan un grado bastante alto de fiabilidad acerca del número de citas que contiene la obra en sus diferentes volúmenes, así como los autores u obras de los que proceden. Limitándome ahora a lo más destacado, el DA contiene unas 71 000 citas textuales (sin contar las no textuales, que Freixas denomina “fuentes”), distribuidas de modo que alrededor de dos terceras partes de las entradas correspondientes a acepciones en sentido estricto (es decir, sin locuciones, expresiones complejas ni refranes) están apoyadas en una o más autoridades (para más detalles, cf. Rojo 2014: 150).

Como se muestra en el trabajo mencionado, los textos de autores pertenecientes a lo más destacado de la tradición literaria constituyen el fondo fundamental del que los académicos extraen la mayor parte de las citas: Quevedo, Cervantes, Lope de Vega y Fray Luis de Granada, que son los más utilizados, suponen en conjunto el 16,62% de todas las citas del DA (cf. Rojo 2014, apéndice 3). A su lado figuran, también con unos índices muy altos de empleo, obras y autores no pertenecientes a la tradición literaria en sentido estricto, como es el caso de la adaptación del *Dioscórides* de Andrés Laguna, la *Nueva*

recopilación de leyes del reino o la *Pragmática de tasas*, texto muy explotado por los académicos, como ha mostrado en varias ocasiones Álvarez de Miranda (2005 y en prensa, entre otros).

En este trabajo me propongo aplicar una metodología similar, pero referida ahora no a los autores, sino a las obras que los primeros académicos utilizan para garantizar el empleo de las diferentes voces en las distintas acepciones que recogen en su obra. No puedo pasar aquí, por limitaciones de espacio, de una presentación general de las grandes cuestiones que se suscitan (apdo. 2) y un rápido análisis de los aspectos más importantes que derivan de los resultados obtenidos (apdo. 3).

2. EL ANÁLISIS DE LAS OBRAS

La lista de abreviaturas utilizadas que figura al comienzo de cada uno de los seis tomos, considerablemente más amplia que la relación de obras seleccionadas por la RAE, muestra que los académicos diseñaron un sistema muy congruente para organizar el modo de referirse a ellas. Las claves de autor (y obra cuando se citan directamente) consisten en una o más palabras con versal inicial y versalitas en el resto de la secuencia seguidas de un punto cuando hay abreviatura y de una coma cuando se da el apellido completo (QUEV./ FR. L. DE GRAN./ C. LUCAN./ AYALA,/)³. Las claves de obras que acompañan a la mención de autor están formadas por una o más palabras con puntos para marcar las abreviaturas; la primera palabra empieza con mayúscula y sigue en minúsculas y el resto, si lo hay, respeta la alternancia habitual de mayúsculas y minúsculas (ALDR. Antig./ ALDR. Orig. de la leng. Castell./ ALVAR. DE TOLLED. Hist. de la Igles./ ARDEMANS. Gobiern. Polit. de las fábricas/)⁴. Pero el recuento exhaustivo de autores y obras mencionadas directamente que se incluye en Rojo (2014: apéndice I) muestra en toda su amplitud lo que se sabía ya desde hace tiempo: las listas de abreviaturas no contienen todas las

3. Para la reproducción de las claves utilizadas en el *DA* empleo aquí tipos de espacio fijo. En la edición electrónica se ha prescindido de la diferencia entre versales y versalitas, convirtiendo todas en versales. En consecuencia, eso mismo hago a partir de este momento. Dado que los puntos y las comas forman parte de la clave, para evitar confusiones con los signos de puntuación del texto utilizo una barra (/) para separar las diferentes claves y también para marcar su final.

4. Hay muchas más posibilidades, por supuesto, pero los ejemplos anteriores serán suficientes para mostrar que es relativamente sencillo extraer automáticamente las secuencias correspondientes a claves de autores y obras citadas directamente (son secuencias de caracteres en mayúsculas no situados a comienzo de párrafo). En cambio, detectar automáticamente dónde termina la mención de la obra y comienza el texto de la cita resulta mucho más complicado o simplemente imposible (cf. *infra*).

correspondientes a las citas realmente utilizadas en el cuerpo de la obra. Como consecuencia, el hecho de poder utilizar una versión electrónica del *DA* como la situada en la página web de la RAE supone un avance importantísimo, pero no garantiza la recuperación de todos los datos relevantes para muchas de las búsquedas de interés. Por ejemplo, Tirso de Molina no figura en la lista de abreviaturas, pero el *DA* contiene ocho textos de este autor bajo diferentes claves⁵. Las claves de *Bocados de oro* que figuran en las listas son BOC. DE OR./BOCAD. DE OR./, pero si nos limitamos a introducir esas secuencias para obtener las citas de la obra perderemos las correspondientes a BOC. DE ORO./BOCAD. DE ORO./, con lo que se nos habría escapado alrededor del 50% de las citas de esta obra en el texto del *DA* (cf. Rojo 2014: apéndice I).

El análisis de las citas de la obra requiere, por tanto, un trabajo previo consistente en la detección automática de las secuencias que pueden contener claves de autores y obras. Dado que una parte de ese trabajo, el relativo a los autores y obras mencionados directamente, ya fue realizado con anterioridad (cf. Rojo 2014), en esta ocasión he podido partir de una lista contrastada y preparar una nueva rutina capaz de recorrer el texto de los seis tomos y detectar las cadenas en las que figuraba cada clave de autor (u obra) y extraer también la secuencia de caracteres que viene a continuación, que es la candidata a contener la clave de la obra. Esta rutina produjo una enorme cantidad de líneas (en realidad, tantas como casos detectados) que fueron reducidas automáticamente mediante la supresión de las repetidas. A pesar de esa reducción, quedaba todavía una gran cantidad de cadenas que contenían la misma clave de obra, pero también algunas indicaciones adicionales (página, folio, tomo, etc.) o los primeros caracteres del texto de la cita. Fue necesario depurar manualmente esa lista hasta lograr reducirla a las combinaciones de autor y obra utilizadas en el diccionario. En el paso siguiente, una nueva rutina realiza el recuento de los casos de las diferentes claves en cada uno de los tomos del *DA*.

El análisis de los resultados obtenidos muestra desajustes similares a los observados en las listas de autores y obras mencionadas directamente, ahora multiplicados como consecuencia del mayor grado de especificidad. Veamos los tipos fundamentales:

(i) Autores y, sobre todo, obras mencionadas directamente que no aparecen en la lista de abreviaturas incluida al comienzo de cada tomo. Así, la revisión de la zona correspondiente a las *Ordinaciones y Ordenamientos* del apéndice I de Rojo (2014) muestra, entre otras, las claves ORDEN. DE ALCAÑ./ ORDEN. DE ALCANIZ./ ORDEN. DE DAROC./ ORDEN DE DAROCA./ ORDEN. DE HUES./ ORDEN. DE HUESC./ ORDEN. DE LA CHANC. DE GRAN./ ORDEN DE LA INFANT./ ORDEN. DE LA PLAN.

5. Cf. Florit (2001), Rojo (2014: 154 y n. 25), Prieto García-Seco (2014) para el análisis de esta interesante cuestión.

MAYOR DEL EXERC./ ORDEN. DEL EXERC. DE FLAND./ ORDEN DEL REGIM. DE GUARD./ ORDEN DEL VALLE DE ARAM./ ORDEN. PARA EL DIRECT. DE LA INFANT./ ORDEN. PARA EL EXERC. DE FLAND./ ORDEN. PARA LA FORMAC. DE LAS MILIC. DEL REINO/, que no están registradas en la lista de abreviaturas ni en esta forma ni en ninguna variante. Un ejemplo llamativo de obra ausente en un autor del que figuran algunas otras es el constituido por los *Entremeses* de Cervantes, que tienen cuatro citas en el texto del *DA* con la clave CERV. Entrem./.

(ii) Claves de autor y obra diferentes en mayor o menor grado de las que figuran en la lista de abreviaturas, como sucede en el caso de los *Bocados de oro* ya mencionado.

Más interesantes para lo que aquí estudiamos son los desajustes que surgen como consecuencia de la indicación de autor y obra:

(iii) Con cierta frecuencia, la lista de abreviaturas contiene indicaciones genéricas del estilo “sus obras”, “su obra poética” o algo parecido. En casos de este tipo, la referencia que figura en el texto puede ser del estilo adoptado en las 90 citas de Álvaro Gómez de Ciudad Real (el joven)⁶, que son todas ellas del tipo ALV. GOM. Cant. 3. Oct. 15 (con variantes en la clave de autor). Resulta también muy llamativo el caso de Ambrosio de Morales, que es uno de los autores más citados del *DA*: de los 1045 textos extraídos de sus obras, 886 (es decir, el 84,78%) aparecen con la simple abreviatura del autor, lo cual resulta bastante confuso en una obra tan extensa⁷. Por fin, hay que mencionar casos como el representado por Calderón, que en la lista de abreviaturas aparece como CALD./ CALDER./ seguido de indicaciones genéricas del tipo Aut./ Com./, pero que en el cuerpo de las entradas lleva casi siempre indicación de la obra concreta de la que procede el ejemplo⁸.

El resultado de todo ello es una gama bastante heterogénea de organización de las obras de los diferentes autores, complicada además por la existencia de variantes gráficas (resultado de las diferentes opciones adoptadas por los académicos, las erratas del texto impreso y los errores de la edición digitalizada). A las diferencias en la mayor o menor amplitud de las obras de un cierto autor se añaden las que resultan de las utilizadas por los académicos para organizarlas. Así, en las listas de abreviaturas, las obras de Cervantes se organizan

6. Poeta renacentista en latín. Supongo que los ejemplos incorporados proceden de su traducción al español del *Triunfo de amor*, de Petrarca (cf. Recio 1996).

7. Las referencias genéricas van seguidas normalmente de indicaciones del tipo tom./ o lib./, algo diferente es lo que sucede con Pedro de Abarca: 12 de las 23 citas tienen únicamente clave de autor a pesar de que en la lista de abreviaturas figura siempre la referencia a los *Anales de Aragón* con diversas variantes. Además, una de las referencias con indicación de obra dice ABARC. Mahometismo./, que parece ser el primero de los capítulos de los *Anales*.

8. Por ejemplo, de las 163 citas de CALD. Aut./ CALDER. Aut./, únicamente 2 carecen de la indicación de la obra concreta.

en *Comedias*, *La Galatea*, *Novelas ejemplares*, el *Persiles*, el *Quijote* y el *Viaje del Parnaso*. Las citas desarrollan las comedias (he registrado cinco distintas, sin tener en cuenta las variantes gráficas) y las novelas (citadas mediante un número de orden casi siempre), además de añadir los ejemplos tomados de los *Entremeses* (4 citas tomadas de dos textos)⁹. El más llamativo es, sin duda, el caso de Calderón. La lista de abreviaturas se limita a diferenciar entre comedias y autos sacramentales (cf. *supra*), mientras que las citas llevan casi siempre la indicación de la obra concreta, de modo que se pueden identificar (sin tener en cuenta variantes gráficas) cerca de 60 autos sacramentales y casi 90 comedias diferentes¹⁰.

La obtención de los datos a partir de lo que contienen los tomos y, sobre todo, las diversas revisiones a que hay que someterlos constituye un proceso pesado y aburrido, pero su resultado será de gran utilidad para quienes quieran profundizar en el análisis de las citas del *DA* o bien en el modo en que sus redactores emplearon las procedentes de un determinado autor u obra. Una muestra clara de la importancia que posee viene de la simple comparación de las cifras. La lista unificada de abreviaturas de los seis tomos (es decir, una vez eliminadas las líneas duplicadas, aunque manteniendo, por razones obvias, las variantes gráficas que afectan a las claves, pero no a sus desarrollos) contiene 930 líneas, mientras que la resultante de la extracción directa del texto del *DA* consta de 4253, cifra que supone más que cuadruplicar la lista de partida. Es decir, el texto del *DA* contiene 3323 claves de autor y obra que no están en la lista de abreviaturas. Por supuesto, no se puede olvidar que las adiciones vienen mayoritariamente de variantes gráficas de las que sí constan en la lista, erratas de la edición impresa o errores del texto digitalizado, pero también hay algunos casos importantes de simple ausencia de la lista de abreviaturas de cualquier referencia a la obra o de concreción de referentes genéricas. En cualquier caso, es evidente que la lista es necesaria para saber qué obras están citadas y también para garantizar que las variantes puramente gráficas no impiden que la extracción de las entradas en las que se cita una determinada obra recoja todos los casos¹¹.

9. Las comedias son *La entretenida*, *Pedro de Urdemalas*, *El rufián dichoso*, *El gallardo español* y *La prudencia en la mujer*. Esta última, por cierto, en la entrada ALANZADA, no es de Cervantes, sino de Tirso de Molina, como ha señalado Prieto García-Seco (en prensa) (cf. Rojo 2014: 154, n. 25). Los entremeses son *El rufián viudo* y *El juez de los divorcios*.

10. Teniendo en cuenta las muchas variantes localizadas, el número de claves correspondientes a Calderón de la Barca se eleva a 429.

11. Dado que las limitaciones de espacio habituales impiden la publicación de listas tan voluminosas, he creado un recurso electrónico en el que es mi intención mantener, en varios formatos, las listas completas de autores y obras citados en el *DA* que he ido elaborando en estos últimos meses, a las que añadiré otros materiales de interés. El acceso a todo ello puede hacerse

El resto del trabajo de preparación es sencillo: una vez elaborada la lista de claves de autor y obra, otra rutina se encarga de detectar su frecuencia de aparición en cada uno de los seis tomos. Lo mismo que en el caso de las relaciones de autores con sus frecuencias (cf. Rojo 2014: 141 y sigs.), la lista resultante ha de ser todavía revisada para descontar los casos en los que se puede haber producido un doble recuento¹².

3. PRIMER ANÁLISIS DE RESULTADOS

Los datos globales resultantes son, como es lógico, muy próximos a los ya publicados a partir de los recuentos de autores (cf. Rojo 2014). Incluyo en el cuadro 1 los obtenidos en este nuevo recuento y los pongo en relación con el número de lemas y acepciones¹³ de cada tomo (cf. Rojo 2014: 152) para facilitar así un elemento de normalización de las cifras.

Tomo	Citas	Lemas	Acepciones	Citas por lema	Citas por acepción
I	12 304	7567	10 033	1,62	1,23
II	13 625	5218	7233	2,60	1,88
III	14 740	8887	11 475	1,66	1,28
IV	11 478	7446	10 306	1,54	1,12
V	10 392	6877	10 420	1,51	1,00
VI	8853	6373	8585	1,39	1,03
Total	71 392	42 368	59 248	1,68	1,20

Cuadro 1. Número de citas, lemas y acepciones en los diferentes tomos del *DA*

Las cifras son ligeramente distintas de las obtenidas en el recuento de autores, donde aparece un total de 70 802 citas. En el recuento por obras la cantidad se incrementa ligeramente (590 casos más), sin duda por efecto de errores cometidos en el proceso de unificación de listas o detección de casos de doble recuento. En cualquier caso, la cifra se sitúa en torno a las 71 000 citas, que es,

desde la sección *Recursos* de mi página electrónica personal <<http://gramatica.usc.es/persoas/guillermo.rojo>>.

12. De las cifras obtenidas para los casos de, por ejemplo, CALD./ Com. han de ser descontados todos los del tipo CALD. Com. Afect. de Odio y amor./, etc. Dejo constancia de la realización de esa operación mediante la marca R que figura en la novena columna del cuadro 2.

13. Utilizo los términos *lema* y *acepción* en el sentido actual, aunque, como es bien sabido, el *DA* sitúa las acepciones en párrafos distintos, lo mismo que las locuciones, expresiones complejas y refranes. Estas últimas entradas no entran en los recuentos que doy a continuación.

como se ha señalado repetidamente, una cantidad realmente asombrosa si tenemos en cuenta las condiciones de la época. En los índices de citas por lema y acepciones del cuadro 1 se observa una línea general que asciende del tomo I al tomo II y luego va descendiendo hasta alcanzar su mínimo en el último tomo, como si el cansancio y las dificultades para encontrar los ejemplos adecuados se fueran haciendo cada vez mayores.

En el cuadro 2 figuran las frecuencias parciales y totales de las 75 obras más citadas en el *DA*¹⁴. No hay en ella nada especialmente llamativo con respecto a lo que ya sabíamos, pero su análisis revela algunos factores que pueden ser de interés. La obra más citada es, por supuesto, el *Quijote*, que concentra el 3,35% del total de las citas presentes en la obra. Siguen, a cierta distancia, la *Nueva Recopilación de las Leyes del Reino* (1567), las que aquí se agrupan como *Musas* de Quevedo y la adaptación del *Dioscórides*, de Andrés Laguna¹⁵. El conjunto de estas 75 (variantes de) obras supone un porcentaje acumulado del 52,54%, similar al que se alcanza con los 35 autores más citados (cf. Rojo 2014: apéndice II). La acumulación es, con toda lógica, claramente inferior a la que se observa con los autores, pero alcanza unos niveles importantes, lo cual revela una notable concentración no ya en ciertos autores, sino también en ciertas obras. Se rebasa ligeramente el 50% de todas las citas con solo las 69 (variantes de) obras más utilizadas.

Obra	T. I	T. II	T. III	T. IV	T. V	T. VI	Total	R	%	% acum.
CERV. Quix.	554	378	717	392	195	166	2402		3,36	3,36
RECOP.	156	513	422	188	190	59	1528	R	2,14	5,50
QUEV. Mus.	400	196	333	223	157	210	1519		2,13	7,63
LAG. Diosc.	216	382	215	244	217	179	1453		2,04	9,67
SAAV. Empr.	284	284	219	106	103	102	1098		1,54	11,21
SOLIS, Hist. de Nuev. Esp.	182	191	290	176	148	89	1076		1,51	12,71
MARIAN. Hist. Esp.	159	275	360	118	137	23	1072		1,50	14,21
OV. Hist. Chil.	422	145	152	139	59	28	945		1,32	15,54
AMBR. MOR.	129	121	135	143	198	148	874	R	1,22	16,76
ALFAR.	202	141	172	122	99	118	854		1,20	17,96
RIBAD. Fl. Sanct.	0	251	286	105	158	23	823		1,15	19,11
PRAGM. DE TASS. año 1680.	147	169	102	122	96	79	715		1,00	20,11

14. La clave *R* indica que se ha hecho el reajuste necesario por alguna coincidencia parcial. Puede obtenerse la lista completa en Rojo (2015).

15. En el cuadro figuran las variantes de estas obras que concentran mayor número de apariciones, pero, en estos casos, la consideración del total de las variantes no introduce diferencias significativas.

FR. L. DE GRAN. Symb.	0	159	183	142	126	86	696		0,97	21,09
PIC. JUST.	103	134	168	150	78	58	691		0,97	22,06
ACOST. Hist. Ind.	125	109	108	82	172	58	654		0,92	22,97
COMEND. sob. las 300.	215	130	76	146	55	19	641		0,90	23,87
ESTEB.	213	133	86	43	57	101	633		0,89	24,76
GRAC. Mor.	171	128	50	74	103	100	626		0,88	25,63
CORN. Chron.	1	0	198	159	188	56	602		0,84	26,48
JACINT. POL.	105	113	103	72	135	63	591		0,83	27,30
CHRON. GEN.	188	147	105	48	35	48	571		0,80	28,10
PART.	52	155	166	44	74	74	565		0,79	28,90
ESPIN. Escud.	47	107	97	134	76	45	506		0,71	29,60
MARM. Descripc.	0	112	74	88	106	120	500	R	0,70	30,30
MEND. Guerr. de Gran.	143	165	96	48	16	10	478		0,67	30,97
HORTENS. Mar.	48	49	145	53	70	104	469	R	0,66	31,63
FUENM. S. Pio V.	94	36	99	74	103	59	465		0,65	32,28
NAVARR. Man.	45	88	126	83	85	31	458		0,64	32,92
PELLIC. Argen.	87	94	66	129	46	34	456		0,64	33,56
LOP. Dorot.	98	59	97	90	77	33	454		0,64	34,20
M. AGRED.	33	51	99	99	85	80	447		0,63	34,82
QUEV. Fort.	79	67	120	68	56	50	440		0,62	35,44
ALCAZ. Chron.	28	84	60	55	117	89	433		0,61	36,05
QUEV. Tacañ.	121	72	148	83	2	0	426		0,60	36,64
VALVERD. Vid. de Christ.	43	132	156	67	15	7	420		0,59	37,23
CERV. Persil.	105	42	69	40	61	91	408		0,57	37,80
ARGENS. Maluc.	92	71	58	103	36	15	375		0,53	38,33
NAVARR. Conserv.	24	139	98	39	38	29	367		0,51	38,84
HUERT. Plin.	0	0	92	143	88	43	366		0,51	39,36
HORTENS. Paneg.	47	55	125	30	26	76	359		0,50	39,86
COVARR.	121	97	53	18	26	32	347	R	0,49	40,34
GONG. Rom.	96	57	73	49	27	37	339		0,47	40,82
FONSEC. Vid. de Christ.	0	151	56	50	48	31	336		0,47	41,29
RECOPI. DE IND.	7	162	35	70	37	24	335		0,47	41,76
SIGUENZ. Vid. de S. Geron.	52	55	127	31	38	30	333		0,47	42,23
NUÑ. Empr.	49	45	76	47	27	73	317		0,44	42,67
PALOM. Mus. Pict.	0	98	49	51	41	78	317		0,44	43,11
TORR. Philos.	0	56	44	45	53	95	293		0,41	43,52
MANER. Apolog.	73	63	34	64	34	17	285		0,40	43,92
HORTENS. Quar.	9	2	75	30	41	119	276	R	0,39	44,31

LOP. Philom.	141	31	53	27	3	15	270		0,38	44,69
QUEV. Cuent.	81	59	42	29	27	31	269		0,38	45,07
ESPIN. Art. Ballest.	0	84	47	67	30	34	262		0,37	45,43
ERCILL. Arauc.	41	28	89	16	23	64	261		0,37	45,80
MEX. Hist. Imper.	2	49	105	27	21	49	253		0,35	46,15
PANT. Rom.	81	49	46	31	12	34	253	R	0,35	46,51
SART. P. Suar.	0	44	65	61	75	3	248		0,35	46,85
ILLESC. Hist. Pontif.	3	109	59	29	31	13	244		0,34	47,20
COLMEN. Hist. Segob.	0	91	93	39	19	1	243		0,34	47,54
BARBAD. Coron.	69	91	29	29	15	6	239	R	0,33	47,87
VILLAV. Mosch.	1	1	0	26	101	109	238		0,33	48,20
FUER. JUZG.	73	41	53	8	36	25	236		0,33	48,53
MEN. Coron.	59	68	42	33	33	0	235		0,33	48,86
CIENF. Vid. de S. Borj.	0	0	51	78	46	56	231		0,32	49,19
GUEV. Epist.	30	76	62	21	28	12	229		0,32	49,51
TEJAD. León Prodig.	0	22	70	91	44	0	227		0,32	49,83
FR. L. DE LEÓN, Nomb. de Christ.	0	56	103	34	33	0	226		0,32	50,14
C. LUCAN.	66	38	50	23	21	27	225		0,32	50,46
PARTID.	132	90	0	0	0	0	222		0,31	50,77
SANDOV. Hist. de Carl. V.	3	56	51	22	18	67	217		0,30	51,07
PUENT. Conven.	33	47	52	21	30	33	216		0,30	51,38
SOLD. PIND.	14	41	33	56	37	30	211		0,30	51,67
ESPIN. Escuder.	209	1	0	0	0	0	210		0,29	51,97
CAST. Hist. de S. Dom.	0	28	18	66	61	33	206		0,29	52,25
LOP. Circ.	81	34	44	29	5	13	206		0,29	52,54

Cuadro 2. Frecuencias y porcentajes de las 75 obras más citadas en el *DA*

De todas formas, por lo ya visto en el apartado anterior es necesario tomar los datos concretos con las precauciones aconsejadas por el diferente tratamiento que el sistema de referencias impone a las obras citadas. Por ejemplo, lo que en el cuadro 2 aparece como Mus./ va en el texto acompañado del número correspondiente a la musa que justifica la agrupación de poemas, lo cual permite saber que esta abreviatura no puede corresponder solo a las composiciones que aparecen en *Las tres Musas últimas castellanas. Segunda cumbre del Parnaso español*, publicadas en 1670, tras la muerte de Quevedo, por su sobrino Pedro Aldrete Quevedo y Villegas, sino que también hay citas de las seis primeras, publicadas con el título *El Parnaso español, monte en dos cumbres dividido, con las nueve Musas* por González de Salas en 1648, pero que solo comprende las seis primeras. Por tanto, esta referencia debe ser interpretada como la correspondiente a la mayor

parte de la obra poética de Quevedo, lo cual tiene una importancia evidente (cf. *infra*). De otra parte, la clave AMBR. MOR./ se refiere al conjunto de las obras de Ambrosio de Morales, de modo que no es fácil evaluar el peso que hay que atribuir a este dato al compararlo con los referentes a obras concretas que tiene a su alrededor. Una consecuencia clara de los desajustes causados por la falta de uniformidad en las claves la evidencia la línea correspondiente al *Símbolo de la fe* de Fr. Luis de Granada. La sorprendente ausencia de citas de esta obra en el tomo I se explica sencillamente por el hecho de que 402 de las 404 citas de este autor en el primer tomo aparecen bajo la clave FR. LUIS DE GRAN./, que solo se utiliza dos veces más en el resto de los tomos. Si añadimos las 155 citas de la obra que figuran con esta otra clave en el tomo I se obtiene la cifra de 851, que la sitúa inmediatamente después del *Guzmán*. Algo diferente sucede con el *Flos sanctorum*, de Pedro de Ribadeneira, publicado entre 1599 y 1601, que alcanza un número importante de apariciones, pero ninguna en el tomo I. En este caso, lo que sucede es que los académicos no empiezan a utilizar con frecuencia las obras de este autor hasta el segundo tomo: de las 1123 citas de Ribadeneira en la totalidad del *DA*, solo 21 corresponden al tomo I (12 de ellas proceden del *Flos Sanctorum*). No muy diferente es lo que sucede con Damián Cornejo: de las 770 citas de su *Crónica de san Francisco* (1682-1698) en el *DA*, únicamente 15 están en el primer tomo, pero en el segundo hay ya 146. Fundir las variantes de las referencias a esta obra la situaría por encima de la *Pragmática de tasas*. Como última muestra de las precauciones con que hay que interpretar los resultados que figuran en la lista de frecuencias por obras, téngase en cuenta que el conjunto de las *Novelas ejemplares* de Cervantes tiene un total de 599 citas, de modo que su consideración conjunta las situaría entre las veinte más citadas.

A pesar de todos estos inconvenientes, superables con el análisis detenido de las listas y el conocimiento de las circunstancias que las rodean, los datos que figuran en el cuadro 2 confirman, ahora con datos totales, lo que se ha venido diciendo acerca del tipo de textos utilizados por los primeros académicos para apoyar las definiciones incluidas en el *DA*. De las diez obras a las que recurren con más frecuencia, solo tres se sitúan entre las de creación literaria en sentido estricto¹⁶.

También es significativo de la orientación descriptiva y filológica que tiene la obra la importancia otorgada a los textos medievales, que hace que la *Crónica general* y las *Partidas* figuren entre las más citadas¹⁷. En realidad, su importancia es mayor que la mostrada en el cuadro 2, ya que, si bien la *Crónica general* apenas tiene casos con variantes distintas de CHRON. GEN./, las

16. Debe tenerse en cuenta, además, que las diversas variantes de la *Pragmática de tasas de 1680* suponen, como mínimo, un total de 774 citas.

17. Y a cierta distancia de ellas, el *Fuero Juzgo*, con 242 casos, y el *Fuero Real*, con 111.

Partidas aparecen también citadas como PARTID./ 222 veces, todas ellas entre los tomos I y II. Sumando ambas cifras, el total se eleva a 787 apariciones.

La diferenciación de las frecuencias por obras permite tener también una idea mucho más ajustada del peso de cada autor y, hasta cierto punto, de las causas de su utilización como justificación de un determinado significado. El análisis de las frecuencias de las citas tomadas de obras de Cervantes que figuran en el cuadro 3, fundidas en algunos casos para mayor claridad, muestra que el *Quijote*, con un notable equilibrio entre las dos partes, supone alrededor de los dos tercios de todas las citas de este autor. El conjunto de las *Novelas ejemplares* y el *Persiles* tienen también un peso importante, lo cual implica, como es lógico, que las comedias, los entremeses, *La Galatea* y el *Viaje del Parnaso* tienen una importancia considerablemente menor (en torno al 5% en total).

	T. I	T. II	T. III	T. IV	T. V	T. VI	Total
Comedias	4	0	1	1	3	2	11
Entremeses	3	0	0	1	0	0	4
<i>La Galatea</i>	1	0	4	22	14	1	42
<i>Novelas ejemplares</i>	125	74	71	98	94	139	601
<i>Persiles</i>	105	43	70	40	61	91	410
<i>Quijote</i> (primera parte)	280	151	390	184	102	53	1160
<i>Quijote</i> (segunda parte)	269	219	316	204	90	113	1211
<i>Parnaso</i>	18	26	30	16	14	19	123
Totales	805	513	882	566	378	418	3562

Cuadro 3. Distribución de las citas de Cervantes en el *DA*

Reveladoras de cierta parte de la organización del *DA* resultan las citas de las *Musas* de Quevedo. Según muestra el cuadro 4¹⁸, proceden de este conjunto de composiciones poéticas 1551 citas, que suponen nada menos que el 35,77% del total de los textos tomados de este autor. Pero más importancia tiene el análisis de la procedencia completa de estos ejemplos. Resulta que 1280 casos, es decir, el 82,52% de las citas procedentes de las *Musas* se concentran en las musas quinta (Terpsícore, con 426 ejemplos) y, sobre todo, la sexta (Talía, con 854 textos de apoyo). Son las composiciones que contienen los poemas satíricos y burlescos, jácaras, bailes y bromas. Estas 854 citas suponen casi el 20%

18. He acumulado los resultados de las variantes utilizadas para las *Musas*, de ahí sus discrepancias con lo que puede observarse en la segunda línea del cuadro 2.

de todas las procedentes de Quevedo¹⁹. Creo que aquí tenemos un importante factor explicativo de los altos índices de utilización de textos quevedescos: esos dos conjuntos de composiciones poéticas son, sin duda, la fuente de gran cantidad de palabras o acepciones que los académicos redactores califican como “voz inventada”, “voz festiva”, “voz jocosa”, “voz familiar”, etc.²⁰. Un recuento muy superficial muestra que hay en el *DA* 272 voces que llevan la indicación *jocosa*. De ellas, 126 (el 46,32%) contienen un texto de apoyo tomado de Quevedo y 57 (20,95% del total y 45,23% de las procedentes de este autor) corresponden a las Musas 5 o 6. Parecen cifras reveladoras de la lógica atención a estos textos para autorizar ciertos usos²¹.

	T. I	T. II	T. III	T. IV	T. V	T. VI	Total
Musa 1	6	0	10	0	1	4	21
Musa 2	26	4	21	7	1	7	66
Musa 3	2	0	10	3	1	1	17
Musa 4	9	1	30	7	8	5	60
Musa 5	128	51	70	69	43	65	426
Musa 6	231	125	178	107	99	114	854
Musa 7	17	9	8	25	2	6	67
Musa 8	3	4	3	2	0	2	14
Musa 9	11	2	3	3	1	6	26
Totales	433	196	333	223	156	210	1551

Cuadro 4. Distribución de las citas procedentes de las *Musas* de Quevedo

Es fácil ver que las cuestiones aquí simplemente esbozadas son solo una parte minúscula de las muchas que la investigación detenida de la distribución de las citas textuales presentes en el *DA* podrá poner de relieve en el futuro. Lo que sabemos hasta el momento refuerza, cada vez con más intensidad, el carácter excepcional del trabajo realizado por los primeros académicos.

19. El porcentaje de citas procedentes de las *Musas* con respecto al total de las tomadas de textos de Quevedo se sitúa en torno al 40% en casi todos los tomos. Las excepciones son el segundo (30,76%) y el tercero (28,73%).

20. Cf. Candelas (2004: 83) para una consideración similar. Para un análisis de las palabras ‘jocosas’ e ‘inventadas’, vid. Ruhstaller (2004, 2006) y Prieto García-Seco (2010).

21. Los recuentos anteriores han sido realizados mediante la aplicación de consulta del *DA* en la página electrónica de la RAE. Candelas (2004: 84 y ss.) da una relación de las palabras del tomo I que están apoyadas con textos de Quevedo. Es fácil recuperar la totalidad de estas entradas usando operadores booleanos en la página de consulta de la RAE. Naturalmente, hay que tener en cuenta las variantes gráficas utilizadas a lo largo del texto, a las que se puede acceder consultando la lista completa de claves de autor y obra incluida en Rojo (2015).

REFERENCIAS BIBLIOGRÁFICAS

- ÁLVAREZ DE MIRANDA, PEDRO (2004): “Quevedo en la lexicografía española”, *Edad de Oro*, XXIII, 389-416.
- (2005): “La proeza de un diccionario con textos. Algo más sobre el *Diccionario de Autoridades*”, en T. Martínez Romero (ed.), *Les lletres hispaniques als segles XVI, XVII i XVIII*, Castellón: Universitat Jaume I/Fundació Germà Colón, 73-92. Cito por su reedición en Álvarez de Miranda (2011), 17-36.
- (2011): *Los diccionarios del español moderno*, Gijón: Trea.
- (en prensa): “Sobre fuentes no literarias del *Diccionario de Autoridades*, con especial atención a la ‘Pragmática de tasas’ (1680)”, en M. Fernández Alcaide, E. Leal Abad y Á. Octavio de Toledo Huertas (eds.), *En la estela del Quijote: cambio lingüístico, normas y tradiciones discursivas en el siglo XVII*, Frankfurt a.M: Peter Lang.
- BREUER, JUDITH (2000): *Normen und normative Konzepte der Real Academia Española. Diasystematische Markierungen im ‚Diccionario de Autoridades‘*, Bonn: GRIN Verlag.
- CAMPOS SOUTO, MAR y JOSÉ IGNACIO PÉREZ PASCUAL (2010): “Los elementos gallegos en el *Diccionario de Autoridades*”, en A. Nomdedeu Rull, E. Forgar Berdet y M. Bargalló Escrivà (eds.), *Avances de lexicografía hispánica*, Tarragona: Publicacions URV, vol. I, 259-270.
- CANDELAS COLODRÓN, MANUEL ÁNGEL (2004): “Quevedo y el *Diccionario de Autoridades*”, en L. Schwartz (ed.), *Studies in honor of James O. Crosby*, Newark: Juan de la Cuesta, 69-89.
- FLORIT DURÁN, FRANCISCO (2001): “La nómina del *Diccionario de Autoridades*: el caso de Tirso de Molina”, en I. Pardo Molina y A. Serrano (eds.), *En torno al teatro del Siglo de Oro. XV jornadas de teatro del Siglo de Oro*, Almería: Instituto de Estudios Almerienses, 71-84.
- FREIXAS ALÁS, MARGARITA (2003): *Las autoridades del primer diccionario de la Real Academia Española*, Barcelona: Universitat Autònoma de Barcelona. Tesis doctoral [en línea] <<http://www.tdx.cesca.es/TDX-0611104-150443>>.
- (2010): *Planta y método del Diccionario de Autoridades. Orígenes de la técnica lexicográfica de la Real Academia Española (1713-1739)*. Anexo 14 de la *Revista de Lexicografía*.
- GUTIÉRREZ RODILLA, BERTHA M. (1994-95): “Construcción y fuentes utilizadas para los términos médicos en el *Diccionario de Autoridades*”, *Revista de Lexicografía*, I, 149-162.
- JAMMES, ROBERT (1996): “Góngora en el *Diccionario de Autoridades*”, en *Philologica. (Homenaje al profesor Ricardo Senabre)*, Cáceres: Universidad de Extremadura, 247-272.
- LÁZARO CARRETER, FERNANDO (1972): *Crónica del Diccionario de Autoridades. (1713-1740)*. Discurso de ingreso en la Real Academia Española. Madrid, 1972. Cito por su reedición en la serie *Discursos de ingreso en la Real Academia Española*. Madrid: Biblioteca Nueva, 2014.
- PRIETO GARCÍA-SECO, DAVID (2010): “La procedencia de las voces ‘caprichosas’ y ‘jocosas’ del *Nuevo diccionario de la lengua castellana* (1846) de Vicente Salvá”, en A. Medina Guerra y M. C. Ayala Castro (ed. y coord.), *Los diccionarios a través*

- de la historia, Málaga: Servicio de Publicaciones de la Universidad de Málaga, 505-525.
- (2014): *Cuatro siglos de lexicografía española. La recepción de Tirso de Molina en los diccionarios del español*. Anexo de la *Revista de Lexicografía*. Universidade da Coruña.
- (2015): “La pícaro Justina en el Diccionario de Autoridades”, en J. M. García Martín, F. J. de Cos Ruiz y M. Franco Figueroa (eds.), *Actas del IX Congreso internacional de historia de la lengua española (Cádiz, 2012)*. Madrid/Frankfurt: Iberoamericana/Vervuert, 1593 – 1614.
- REAL ACADEMIA ESPAÑOLA (1626-1639[2013]): *DA (Diccionario de Autoridades): 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 tomos. Madrid. 1626-1639. Edición facsimilar, Real Academia Española, 2013 [en línea] <<http://web.frl.es/DA.html>>.
- (1626-1639[2013]): “Estatutos de la Real Academia Española”, en *DA*, I, xxiii-xxx.
- RECIO, ROXANA (1996): *Petrarca y Álvaro Gómez: la traducción del Triunfo de amor*, Nueva York: Peter Lang.
- ROJO, GUILLERMO (2014): “Análisis cuantitativo de las citas del *Diccionario de Autoridades*”, *BRAE*, XCIV, 137-196.
- (2015): “Cuantificación de las citas textuales del *Diccionario de Autoridades*”. Publicación electrónica en <<http://gramatica.usc.es/persoas/guillermo.rojo/>>.
- RUHSTALLER, STEFAN (2001): “Las autoridades del *Diccionario de Autoridades*”, en S. Ruhstaller y J. Prado Aragonés (eds.): *Tendencias en la investigación lexicográfica del español*, Huelva: Universidad de Huelva, 193-224.
- (2004): “Voces ‘inventadas’ y ‘voluntarias’ en el Diccionario de Autoridades. Su caracterización en la microestructura”, en Prado Aragonés, Josefina y M.^a Victoria Galloso Camacho (eds.): *Diccionario, léxico y cultura*. Servicio de Publicaciones Univ. de Huelva, 175-187.
- (2006): “Voces ‘inventadas’ y ‘voluntarias’ en la macroestructura del Diccionario de Autoridades”, en J. J. de Bustos Tovar y J. L. Girón Alconchel (eds.), *Actas del VI Congreso internacional de historia de la lengua española*, Madrid: Arco Libros, vol. II, 1673-1682.
- SALVADOR ROSA, AURORA (1985): “Las localizaciones geográficas en el *Diccionario de Autoridades*”, *LEA*, 8, 103-139.