

Las nuevas herramientas para la investigación en gramática histórica

INMACULADA C. BÁEZ MONTERO
Universidade de Vigo

0. INTRODUCCIÓN

La evolución de las lenguas, como siempre, evidencia los cambios que se producen en la sociedad. El rastro de dichos cambios no se detecta solamente en las lenguas; también las herramientas de las que nos servimos para trabajar en ellas son una prueba de las huellas indelebles del paso del tiempo.

Las herramientas tradicionales para la investigación en diacronía han seguido el ritmo de los tiempos y se han adaptado, se han hecho más manejables. Las nuevas tecnologías han creado nuevos instrumentos que nos permiten profundizar más en el estudio de las lenguas en estadios anteriores, agilizan las pesadísimas tareas de recolección, almacenamiento y gestión de datos y las hacen más llevaderas, más certeras y más exhaustivas. Actualmente disponemos de herramientas de extrema utilidad para la investigación, tanto en historia de la lengua como en gramática histórica, por lo que se han abierto nuevas expectativas al lento y solitario trabajo del filólogo.

Los investigadores de la diacronía seguimos usando la pluma de ave para matizar informaciones que nos llegan a través de la RED; la lingüística informática y la lingüística computacional, junto con las nuevas "rutas" de información, "vuelan" sobre la gramática histórica e historia de la lengua facilitando en algunos aspectos el desarrollo de la investigación en diacronía.

Las herramientas tradicionales modernizadas, las de nueva creación o en proceso de elaboración y los grupos de investigación que crean y renuevan las herramientas y el proceso investigador en gramática histórica e historia de la lengua española serán objeto de nuestro estudio.

Afortunadamente las ideas no son pocas y continuamente se generan nuevos proyectos; por eso, no es sino desconocimiento u olvido, solo a nosotros imputable, el que aquí no aparezcan algunos.

1. LA DIACRONÍA EN LOS GRUPOS DE INVESTIGACIÓN EN LINGÜÍSTICA INFORMÁTICA.

Tradicionalmente, la investigación en lingüística histórica era individual, pero la magnitud de los proyectos y la envergadura de los mismos han propiciado la formación de grupos de investigación. El informe sobre grupos de investigación en procesamiento del lenguaje y del habla en España¹ traza el panorama de las actividades relacionadas con el procesamiento del lenguaje y el habla que actualmente se llevan a cabo en el estado español y, aunque el núcleo central de los trabajos en lingüística informática² lo constituye el español actual, los estadios anteriores del español son también centro de interés de algunos de los grupos³ que han respondido a la encuesta base del citado informe: el grupo de Lexicografía de la Real Academia Española, el Laboratorio de Lingüística Informática de la Universidad Autónoma de Madrid y el Departamento de Filología Española, Teoría de la Literatura y Lingüística General de la Universidad de Santiago de Compostela.

- El Grupo de Lexicografía de la Real Academia Española⁴ en la actualidad desarrolla dos proyectos directamente relacionados con la diacronía del español: la creación de un Corpus Diacrónico del Español (CORDE) (cfr. *infra* § *Corpora*) y la creación de un sistema informatizado de elaboración de diccionarios (cfr. *infra* § *Diccionario*).
- El Laboratorio de Lingüística Informática de la Universidad Autónoma de Madrid⁵ cuenta con varios proyectos pero el más directamente relacionado con la diacronía es el proyecto ADMYTE (cfr. *infra* §§ *Selecciones bibliográficas, Edición de textos en formato electrónico, Otras herramientas*).

¹ Informe publicado conjuntamente por la SEPLN (Sociedad Española para el Procesamiento del Lenguaje Natural) y el OEIL (Observatorio Español de Industrias de la Lengua) en septiembre del presente año. También está disponible en la página WEB del Observatorio Español de Industrias de la Lengua (<http://www.cervantes.es/oeil/Oeiles.htm>). La página de la SEPLN es <http://sisx04.si.ehu.lsi/Activities/SEPLN>.

² En el informe SEPLN y OEIL se recoge información sobre 21 grupos I+D, ninguno de los cuales se dedica en exclusiva a la diacronía. El resultado del informe debe ser tenido en cuenta, aunque el sistema de obtención de información no haya sido diseñado específicamente para recoger datos sobre grupos de investigación en diacronía del español.

³ En ese mismo informe se hace referencia a trabajos de otros grupos pero solamente aparece la información recogida mediante un cuestionario difundido en su día a través de la página WEB del Observatorio de Industrias de la Lengua.

⁴ Los miembros del grupo son: Guillermo Rojo, Rafael Rodríguez Marín, Abraham Madroñal, Fernando Sánchez León, Marta Pino, Paula Santalla, Amalio Nieto, Jordi Porta, José Luis Sancho, Carlos Domínguez, Ángeles Blanco, Laura Marso de Frutos, Inmaculada García, Marta García, Silvia Fernández Alonso y Encarnación Raigal.

⁵ El grupo está constituido por Francisco Marcos Marín, Antonio Moreno Sandoval, Fernando Sánchez León y Flora Ramírez Bustamante, así como los becarios Enrique López Díaz, Isabel Nieto, Anatoli Adamidou y Nieves Suárez.

- También tenemos constancia de las investigaciones que se llevan a cabo en el Departamento de Filología Española, Teoría de la Literatura y Lingüística General de la Universidad de Santiago de Compostela⁶, en las que también participan profesores de la Universidad de Vigo. El objetivo del grupo dedicado a diacronía del español es la creación de un corpus sintáctico del español medieval y clásico (cfr. *infra* § *Corpora*) que será la base para la elaboración de un diccionario de construcción y régimen en español medieval y clásico.
- El Seminario de Lingüística Informática de la Universidad Autónoma de Barcelona⁷, según el Observatorio de Industrias de la Lengua, cuenta con dos proyectos orientados hacia el estudio histórico del vocabulario español: 1) la informatización del Diccionario Crítico Etimológico Castellano e Hispánico de J. Corominas y J. A. Pascual (cfr. *infra* §§ *Diccionarios* y 2) la creación de un corpus del español medieval para el estudio del español jurídico: el Archivo Informatizado de Textos Jurídicos Medievales (AITJUM) (cfr. *infra* § *Corpora*)

2. SELECCIONES BIBLIOGRÁFICAS

Hasta ahora, una de las tareas más difíciles en la investigación era la de conseguir bibliografía. Gran parte de los interesantísimos trabajos presentados y publicados en congresos o seminarios no son fácilmente accesibles y su localización en las bibliotecas a veces resulta una tarea desesperante. La informática no resuelve el problema, pero nos facilita las cosas. La selección y localización de publicaciones sobre diacronía del español es ahora más rápida y eficaz merced a los repertorios y las bases de datos bibliográficas existentes en CD-ROM así como por las búsquedas en catálogos informatizados de acceso remoto.

2.1. Repertorios bibliográficos

Repertorio Bibliográfico Básico de Sintaxis Histórica del Español

Antonio Narbona Jiménez publicó en el segundo número de la revista de la Universidad de Córdoba, *Alfinge*, un “Repertorio Bibliográfico Básico de Sintaxis Histórica del Español”. Como nos explica el autor en la *Addenda* editada en el siguiente número de la revista, lo que pretendía era “poner a disposición de los estudiantes de filología una información necesaria para adentrarse en una disciplina que carece de una obra general de carácter sistemático” (1985: 277). Esta selección de es-

⁶ El grupo dedicado al español medieval y clásico dirigido por Emilio Montero Cartelle, está constituido por Elena Rivas, Marta Blanco, Teresa Díaz, Victoria Crego e Inmaculada Báez.

⁷ Dirigido por J. M. Bleuca, cuenta con varios proyectos. En algunos de ellos participan Gloria Clavería, Carlos Sánchez Lancis, Julia Giménez y Joan Torruella.

tudios sobre problemas de sintaxis histórica del español también incluía un propósito latente: “invitar a cuantos trabajan en este campo a completarla, actualizarla, cribarla, corregirla, etc.” (1985: 277).

Boletín Bibliográfico de Historia de la Lengua Española

Tras la publicación de este repertorio se crea un nuevo vacío. Será necesario esperar hasta 1994 para que aparezca el *Boletín Bibliográfico de Historia de la Lengua Española*, un trabajo de los profesores José Manuel Fradejas Rueda (UNED), Pilar Díez de Revenga Torres (Universidad de Murcia) y José Ignacio Pérez Pascual (Universidade da Coruña), publicado bajo los auspicios de la Asociación de Historia de la Lengua Española. El objetivo del boletín es “recopilar toda la producción bibliográfica publicada en revistas, misceláneas, actas de congresos, homenajes, libros y cualesquiera otras publicaciones que den cabida a trabajos sobre Historia de la Lengua Española” (1994: 5). En el boletín, que cronológicamente llega hasta 1990, se contempla la inclusión de trabajos sobre el Español de América porque tiene “sus foros especializados”.

Los autores organizan la información en apartados dedicados a: 1) Bibliografía, 2) Generalidades, 3) Teoría y métodos de la lingüística histórica, 4) Obras generales, 5) Historia de la lengua literaria, 6) Historiografía de la lengua española, 7) Fonética, fonología, grafémica y ortografía, 8) Morfosintaxis, Lexicografía y semántica, 9) Onomástica, 10) Dialectología y sociolingüística, 11) Enseñanza. La obra se completa con tres utilísimos índices que además de dar nuevos inventarios temáticos facilitan las búsquedas: 1) índice de autores, 2) de obras y 3) de palabras.

El esfuerzo de recolección es enorme y los resultados así lo confirman, si bien se produce un pequeño desfase cronológico al existir un intervalo de cuatro años entre la fecha límite de la selección y la fecha de publicación. Como los autores no se han limitado a la sintaxis, se ha superado el primer objetivo propuesto por Narbona, aunque las dificultades técnicas siguen siendo una barrera que solo conseguiremos superar con la utilización de las nuevas tecnologías.

Boletín Bibliográfico de Historia de la Lengua Española en Internet

En 1997 apareció, en el servidor de la UNED, una versión electrónica del *Boletín Bibliográfico de Historia de la Lengua Española*. A diferencia del “primer boletín” en formato tradicional (en papel), incorpora información no contrastada extraída de catálogos de librerías, fotocopias de artículos y separatas insuficientemente identificadas. Es la única recopilación específica para la Historia de la Lengua Española e incorpora más de 1.350 referencias bibliográficas desde 1990 hasta hoy. Supone, pues, un cambio radical en la consulta bibliográfica de la historia de la lengua española porque: 1) ofrece la información actualizada (la última revisión es de marzo de 1997 y probablemente será revisada cada poco tiempo) de “toda la producción bibliográfica que sobre los diversos aspectos de la historia de la lengua española se

publican, bien sean monografías, artículos o colaboraciones en revistas, homenajes, misceláneas, etc.”, 2) convierte a los usuarios de la base bibliográfica en colaboradores: nos solicitan ayuda, ya que algunas referencias o bien son incompletas o bien no han sido contrastadas, 3) es una obra abierta, ofrecen una dirección de correo electrónico donde enviar los mensajes con “las correcciones o las aportaciones, o los comentarios que crean pertinentes”.

El objetivo es conseguir “una buena bibliografía de Historia de la Lengua Española” (<http://www.uned.es>). Mediante el tratamiento informático, los autores se proponen superar las dificultades que hasta hace poco tiempo creaba el manejo de grandes volúmenes de datos. Al mismo tiempo, la ubicación de los mismos en internet facilita las consultas y permite la cooperación interactiva de los usuarios, a quienes convierte en colaboradores inmediatos.

Aunque han pasado más de diez años desde la publicación del repertorio bibliográfico de Narbona Jiménez, la versión electrónica del *Boletín Bibliográfico de Historia de la Lengua Española* supone la germinación de la idea allí lanzada. Con el nuevo soporte el “propósito latente” que antes mencionábamos se hace patente; las nuevas tecnologías convierten la corrección, tradicionalmente tarea *a posteriori*, en tarea inmediata⁸.

Philobiblon

Es una base bibliográfica interactiva, con más de cuatrocientos cincuenta campos y once tablas, publicada en el CD-ROM instrumental del proyecto ADMYTE elaborado por el grupo de investigación *Laboratorio de Lingüística Informática* con la colaboración de la Universidad de California en Berkeley y la Universidad Complutense de Madrid, que incluye:

- La *Bibliografía Española de Textos Antiguos (BETA/BOST)*
- La *Bibliografía de Textos Antiguos Portugueses (BITAPA/BOPT)*
- La *Bibliografía de Textos Catalanes Antiguos (BITECA/BOCT)*.

2.2. Bases de datos bibliográficas en CD-ROM o por acceso remoto

En la pequeña muestra que recogemos a continuación incluimos algunas bases de datos especialmente interesantes para facilitar las pesquisas bibliográficas en la diacronía del español. Aunque no nos hemos restringido a las escritas en español, el repertorio que sigue no es más que una modesta selección. Simplemente pretendemos hacer pequeñas calas que reflejen la disponibilidad y accesibilidad de datos bibliográficos. La proliferación de estas bases de datos es tan rápida como fructífera, por lo que requieren un tratamiento exclusivo.

⁸ A la hora de cerrar este trabajo el *Boletín bibliográfico* fue retirado de la página WEB en la que se hallaba. Esperamos que su ausencia sea temporal.

Las bases de datos del Ministerio de Educación y Ciencia

Este Ministerio ha iniciado la publicación, en CD-ROM y por acceso remoto, de bases de datos que contienen información sobre archivos, libros publicados en España, tesis doctorales leídas en las universidades españolas desde 1976 (Teseo), etc. Algunas de estas bases de datos son:

Base de datos ISBN:

<http://mec.es/isbn>

Base de datos BARC Archivos:

<http://mcu.es/pic/spain/barc.html>

Base de datos EDIT Censo editorial:

<http://mcu.es/pic/spain/edit.html>

Base de datos TESEO Tesis Doctorales Españolas:

<http://mec.es/teseo>

Bibliografía Nacional Española desde 1976

Esta base de datos en español producida por la Biblioteca Nacional⁹ contiene el registro de las monografías ingresadas en la Biblioteca Nacional por depósito legal desde 1976 hasta a la actualidad. No se reduce al ámbito de la filología o la lingüística sino que es multidisciplinar. Es actualizada trimestralmente.

Otras bases de datos

También hemos seleccionado algunas bases de datos no españolas cuya consulta también puede ser una gran ayuda en la investigación.

Dissertation Abstracts

La temática de esta base de datos no se restringe a la lengua y lingüística, es multidisciplinar e internacional (la lengua de consulta es el inglés) pero se circunscribe al ámbito universitario. Contiene información sobre tesis doctorales y ponencias de masters presentada en universidades de todo el mundo (de 1861 a 1987 es información referencial y sólo cubre el ámbito norteamericano; a partir de 1988 cubre ampliamente el ámbito europeo y desde 1980 además de la información referencial adjunta un resumen). Se incrementa con más de 50.000 títulos. Tiene una actualización trimestral. Editada por The University Microfilms International (UMI) Ann Arbor.

⁹ Además de la Biblioteca Nacional podemos acceder a otras bibliotecas españolas a través de la red de Internet: Catálogos de Bibliotecas Públicas Españolas, Consejo Superior de Investigaciones Científicas, Red de bibliotecas, Fundesco, Bibliotecas Universitarias Españolas, Ruedo (catálogo colectivo de la red universitaria española dobis/libis), Bibliotecas Universitarias de Cataluña (catálogo colectivo vtls), UNED, etc.

Linguistics and Language Behavior Abstracts (LLBA)

En esta base de datos se recogen referencias bibliográficas con resumen de artículos aparecidos en más de 2.000 publicaciones periódicas, monografías recientes, tesis y literatura del ámbito temático de la lingüística, psicolingüística, sociolingüística y disciplinas relacionadas, desde 1973 a la actualidad, con una actualización semestral. Producida por Sociological Abstracts.

The Modern Language Association (MLA)

Las referencias bibliográficas (en inglés) de artículos sobre literatura, lenguas modernas, lingüística y folclore aparecidos desde 1963 hasta la actualidad, en más de 3.000 publicaciones periódicas especializadas, monografías, colecciones, actas de congresos, tesis, etc., constituyen el contenido de esta base de datos producida por The Modern Language Association y cuya actualización es trimestral.

2.3. Revistas

Aunque no están disponibles por el momento en formato electrónico, la *Revista de Filología Española (RFE)*, la *Nueva Revista de Filología Hispánica (NRFH)* y la *Revista de la Sociedad Española de Lingüística (RSEL)* tienen una sección dedicada a recopilar toda la bibliografía que ha aparecido desde la publicación del número anterior y están elaborando una página WEB. En la Red hemos encontrado varias revistas dedicadas a la literatura medieval y del Renacimiento. Especialmente interesantes para la investigación lingüística son *LEMIR*, *INTERLETRAS* e *INTERVISTA*.

INTERVISTA es una revista comercial que ofrece una curiosa reproducción de la edición facsimilar del Quijote. Está editada por Jaime Suances Torres.

LEMIR Es la revista de literatura medieval y renacentista de la Universidad de Valencia, dirigida por José Luis Canet, que incluye numerosas referencias a páginas literarias, especialmente medievales.

La revista *INTERLETRAS* (<http://unizar.es/gcorona/interlet.htm>), dirigida por Gonzalo Corona en la Universidad de Zaragoza está dedicada a la filología española, aunque no especializada en épocas o autores. Hasta la actualidad ha publicado tres números; el primero de ellos en mayo de 1996. Presenta un área de intercambio de textos electrónicos¹⁰.

¹⁰ Los textos disponibles y especialmente interesantes para la diacronía del español son: Berceo, Gonzalo de, *El milagro de Teófilo*, Calderón de la Barca, Pedro, *La vida es sueño*, *El médico de su honra*. Cervantes Saavedra, Miguel de, *Entremés de El hospital de los podridos*, *Entremés de El juez de los divorcios*, *Entremés de Los mirones*. Espronceda, José de, *El Estudiante de Salamanca*. Quevedo, Francisco de, *Selección*. Tirso de Molina (Gabriel Téllez), *El burlador de Sevilla*, *El condenado por desconfiado*. Vega y Carpio, Félix Lope de, *El Arte Nuevo de hacer comedias*, *El caballero de Olmedo*.

3. CORPORA

Los conjuntos homogéneos de documentos lingüísticos¹¹ de diferentes tipos que se toman como modelo de un estado o nivel de lengua anterior al actual son fruto de un lento y arduo trabajo. Su desarrollo viene motivado por las posibilidades de aplicabilidad y reutilización que ofrecen. La mayor parte de los dedicados a la diacronía del español aún no están disponibles.

3.1. Corpus de Español Medieval de la Universidad de Santiago de Compostela¹²

Se inició a principios de los años 80, cuando la informática aún no estaba al alcance de todos, pero su estructura, su tamaño, la organización de la información, el sistema de fichado, etc. hacen que haya sido un instrumento muy eficaz para la investigación en sintaxis diacrónica.

El material sintáctico que conforma el *corpus*, estructurado y organizado como una base de datos sintáctica, es el resultado del análisis funcional de la unidad cláusula entendida como unidad integradora del predicado y sus acompañantes, sean estos argumentales (sujeto, predicativo, complemento directo, complemento indirecto y complemento preposicional) o marginales (complementos circunstanciales y predicativos opcionales entre otros).

En esta base de datos no informatizada se analizan los constituyentes funcionales (tipo de unidad, categoría gramatical, etc.) de cada una de las cláusulas que componen el corpus, además de algunos rasgos que afectan globalmente a las mismas, como el tipo, la polaridad, la modalidad y el orden de constituyentes.

Las obras que componen el corpus son textos, no siempre completos, del español arcaico, alfonsí y del XIV, obras en prosa (*Primera Crónica General* y parte de *El Conde Lucanor*) y en verso (*Poema de Mio Cid*, *Libro de Alexandre*, *Fazienda de Ultramar*, *Vida de Santo Domingo de Silos*, *Vida de San Millán de la Cogolla*, *Libro de Apolonio*, *Milagros de Nuestra Señora*, *Poema de Fernán González* y *Libro de Buen Amor*) que, por su diversidad formal y temática, pueden aportar una información fiable acerca de la sintaxis de la etapa cronológica anterior a 1335¹³.

¹¹ Aunque no es un trabajo específico para el español medieval, los conceptos y características definidas en Alvar Ezquerro, M. y Villena Ponsoda, J. (1994) sobre corpora lingüísticos han sido tenidos en cuenta para la selección de corpora que incluimos aquí.

¹² Dirigido por el Dr. Jesús Pena Seijas y llevado a cabo por miembros del departamento de Filología Española, Teoría de la Literatura y Lingüística General de nuestra antigua Universidad de Santiago, así como por estudiantes de tercer ciclo.

¹³ Seguimos, para la datación de los textos, a Viña Liste (1991).

Distintos investigadores se han servido de este corpus (total o parcialmente) para elaborar tesinas, tesis doctorales, monografías y trabajos en revistas sobre los constituyentes de la cláusula¹⁴.

3.2. Corpus Diacrónico del Español (CORDE).

Está siendo elaborado por el Instituto de Lexicografía de la Real Academia Española con el objetivo de dar una visión representativa del español escrito (también del español de América) desde sus orígenes hasta 1975: Edad Media (hasta 1499), Siglos de Oro (1500-1699) Edad Moderna y Contemporánea (1700-1975).

Estiman que contará con un cómputo final entre 50 y 70 millones de palabras, anotadas morfosintácticamente, extraídas de 10.000 textos. Prescinden de las variantes textuales y de los manuscritos y trabajan sobre ediciones modernas de textos completos recogidos en formato electrónico (los textos más representativos de cada período) y ejemplos de textos.

La disponibilidad será progresiva a partir del segundo cuatrimestre de 1998, iniciándose con búsquedas de tipo léxico. Estará disponible con fines no comerciales para organismos públicos de investigación y prevén su acceso por Internet.

3.3. Archivo de textos hispánicos de la Universidad de Santiago (ARTHUS)

Es un corpus de lengua general (3.000.000 de palabras) que incluirá, además de lengua oral y escrita actual, textos de diferentes períodos, sobre todo medievales y clásicos. La parte contemporánea está compuesta de treinta y cuatro textos narrativos, teatrales, ensayísticos, periodísticos y orales procedentes de España e Hispanoamérica¹⁵.

Más lentamente se trabaja en la parte medieval y clásica. Por el momento, contiene trece textos con más de un millón de palabras. Las obras incluidas son: *Rimado de Palacio*, Las obras completas de D.Juan Manuel: *Libro del cauallero et del escudero*, *Libro de las armas*, *Libro enfenido*, *Libro de los estados*, *Tractado de la Asunción de la Virgen María*, *Libro de la caza*, *El Conde Lucanor*, *Crónica abreviada*, *Li-*

¹⁴ Sirvan de ejemplo los excelentes trabajos de Folgar, C, *Diacronía de los objetos directo e indirecto (del latín al castellano medieval)*. Santiago de Compostela: Universidad, 1993 (anejo 37 de *Verba*), Rivas, E, *Construcciones monoactanciales y transitivas biactanciales en castellano medieval: sus empleos alternativos*. Santiago de Compostela: Universidad, 1996 (anejo 41 de *Verba*) o Suárez Fernández, M. *El complemento predicativo en castellano medieval (Época Prealfonsí)* Santiago de Compostela: Universidad, 1996 (anejo 42 de *Verba*) entre otros.

¹⁵ Todas las cláusulas de este conjunto textual han sido fichadas y los datos correspondientes incorporados a la Base de datos sintácticos del español actual (BADSEA). Página web: <http://www.usc.es/~syntax/welcome/html>.

bro de buen amor, *Corbacho*, *Amadís de Gaula* y Obra poética de Francisco de Quevedo¹⁶. Ya se han fijado y probado los criterios de fichado y actualmente se procede al definitivo fichado de cláusulas que permitirá construir una base de datos sintáctica semejante a la del español actual.

3.4. Archivo informatizado de textos jurídicos medievales (AITJM)

La responsable de este corpus, no anotado, es Julia Ángeles Giménez Jurado del Seminario de Filología e Informática de la Universidad Autónoma de Barcelona. Según el *Informe sobre recursos lingüísticos para el español. Corpus escritos y orales disponibles en España y en desarrollo*, elaborado por el OEIL del Instituto Cervantes, pretende el análisis, principalmente léxico, de la lengua jurídica medieval y la determinación de la terminología legal del derecho hispánico. El contenido del corpus está formado exclusivamente por textos legales de la época medieval¹⁷.

3.5. Otros

En algunos trabajos sobre bases de datos, grupos de investigación en lingüística informática, etc. se citan otros corpora cuya adscripción a estadios anteriores al actual en español hacen que tales referencias se muestren especialmente interesantes para nosotros:

- *Corpus de prosa del XVI*, elaborado por el profesor Eugenio de Bustos, de la Universidad Complutense.
- *Corpus de la prosa histórica castellana* (1200-1800); Mark Davies, de la Illinois State University ha confeccionado este corpus que cuenta con 5.300.000 palabras y 114 obras.
- *Corpus del español medieval de la Universidad de Wisconsin*, con anotación morfológica y sintáctica.
- La mayor parte de los investigadores han optado por la creación de sus propios repertorios documentales. Bases de datos de diferentes tamaños, épocas, etc. para fines específicos. Ejemplo de ello es la base de datos de C. Sánchez Lancis para el estudio de elementos lingüísticos espacio-temporales (cfr. Sánchez Lancis, 1992).

¹⁶ <http://www.usc.es/quevd>, realizada y diseñada por Miguel A. Gómez Segade y Santiago Fernández Mosquera

¹⁷ No hemos conseguido información sobre disponibilidad, tamaño, etc.

4. DICCIONARIOS

Tanto los investigadores en lexicografía como los usuarios de diccionarios se benefician de la aplicación de los recursos informáticos. Las nuevas formas de organización y gestión de la información contenida en los diccionarios tradicionales junto con los programas para la elaboración de diccionarios son las dos muestras más evidentes de las ventajas que suponen la explotación y el aprovechamiento de las novedades disponibles. No podemos olvidarnos de la facilidad de consulta, de la versatilidad, ni de otras ventajas más prosaicas como el peso (hemos pasado de kilos a gramos), tamaño (hemos pasado de varios gruesos volúmenes a centímetros), precio, etc., pero sobre todo hemos logrado que la información contenida sea mejor aprovechada y que los diccionarios sean cada vez más consultados.

4.1. Sistema informatizado de elaboración de diccionarios

Para la diacronía del español la transformación en soporte electrónico de diccionarios existentes solo en algunas bibliotecas, así como la creación de otros nuevos diccionarios etimológicos e históricos (todavía no poseemos ningún diccionario histórico de la lengua española completo) supondrá una importante contribución al cambio que estamos presenciando. La creación de programas de gestión de diccionarios además permitirá acceder más fácilmente a la información contenida en ellos.

El grupo de Lexicografía de la Real Academia Española pretende desarrollar un conjunto de aplicaciones informáticas (a partir de la vigésima primera edición del DRAE) para la construcción de un sistema informatizado de elaboración de diccionarios cuyos componentes principales son: 1) bases de datos, 2) gestor de consultas, 3) sistemas de redacción y 4) herramientas de edición.

4.2. Informatización de diccionarios

La diacronía del español se beneficia de la informatización de diccionarios históricos, etimológicos, de variedades, etc., pero también de las posibilidades de recuperación de la información diacrónica contenida en los diccionarios del español actual.

4.2.1. Diccionarios etimológicos, históricos y glosarios

Diccionario Crítico Etimológico Castellano e Hispánico (DCECH).

La informatización del DCECH es una ardua tarea que los miembros del Seminario de Filología e Informática de la Universidad Autónoma de Barcelona llevan a cabo desde hace años. El proyecto de "Informatización del Diccionario Crítico Etimo-

lógico Castellano e Hispánico (DCECH)” de J. Corominas y J. A. Pascual¹⁸ no se ocupa simplemente de la informatización sino que “pretende la digitalización de parte de sus informaciones con el fin de tenerlas organizadas y sistematizadas en diversas bases de datos relacionales / interrelacionadas que permitirán consultas más complejas de las que se pueden hacer en cualquier diccionario, a la par que una ampliación, revisión, y reutilización continuas.” (Clavería, 1997: 1631). Los responsables del proyecto pretenden crear un corpus que sirva de base para complementar los datos que ofrece el DCECH. Se prevé su conclusión en el 2001.

Diccionario del Español Antiguo (DOSL)

El *Hispanic Seminary of Medieval Studies* de la Universidad de Madison, Wisconsin, trabaja desde hace más de cincuenta años en la elaboración del *Dictionary of the Old Spanish Language*, también denominado DOSL/DEA.

Form-Lex

Es un instrumento de recuperación de información, un diccionario interno de lemas y formas. Ha sido elaborado por el grupo de investigación *Laboratorio de Lingüística Informática*, dentro del proyecto ADMYTE con la colaboración del *Hispanic Seminary of Medieval Studies* de la Universidad de Madison, Wisconsin, EEUU. Está publicado en CD-ROM, por lo que el usuario ve facilitada la búsqueda de formas en los documentos. Contiene cerca de 47.000 formas y más de 11.000 lemas y puede ser consultado bien en su totalidad (número de veces y posiciones de la palabra *amor*, por ejemplo), bien de forma sesgada: por obras (la palabra *amor* en la *Vida de Santa María Egipciaca*), tipo de obras (la palabra *amor* en las enciclopedias, en las obras en prosa, en las obras anteriores a 1400, etc.).

4.2.2. *Diccionarios del español actual en CD-ROM*

La aparición de los diccionarios en CD-ROM ha supuesto un cambio radical en el modo de consulta. Las versiones en CD-ROM de los diccionarios tradicionales multiplican los resultados de las búsquedas porque gestionan todo el diccionario como texto de una base de datos permitiéndonos efectuar búsquedas no solo en las entradas del diccionario sino también en las definiciones. Podemos utilizar más criterios que el orden alfabético en el inicio de cada palabra e incluso combinar más de un criterio de búsqueda, etc., operaciones estas imposibles en los diccionarios en papel.

Diccionario de la Real Academia Española

La vigésima primera edición del DRAE se publicó en papel en 1992 y en formato CD-ROM en 1995. Salvo pequeñas modificaciones, ambas versiones mantienen la misma información, la diferencia radica en los recursos para obtener la información

¹⁸

La responsable del proyecto es Gloria Clavería de la Universidad Autónoma de Barcelona.

de los diccionarios. Con el CD-ROM podemos lograr una mayor y mejor accesibilidad a la información en ellos contenida, pudiendo consultar tanto la entrada como la definición.

La consulta de las entradas puede hacerse bien como en un diccionario tradicional, bien como en un diccionario inverso¹⁹ (palabras terminadas en *-eco*, por ejemplo). Además, permite la búsqueda de formas complejas (expresiones que contienen cualquiera de las palabras que componen una forma compleja) o incluso combinar dos parámetros, logrando así las llamadas búsquedas múltiples, que consisten en la combinación de elementos de los distintos modos de consulta: diccionario inverso, formas complejas, etc. mediante operadores lógicos (palabras acabadas en *-arro* siempre que en el árbol de categoría gramatical no sea un sustantivo, es un ejemplo de búsqueda). Podemos hacer consultas de los contenidos de las entradas y solicitar, por ejemplo, índices de palabras (definiciones en que aparece la palabra *alma*) con filtros sobre las definiciones, etimologías, etc.

La información recogida en cada una de las definiciones está clasificada y se puede acceder a ella por medio de los llamados "árboles". A través de la estructura de carpetas y subcarpetas se jerarquiza información sobre categoría gramatical, procedencia de los vocablos, usos y ámbito geográfico, etc. Se presentan los siguientes árboles: árbol de lenguas, de categoría gramatical, de usos: ámbito geográfico, de usos: materia y nivel y el árbol de opciones avanzadas, etc. Toda la información aparece estructurada y jerarquizada en carpetas y subcarpetas²⁰, además del número de apariciones de cada tipo y el número de marcas. Las aportaciones del DRAE al estudio de la historia de la lengua y de la gramática histórica han sido acrecentadas con la versión en CD-ROM.

Diccionario de uso del Español, M. MOLINER (1996).

El tratamiento informático del texto impreso en el *Diccionario de uso del español* presenta las mismas ventajas que la versión en CD-ROM del DRAE. La información es la misma que en la versión impresa (lema, anotaciones sobre la etimología, pronunciación, uso gramatical y lingüístico, presencia en el DRAE, referencias gramaticales, acepciones, sinónimos, familias de palabras, expresiones pluriverbales, etc.) pero las posibilidades de obtención de la información contenida se multiplican vertiginosamente.

¹⁹ El español ya contábamos con un diccionario inverso: Bosque, I. y Pérez Fernández M., *Diccionario inverso de la lengua española*, Madrid, Gredos, 1987.

²⁰ Sirva de ejemplo la consulta efectuada en el árbol de lenguas sobre el origen de los vocablos y el número de apariciones y marcas: 1.- *Lenguas originarias*: (44 apariciones / 6 marcas) 2.- *Castellano o español y sus dialectos*: (220 apariciones / 40 marcas) 3.- *Otras lenguas peninsulares* (106 apariciones / 5 marcas), 4.- *Árabe* (1.290 apariciones / 17 marcas), 5.- *Latín* (19.501 apariciones / 56 marcas), 6.- *Griego* (4.298 apariciones / 11 marcas), 7.- *Romances* (3.016 apariciones / 53 marcas), 8.- *Indígenas americanas* (903 apariciones / 32 marcas), 9.- *Otras* (1.173 apariciones / 112 marcas).

El CD-ROM del M. Moliner nos ofrece tres modos de consulta: 1) búsqueda en la lista de entradas (índice alfabético en el que se encuentran todas las entradas del diccionario), 2) búsqueda en las entradas que se dividen en principales y subentradas (derivadas de una entrada principal) y 3) búsqueda en las definiciones; además nos indica de forma automática el número de casos encontrados que contienen la palabra solicitada.

Además de la información contenida en la versión en papel la versión en CD-ROM nos ofrece sugerencias cuando se ha producido un error mecanográfico o de otro tipo, permite el uso de comodines de una letra o una serie de letras, anagramas, etc. También permite establecer orden de prioridades en las palabras que constituyen la búsqueda de expresiones, rango de proximidad entre las palabras de una expresión, combinaciones de palabras con operadores lógicos, etc. Las variaciones en la tipografía y los símbolos informan sobre el mayor / menor uso de una palabra, frases, modismos, información gramatical, etc.

El diccionario en CD-ROM puede ser consultado como una herramienta más de las últimas versiones de los procesadores de textos elidiendo, de esta forma, los problemas causados por la mala calidad de los diccionarios en ellos incluidos (Vid. Millás, 1997).

5. EDICIONES DE TEXTO EN FORMATO ELECTRÓNICO

Las técnicas de recuperación textual mediante el escáner, el reconocimiento de caracteres a través del OCR, el almacenado en soporte informático, la impresión digital, etc. son los primeros pasos en la edición de textos en formato electrónico (*vid.* García García, 1992). Estas herramientas, ya casi tradicionales e imprescindibles en la diacronía del español, eran hasta hace poco eran el único medio de disponer de textos en formato electrónico. Hoy día podemos consultar y manejar ediciones de textos del español medieval y clásico gracias a la generosidad de sus autores. Las ediciones de textos publicadas en las páginas WEB de algunas universidades y centros de investigación ponen a nuestra disposición textos en formato electrónico y de esta forma permiten que la creación de bases de datos, concordancias, etc. de esos textos sea una tarea mucho más fácil.

ADMYTE

Contiene títulos, la mayoría de ellos de incunables, aunque hay algunos impresos antiguos, todos ellos de la Biblioteca Nacional de España. El Archivo Digital de Manuscritos y Textos²¹ almacena textos e imágenes además de un conjunto de pro-

²¹ El comité de realización está compuesto por: Francisco Marcos Marín, Gerardo Meiro Martínez, Charles B. Faulhaber, Angel Gomez Moreno, Aurora Martín Santa Olalla, Julián Martín Abad, John J. Nití, Carmen Foz Baldiz, Félix Calero Sanchez, Cristóbal J. Florido García, Enrique de Lema Garzón, Luis Miguel Martínez Guiñales.

gramas que facilitan al usuario el acceso a los datos. Hasta la fecha han sido publicados dos volúmenes, pero el objetivo es publicar tantos como sea necesario para recoger toda la producción en castellano hasta 1615. Para la transcripción se han seguido los criterios de Nitti para la recopilación de textos del *Hispanic Seminary of Medieval Studies* en Madison, Wisconsin. Según sus autores incluye “lo que podría haber sido una biblioteca de un contemporáneo de los navegantes que partieron de Palos rumbo a las Indias” (Marcos Marín, 1992: 6). La intención ha sido la de presentar “Ese mundo del humanismo en el que el latín va dejando su paso al castellano, se muestra en las gramáticas de Nebrija, o sus diccionarios, de ambas lenguas, en las Partidas alfonsíes, base del ordenamiento jurídico español e hispanoamericano, en las ordenanzas promulgadas por los Reyes Católicos, en textos científicos que van desde la medicina a la veterinaria o la cosmografía, desde los autores árabes a los castellanos” (1992: Manual del ADMYTE en español).

El volumen 1 de ADMYTE, destinado al público en general pero también a investigadores, incluye la edición facsímil y transcripción a texto completo (en código ASCII, imágenes facsímiles en blanco y negro y miniaturas en color) de 61 incunables españoles de finales del XV. Incluye Enciclopedias²², Diccionarios y gramáticas²³, Textos legales²⁴, Textos científicos²⁵, Libros de viajes²⁶, Crónicas y biografías²⁷, Tra-

²² Alfonso de la Torre. *Visión delectable*. Burgos: Friedrich Biel, 1485 [?], 1486 [?] y Bartholomaeus Glanville. Y *De proprietatibus rerum. Propiedades de las cosas*. Tr. fray Vicente de Burgos. Toluca: Enrique Meyer.

²³ Alfonso de Palencia. *Universal vocabulario en latín y romance*. Sevilla: Pablo de Colonia, Juan Pegnitzer de Nuremberga, Magno Herbst de Fils & Tomás Glockner, 1490. Antonio de Nebrija. *Dictionarium hispano-latinum*. Salamanca: Impresor de la Gramática de Nebrija, 1495. Antonio de Nebrija. *Dictionarium latino-hispanicum*. Salamanca: Impresor de la Gramática de Nebrija, 1492. Antonio de Nebrija. *Gramática castellana*. Salamanca: Impresor de la Gramática de Nebrija, 1492-8-18. Antonio de Nebrija. *Introductiones latinae*. Salamanca: Impresor de las Introductiones latinae, 1481-1-16. Rodrigo Fernández de Santaella. *Vocabulario eclesiástico*. Sevilla: Juan Pegnitzer de Nuremberga, Magno Herbst de Fils & Tomás Glockner para Juan Lorenzo, 1499-2-14.

²⁴ Alfonso X. *Siete partidas*. Sevilla: Meinardo Ungut & Estanislao Polono, 1491-10-25. Alfonso Díaz de Montalvo. *Ordenanzas reales*. Huete: Alvaro de Castro, 1484-11-11. Fernando V & Isabel I. *Capítulos de gobernadores, asistentes y corregidores (1500-6-9)*. Sevilla: Juan Pegnitzer de Nuremberga & Magno Herbst de Fils o Estanislao Polono para García de la Torre & Alfonso Lorenzo, 1500-6-9 a quo. Fernando V & Isabel I. *Cuaderno de las leyes nuevas de la hermandad (1486-7-7)*. Sevilla: Jacobo Cromberger, c. 1511. Fernando V & Isabel I. *Leyes hechas por la brevedad y orden de los pleitos (1499-5-26)*. Salamanca: Impresor de la gramática de Nebrija para Fernando de Jaén, 1500?. Fernando V & Isabel I. *Leyes hechas por la brevedad y orden de los pleitos (1499-5-26)*. Salamanca: Juan de Porras, c. 1511. Fernando V & Isabel I. *Ordenanzas sobre los abogados (1499-2-11)*. Salamanca: Juan de Porras, c. 1511. Fernando V & Isabel I. *Ordenanzas de la ciudad de Sevilla (1492-5-30)*. Sevilla: Meinardo Ungut & Estanislao Polono, c. 1492. Fernando V & Isabel I. *Ordenanzas sobre los escribanos de la ciudad de Sevilla (1492-5-30)*. Sevilla: Meinardo Ungut & Estanislao Polono, c. 1492. Fernando V & Isabel I. *Ordenanzas sobre los paños (1500-9-15)*. Sevilla: Estanislao Polono para Niculoso de Monardi-s, 1500-11-26.

²⁵ Al-Mubashshir ibn Fatik. Mukhtar al-hikam, *Bocados de oro*. Sevilla: Meinardo Ungut & Estanislao Polono, 1495-5-16. Antonio de Nebrija. *Tabla de la diversidad de los días y horas*. Alcalá de Henares: Arnaldo Guillén de Brocar, c. 1516-1517?. Bernardus de Gordonio. *Lilium medicinae, Lilio de me-*

tados de caballería y nobleza²⁸, Textos poéticos²⁹.

dicina. ff. 1r-185v. Sevilla: Meinardo Ungut & Estanislao Polono, 1495-4-18. Cat. 315 (=I 2438). Bernardus de Gordonio. *De ingeniis curandorum morborum*. *Tratado de las tablas de los diez ingenios de curar enfermedades agudas*. ff. 186r-190v. Sevilla: Meinardo Ungut & Estanislao Polono, 1495-4-18. Cat. 315 (=I 2438). Bernardus de Gordonio. *De regimine acutorum morborum*, *Regimiento de las agudas*. ff. 190v-194r. Bernardus de Gordonio. *Tratado de los niños, con el regimiento del ama*. ff. 194v-198r. Bernardus de Gordonio. *Prognostica*, *Las pronósticas*. ff. 198v-229v. Fernando Alvarez. *Regimiento contra la peste*. Sevilla: Juan Pegnitzer de Nuremberga & Magno Herbst de Fils, c. 1500. Francisco López de Villalobos. *Sumario de la medicina*. Salamanca: Impresor de la de Nebrija para Antonio de Barreda, 1498. Johannes de Ketham. *Fasciculus medicinae*, *Compendio de la salud humana*. ff. 1-XXXIX. Zaragoza: Pablo Hurus, 1494-8-15. Vasco de Taranta. *De epidemia et peste*. *Tratado de la peste*. ff. 39v-48r. Zaragoza: Pablo Hurus, 1494-8-15. Anónimo. *Tratado de la fisonomía en breve suma contenida*. ff. 48v-63v. Julián Gutiérrez de Toledo. *Cura de la piedra y dolor de la ijada y cólica renal*. Toledo: Pedro Hagenbach para Melchor Gorricio, 1498-4-4. Manuel Díez de Calatayud. *Libre del arte de menescalia*, *Libro de albeitería*. Tr. Martín Martínez de Ampíes. Zaragoza: Pablo Hurus, 1499-10-16.

²⁶ [Jean de Bourgogne]. Juan de Mandevilla. *Libro de las maravillas del mundo y del viaje de la Tierra Santa*. Valencia, 1524-10-13. Martín Martínez de Ampíes. *Tratado de Roma*. ff. 4v-43v. Zaragoza: Pablo Hurus, 1498-1-16. Bernhard von Breydenbach. *Peregrinatio in Terram Sanctam*, *Viaje siquier peregrinación de la tierra sancta*. Tr. Martín Martínez de Ampíes. ff. 44r-199v. Pablo Hurus, 1498-1-16. Marco Polo. *Libro de Marco Polo*. Tr. Rodrigo Fernández de Santaella. Sevilla: Juan Varela de Salamanca, 1518-5-16.

²⁷ Diego Rodríguez de Almela. *Valerio de las historias eclesiásticas y de España*. Murcia: Lope de Roca, 1487-12-6. Diego Rodríguez de Almela. *Compilación de las batallas campales*. ff. 1r-48v. Diego Rodríguez de Almela. *Letra sobre los matrimonios y casamientos entre los reyes de Castilla y de León de España con los reyes y casa de Francia hechos*. ff. 49r-51r. Diego Rodríguez de Almela. *Escritura de cómo y por qué razón no se debe dividir, partir ni enajenar los reinos y señoríos de España*. ff. 51r-55v. Diego de Valera. *Crónica de España*. Sevilla: Alfonso del Puerto para Michael Dachauer & García de Castillo, 1482. Gauberto Fabricio de Vagad. *Crónica de Aragón*. Zaragoza: Pablo Hurus, 1499-9-12. Hernando del Pulgar. *Claros varones de Castilla*. ff. 1v-43v. *Letras*. ff. 43v-88v. Contiene: Hernando del Pulgar. *Letra de los males de la vejez*. ff. 43v-46r. *Para un caballero que fue desterrado del reino*. ff. 46r-48r. *Letra para don Alfonso Carrillo arzobispo de Toledo*. ff. 48r-49v. *Letra para un caballero su amigo de Toledo*. ff. 49v-50v. *Letra para el obispo de Osma*. ff. 50v-51v. *Letra para un caballero criado del arzobispo de Toledo*. ff. 51v-54r. *Letra para el rey de Portugal*. ff. 54r-58r. *Letra para el obispo de Tuy que estaba preso en Portugal*. ff. 58v-59v. *Letra para el doctor de Talavera*. ff. 59v-60r. *Letra para Don Enrique, tío del Rey, cuando le hirieron en Tajara*. f. 60r. *Letra para la reina nuestra señora*. ff. 60r-61r. *Letra para Pedro de Toledo, canónigo de Sevilla*. f. 61r-61v. *Letra para el condestable*. ff. 62r-63r. *Letra para un su amigo de Toledo*. ff. 63r-65v. *Letra para el Cardenal*. f. 65v. *Razonamiento hecho a la reina cuando hizo perdón general en Sevilla*. ff. 65v-68v. *Letra para el señor Don Enrique*. ff. 68v-69r. *Letra para el prior del Paso*. f. *Letra para el conde de Cifuentes que estaba preso en Granada*. ff. 69v-70r. *Letra para don Ñigo de Mendoza, conde de Tendilla*. f. 70r-v. *Letra para un su amigo encubierto*. ff. 70v-72r. *Letra para don Gabriel de Mendoza*. ff. 72r-72v. *Letra para una su hija monja*. ff. 72v-80v. *Letra aconsejando a un señor en el casamiento de su hija*. ff. 80v-81v. *Letra para el obispo de Coria, deán de Toledo*. ff. 81v-84v. *Letra para Fernán Alvarez, tesorero de la reina*. ff. 84v-85v. *Letra para el maestro de la capilla del rey de Portugal*. ff. 85v-86r. *Otra letra para el prior del Paso*. f. 86r-v. *Letra para mosén Alfonso de Olivares, que estaba en la compañía del duque de Plasencia*. ff. 86v-87r. *Letra para Puerto-carrero, el señor de Palma*. f. 87r-v. *Otra letra para el cardenal de España*. ff. 87v-88r. *Letra para el señor don Enrique*. f. 88r-v. Anónimo. *A la reina nuestra señora*. f. 91r. *Crónica popular del Cid*. Burgos: Friedrich Biel para San Pedro de Cardeña., 1512-3-31. R 897.

²⁸ Alonso de Cartagena. *Doctrinal de los caballeros*. Burgos: Friedrich Biel, 1487-6-20. Diego de Valera. *Tratado de las armas*, *Tratado de los rieptos y desafíos*. Valencia: Juan Viñao, c. 1517? R 2302.

TEXTOS - MAD

Colección de textos medievales, cedidos por el *Hispanic Seminary of Medieval Studies*, elaborado por el grupo de investigación *Laboratorio de Lingüística Informática* de la Universidad Autónoma de Madrid, dentro del proyecto ADMYTE (vol. 0) con la colaboración de la Universidad de Madison.

Biblioteca Altera de la Universidad de Salamanca

El objetivo es rescatar mediante impresión digital todos los libros descatalogados de su fondo editorial. Los ámbitos son la historia, el derecho, la filología, la filosofía, el arte y la medicina. Hasta la fecha son 101 ejemplares los que constituyen esta biblioteca, entre los que destacan *El corpus documental de Carlos V*, *El símbolo de la noche de San Juan de la Cruz*, *La poesía de R. Alberti*, *El estudiante de Salamanca*, *El fuero de Béjar*, etc.

Textos medievales y renacentistas de la página de LEMIR³⁰

La página de LEMIR tiene la generosa intención de “poner al alcance de los investigadores, profesores y alumnos de Literatura Española Medieval y del Renacimiento un material específico en su área, en un soporte electrónico y de amplia difusión” (<http://www.uv.es/lemir>). Han incluido en la página WEB una sección de textos que se irá completando con otros textos. Hasta el momento han puesto a nuestra disposición *Comedia llamada Rosabella*, de Martín de Santander, 1550, *La Tebaida de Estacio*, traducida por Juan de Arjona (fines XVI), *La Farsa llamada Alarquina*, Anónimo (mitad del XVI), *La Comedia Poliscena*, de Leonardo Francesco Bruni (II Aretino), *Las siete edades del mundo*, de Pablo de Santa María, *Las Mujeres ilustres*, de Jovanni Boccaccio en forma de diccionario, *Tractado del uso de las mugeres*, de Francisco Núñez de Coria (Oria), 1572.

ff. a2r-c8v (2r-24v.). Diego de Valera. *Ceremonial de príncipes*. ff. d1r-d7v (25r-31v). Fernando Mejía. *Nobiliario vero*. Sevilla: Pedro Brun & Juan Gentil, 1492-6-30.

²⁹ Íñigo López de Mendoza. *Bias contra Fortuna*. Sevilla: Estanislao Polono, 1502-4-3. R 11.886. Juan del Encina. *Coplas del memento homo*. S.l.n.a. [c. 1520-1550] R 3642. ¿Burgos: Juan de Junta?, 1531 ad quem, R 3642 ff. 1va-3rb. Ambrosio Montesino. *Coplas hechas sobre la Pasión*. ff. 3rb-4rb. S.l.n.a. [c. 1520-1550] R 3642. ¿Burgos: Juan de Junta?, 1531 ad quem, R 3642. Juan del Encina. *Villancico nuevo sobre el perdón de Santa María Magdalena*. f. S.l.n.a. [c. 1520-1550] R 3642. ¿Burgos: Juan de Junta?, 1531 ad quem, R 3642. 4va-b. Juan del Encina. *Egloga de Fileno, Zambardo y Cardonio*. S.l.n.a. [c. 1520-1550] R 4993. Juan del Encina. *Egloga de Plácida y Victoriano*. Burgos: Alonso de Melgar, c. 1519. R 4888. ff. a1v-c2va (1va-18va). [Cartagena]. “*Es amor en quien se esfuerza*”. f. c2va (18va). Fernando López de Yanguas. *Nunc dimittis*. ff. c2v-c4vb (19v-20vb). Juan del Encina. *Testamento de amores*. ff. a1r-a4rb (1r-4rb). Juan del Encina. *A su amiga porque se le escondia en viéndola*. f. a4rb-va (4rb-va). [Juan del Encina?]. “*Todos servid al amor*”. f. a4vb (4vb). Anónimo. *Coplas hechas sobre el casamiento de la hija del Rey de España*. Burgos: Friedrich Biel, c. 1496-7?

³⁰ El director es José Luis Canet y el comité de redacción está formado por los profesores Rafael Beltrán, Marta Haro, Evangelina Rodríguez y Josep Lluís Sirera.

Página de la Asociación Internacional de Teatro Clásico Español.

Contiene textos de obras nuestros autores del Siglo de Oro accesibles directamente.

Grupo de investigación del Siglo de Oro (GRISO).

En la página del grupo de Investigación del Siglo de Oro, del Departamento de Literatura Hispánica de la Universidad de Navarra, fundado por Ignacio Arellano (<http://griso.cti.unav.es>), se presentan textos de Quevedo³¹, Calderón³², La Comedia burlesca³³, Tirso de Molina, etc.

³¹ Quevedo, *Poemas de la Musa Clío* Texto fijado por I. Arellano sobre la edición de *El parnaso español*, de González de Salas, Madrid, Diego Díaz de la Carrera, a costa de Pedro Coello, 1648. *Poesía selecta*, Texto tomado de la edición de Lia Schwartz Lerner e Ignacio Arellano, Barcelona, PPU, 1989, *Poemas amorosos, Poemas metafísicos, Poemas satíricos, Orlando enamorado, Prosa, El buscón, historia de la vida del buscón llamado don Pablos, ejemplo de vagamundos y espejo de tacaños*, por Don Francisco de Quevedo Villegas, Edición de Celsa Carmen García Valdés (Texto base: Manuscrito de Don José Bueno), *El gran tacaño, Tercera parte de la vida del gran tacaño*, por Vicente Alemany, Anejos de la Perinola II, Pamplona, Eunsa, 1997, *La cuna y la sepultura para el conocimiento propio, y desengaño de las cosas ajenas*, Año 1634, Edición de Celsa Carmen García Valdés, (Texto base: Ejemplar R. III - F 788, de la Biblioteca Menéndez Pelayo de Santander). *Memorial patronato de santiago, y por todos los santos naturales de españa, en favor de la elección de cristo nuestro señor*, escribete D. Francisco de Quevedo Villegas, caballero del hábito de Santiago, con licencia, en Madrid, por la viuda de Alonso Martín. Año 1628., Ed. Celsa Carmen García Valdés (Según el ejemplar R-17431 de la Biblioteca Nacional de Madrid). *Opúsculos festivos, Sueños y discursos*. Texto fijado por I. Arellano sobre la primera edición impresa: *Sueños y discursos de verdades descubridoras de abusos, vicios y engaños en todos los oficios y estados del mundo*, Barcelona, Esteban Liberos Acosta de Juan Saperá, 1627. *Versión de Juguetes de la niñez*: Texto fijado por I. Arellano sobre la edición de *Juguetes de la niñez y travesuras del ingenio*, Madrid, Viuda de Alonso Martín, 1631. *La Tebaida de Publio Papinio Estacio* traducida por Juan de Arjona Libro I, Edición realizada por Pere-Enric Barreda (Universitat de Barcelona) sobre el texto publicado en 1888, con las variantes de los manuscritos y la correspondencia con el verso del original latino.

³² Los textos que ponen a disposición en esta página proceden de las ediciones críticas publicadas por Reichenberger-Universidad de Navarra dentro del proyecto de edición y publicación de los autos completos calderonianos. Presentan diez autos en edición electrónica. Algunos incluyen las notas críticas correspondientes, otros aparecerán sin notas. Para éstas y para los estudios literarios y textuales, nos remiten a las ediciones en papel o a la próxima edición en CD-ROM que están preparando. También nos ofrecen las concordancias, es decir, desde esta página es posible localizar las apariciones de cualquier palabra o grupo de palabras en los diez primeros autos publicados. Por el momento podemos encontrar una muestra del programa que están preparando y que en breve publicarán; mayor número de autos, posibilidades de búsqueda mucho más amplias, etc. Los textos de Calderón disponibles son: *Andrómeda y Perseo*, Edición crítica de José M. Ruano de la Haza, Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1995. Se incluyen las notas pero no el estudio preliminar, índices, etc. *El año santo de Roma*, Edición crítica de Ignacio Arellano y Ángel L. Cilveti Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1995. No incluye notas ni el estudio preliminar. *El cordero de Isaías*, Edición crítica de M. C. Pinillos, Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1996. Se incluyen las notas pero no el estudio preliminar, índices, etc. *El divino Jasón*, Edición crítica de Ignacio Arellano y Ángel L. Cilveti, Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1992. Se incluyen las notas pero no el estudio preliminar, índices, etc. *El indulto general*, Edición crítica de Ignacio Arellano y J. Manuel Escudero Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1996. Se incluyen las notas pero no el estudio preliminar, índices, etc. *La nave del mercader*, Edición crítica de Ignacio Arellano, Blanca Oteiza, M. Carmen Pinillos, J. Manuel

6. OTRAS HERRAMIENTAS

Incluimos en este apartado algunas de las herramientas más útiles que conocemos para la investigación en la diacronía del español.

Volumen 0 del proyecto ADMYTE

El volumen 0 del proyecto ADMYTE es un disco instrumental destinado a investigadores. Presenta dos programas especialmente útiles para el trabajo en diacronía: *Unite* y *Tact* (también disponible y de acceso libre en la red de Internet), además de incunables e impresos antiguos.

Unite

Conjunto de programas para la construcción de ediciones críticas automatizadas. Elaborado por el grupo de investigación *Laboratorio de Lingüística Informática*, publicado en CD-ROM, dentro del proyecto ADMYTE con la colaboración de la Universidad Autónoma de Madrid y la Alexander von Humboldt Stiftung de Alemania.

Tact

Programa de recuperación de información textual con sistema de creación de la propia base de datos textual, producido por el *Centre for Computing in the Humanities* de la Universidad de Toronto (Canadá). Permite conocer la ubicación de las palabras y sintagmas de una base de datos textual creada a partir de un texto por medio de una de las utilidades (MAKEBASE) que el propio programa posee. Podemos, fácilmente, ver el texto, extraer la lista de palabras, seleccionar una o varias para construir concordancias o índices referenciales, hacer recuentos para estudiar la distribución, etc. Una de las opciones del programa permite crear bases de datos a partir de textos transcritos con las normas HSMS de Madison.

Escudero y Ana Armendariz, Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1996. *No hay instante sin milagro* Edición crítica de Ignacio Arellano, Ildefonso Adeva y Rafael Zafra, Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1995. *El nuevo hospicio*, Edición crítica de Ignacio Arellano, Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1995. *Triunfar muriendo* Edición crítica de Víctor García Ruiz, Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1996. Se incluyen las notas pero no el estudio preliminar, índices, etc. *La viña del Señor* Edición crítica de Ignacio Arellano, Ángel Cilveti, Blanca Oteiza, M. Carmen Pinillos. Pamplona-Kassel, Universidad de Navarra-Reichenberger, 1996. Incluyen las notas pero no el estudio preliminar, índices, etc.

³³ En esta página podemos encontrar bibliografía sobre la comedia burlesca, un banco de textos en el que ponen a disposición de la comunidad internacional los textos de las obras editadas por los miembros del GRISO. Las obras que ya están en formato electrónico son: *Mañanas de abril y mayo*, Edición crítica de Ignacio Arellano Ayuso, Pamplona-Toulouse, Anejos de Criticón, Presses Universitaires du Mirail, 1995. El texto no incluye notas ni el estudio preliminar. *El rey Don Alfonso el de la mano horadada*, Edición crítica de Carlos Mata Induráin, En prensa, Eunsa. El texto no incluye notas ni el estudio preliminar, para los que remitimos a la publicación. Anónimo *La ventura sin buscarla*, Edición crítica de GRISO Pamplona, Anejos Rilce, Eunsa, 1992, El texto no incluye notas ni el estudio preliminar.

Transcalc

Consiste en un programa elaborado por Roman Capsada y Joan Torruella (1995). Dicho programa crea un vínculo informático entre texto, concordancia, y base de datos permitiendo obtener, “a partir de las concordancias de un texto antiguo, una base de datos de sus formas léxicas de manera semiautomática” (Clavería y Sánchez, 1997: 141)

Listas de discusión:

Nos permiten intercambiar opiniones a través de la red por medio de mensajes de correo electrónico. Los temas de discusión son variadísimos. Una lista de listas aparece publicada en Marcos Marín (1994: 717-738), pero también podemos extraer la información directamente de ftp anónimo (<ftp://nisc.sri.com>). Sirva de ejemplo:

INFOLING-L

Es una lista moderada de lingüística española editada por Eulalia Bobes y C. Subirats Rueggeberg, de la Universidad Autónoma de Barcelona, Mar Cruz y Emma Martinell, de la Universidad de Barcelona, con la colaboración de Paola Bentivoglio y Mercedes Sedano, de la Universidad Central de Caracas, Venezuela.

Páginas Web

Pocas son las universidades españolas que, en la actualidad, no están presentes en la red Internet a través de las páginas WEB. Generalmente, los contenidos están relacionados con cuestiones administrativas pero también aparecen descripciones de la estructura de los departamentos, miembros, direcciones de correo electrónico, etc. Algunas cuentan con una sección de novedades para informar sobre congresos, proyectos, etc. pero las cuestiones más específicamente académicas, como los programas de las asignaturas, los contenidos de los cursos de doctorado, etc. en pocos casos están disponibles en este nuevo formato. Las dos páginas que hemos seleccionado, aunque diferentes en sus contenidos y objetivos, son especialmente interesantes para el estudio de la diacronía del español y la gramática histórica.

La Universidad Nacional de Educación a Distancia (<http://www.uned.es>)

Con fines docentes básicamente, la sección de Historia de la Lengua y Gramática Histórica de la UNED ha elaborado una utilísima página WEB. Las propias características del centro hicieron que fuese la UNED una de las primeras universidades españolas en publicar los llamados “libros del alumno” conteniendo los programas de las asignaturas, etc. y nuevamente vuelve a ser innovadora en las modernas redes de comunicación. Los programas de las asignaturas publicados en la red Internet, además de ser consultados por los alumnos, a quien en principio van dirigidos, pueden ser consultados por todos los interesados en la materia, pudiendo acceder fácilmente a los contenidos de los temarios de las materias de Historia del Español y Gramática Histó-

rica y conocer la dirección de correo electrónico del profesor o profesores responsables de la asignatura, etc.

La asignatura de Historia de la Lengua Española, del Departamento de Lengua Española y Lingüística General, además de las utilidades habituales para los alumnos cuenta con apartados dedicados a explicar el programa, con un riguroso y detallado índice temático, objetivos de la asignatura, metodología, criterios de evaluación, lecturas obligatorias, bibliografía, etc., lo cual resulta muy útil para los alumnos, pero no menos ventajoso para los profesores e investigadores de esas asignaturas en otras universidades. Intercambiar, colaborar, es ahora una tarea mucho más fácil.

Como ya hemos dicho anteriormente, además de los programas, en el servidor de la UNED estaba alojado el *Boletín Bibliográfico de Historia de la Lengua Española*. (cfr. *supra* *Boletín Bibliográfico de Historia de la Lengua Española en Internet*).

Página web de Literatura Española Medieval y del Renacimiento (<http://www.uv.es/lemir>)

En el apartado dedicado a la información general de esta WEB aparecen claramente expuestos los objetivos y las aportaciones de la misma. Los reproducimos a continuación como muestra de la generosidad que impera en parte de la RED:

Este WWW ha sido creado con la intención de ser un instrumento más para los investigadores de todo el mundo en Literatura española Medieval y del Renacimiento. En él se podrá encontrar noticias referentes a la especialidad, tanto de cursos, seminarios, planes de estudio o cualquier noticia de interés que sea relevante; así como el acceso a Catálogos de Centros de Documentación, de las Bibliotecas Españolas y Extranjeras, a los resúmenes del Boletín Oficial del Estado y del Diari Oficial de la Generalitat Valenciana. También se ha pensado en crear una Revista electrónica especializada, y un apartado donde editar los Textos Medievales y del Renacimiento, difícilmente publicables en editoriales comerciales. Finalmente se han incorporado secciones temáticas, como el Tirant y Exemplaria ([HTTP://uv.es/lemir](http://uv.es/lemir)).

Además solicitan la colaboración de los “visitantes” que puedan tener información sobre textos, revistas, noticias, congresos, cursos, etc. Ofrecen también una dirección de correo electrónico a la que dirigirse: Lemir@www.uv.es.

REFERENCIAS BIBLIOGRÁFICAS

- ALVAR EZQUERRA, M. & J. VILLENA PONSODA, (1994): *Estudios para un corpus del español*, Málaga, Universidad de Málaga, 1994.
- BOSQUE, I. & M. PÉREZ FERNÁNDEZ (1987): *Diccionario inverso de la lengua española*, Madrid, Gredos, 1987.
- CAPSADA, R. & J. TORRUELLA (1995): *Del manuscrit a la base de dades*, Bellaterra, UAB, 1995.
- CLAVERÍA NADAL, G. (1993): "La informatización lexicográfica en el Diccionario crítico etimológico castellano e hispánico (DCECH) de J. Corominas y J. A. Pascual" en *Actes du XXe Congrès International de Linguistique et Philologie Romanes*, Université de Zürich, 6-11 avril 1992, publiées par G. Hilty, Tübingen, Francke Verlag, 1993, tome IV, pp. 591-604.
- CLAVERÍA NADAL, G., BLECUA, J. M., SÁNCHEZ, C. & J. TORUUELLA, (1995): *Bases para la informatización del Diccionario crítico etimológico castellano e hispánico (DCECH) de J. Corominas y J. A. Pascual*. Bellaterra, UAB, 1995.
- CLAVERÍA NADAL, G. & C. SÁNCHEZ (1997): "La aplicación de las bases de datos al estudio histórico del español", en *Hispania*, 80/1, marzo 1997, 142-152.
- CLAVERÍA NADAL, G., SÁNCHEZ, C. & J. TORUUELLA (1997): "Sobre la informatización del Diccionario crítico etimológico castellano e hispánico (DCECH) de J. Corominas y J. A. Pascual" en *Actas del III Congreso Internacional de la Historia de la Lengua Española*, A. Alonso González, L. Castro Ramos, B. Gutiérrez Rodilla & J. A. Pascual Rodríguez (eds.), Madrid, La Muralla, 1997.
- FRADEJAS RUEDA J., DíEZ DE REVENGA TORRES, P. & I. PÉREZ PASCUAL, (1997): *Boletín Bibliográfico de Historia de la Lengua Española* en versión electrónica hospedado en el servidor de la UNED.
- (1994): *Boletín Bibliográfico de Historia de la Lengua Española*, Asociación de Historia de la Lengua Española, Arco / Libros, 1994.
- GARCÍA GARCÍA M., (1992): "El uso del escáner como ayuda en la lectura de manuscritos antiguos" en *Actas del III Congreso Internacional de Historia de la Lengua Española*, 1645-58. A. Alonso González, L. Castro Ramos, B. Gutiérrez Rodilla & J. A. Pascual Rodríguez (eds.), Madrid, La Muralla, 1997.
- MARCOS MARÍN, F. (1994): *Informática y humanidades*, Madrid, Gredos, 1994.
- (1996): *El comentario filológico con apoyo informático*, Madrid, Síntesis, 1996.
- MILLÁS, J., (1997): "El procesador" en *El País*, 19/IX/97.
- MOLINER, M., (1996) : *Diccionario de uso del Español*, Madrid, Gredos, CD-ROM.
- NARBONA JIMÉNEZ, A., (1984): "Para un repertorio bibliográfico básico de sintaxis histórica del español" en *Alfinge*, Universidad de Córdoba, nº 2, 321-358, 1984.
- (1985): "Para un repertorio bibliográfico básico de sintaxis histórica del español (Addenda)" en *Alfinge*, Universidad de Córdoba, nº 3, 277-279, 1985.

- OEIL, Observatorio Español de Industrias de la Lengua, (1996): *Informe sobre recursos lingüísticos para el español (II). Corpus escritos y orales disponibles y en desarrollo en España*. Instituto Cervantes, Alcalá de Henares, 1996.
- REAL ACADEMIA ESPAÑOLA, (1995): *Diccionario de la lengua Española*, Madrid, Espasa-Calpe, CD-ROM.
- ROJO G. (1993): "Las bases de datos sintácticos del español actual", en *Español Actual*, 59, 1993, pp. 15-20
- SÁNCHEZ LANCIS, C. (1992): *Estudio de los adverbios de espacio y tiempo en español medieval*. Bellaterra, UAB.
- SEPLN y OEIL (Sociedad Española para el Procesamiento del Lenguaje Natural y Observatorio Español de Industrias de la Lengua), (1997): *Grupos de investigación en procesamiento del lenguaje y del habla en España 1997*.
- VIÑA LISTE, J. M. (1991): *Cronología de la literatura española, I: Edad Media*, Madrid, Cátedra, 1991.
- WANNER, D. (1991): "Historical Syntax and Old Spanish Text Files", *Linguistic Studies in Medieval Spanish*. Eds. R. Harris-Northall y Th. D. Cravens. Madison: The Hispanic Seminary of Medieval Studies, 1991, 166-90.

8.- FUENTES ELECTRÓNICAS

- Bases de datos del ministerio de educación y ciencia: <http://mec.es/isbn>, <http://mcu.es/pic/spain/barc.html>, <http://mcu.es/pic/spain/edit.html>, <http://mec.es/teseo>.
- Biblioteca de la Universidad de Columbia: <http://www.columbia.edu/cu/libraries>.
- Biblioteca Nacional: <http://www.bn.es>.
- Grupo de investigación del Siglo de Oro: <http://griso.cti.unav.es>.
- Instituto Cervantes: <http://www.cervantes>.
- Instituto de Lingüística informática: <http://www.llif.uam.es>.
- Interletras: <http://unizar.es/gcorona/interlet.htm>.
- Literatura Española medieval y renacentista: <http://www.uv.es/lemir>.
- Sociedad Española para el procesamiento del lenguaje natural: <http://sisx04.si.ehu.lsi/Activities/SEPLN>.
- Universidad Autónoma de Barcelona: <http://cc.uab.es>.
- Universidad de Santiago de Compostela: <http://www.usc.es>.
- Universidad Jaume I: <http://www.uji.es>.
- Universidad Nacional de educación a distancia: <http://www.uned.es>.