Un ed. Radio Educativa. Hola, buenas tardes. Un saludo desde la Universidad Nacional de Educación a Distancia. A partir de ahora y durante dos horas y media, la UNED les va a acompañar con un programa educativo y cultural que pueden escuchar a través de Radio 3, desde Radio 3.

Nacional de España en FM y también en el 1359 de la Onda Media. Es este un espacio realizado por el CEMAP, Centro de Diseño y Producción de Medios Audiovisuales, en un intento de acercar la universidad a todos nuestros alumnos y, cómo no, a todos aquellos interesados en ampliar su horizonte cultural. Hoy, lunes 4 de mayo, hemos preparado una serie de espacios relacionados con las facultades de Filología, Geografía, Historia y Ciencias. Además, contamos con el Espacio Cultural Tiempo de Músicas y el curso de acceso directo a la universidad para mayores de 25 años, que cerrará, como cada día, nuestra emisión a las once y media de la noche, hora peninsular. Pues bien, si están dispuestos a permanecer en nuestra compañía, vamos a dar paso ya al contenido educativo. Comenzamos con Aula de Filología.

Música Aula de Filología es un espacio de unos 30 minutos de duración aproximadamente que cada lunes dedicamos a tratar en profundidad un tema relacionado con algunas de las materias que forman parte del plan de estudios de Filología Española en la UNED. El espacio de esta noche está dedicado a lengua española, concretamente a Historia de la Lengua Española I. José Manuel Fradejas Rueda, profesor titular de esta asignatura en la Facultad de Filología Española,

Es el responsable de este espacio dedicado a proporcionar algunas notas de lexicografía histórica. Pero les dejo ya con el profesor José Manuel Fradejas. Buenas noches. De nuevo estoy ante ustedes para presentarles uno de los pocos programas que hacemos en Historia de la Lengua Española I a lo largo del curso académico. Como ya es costumbre en esta asignatura, el segundo programa no lo hago yo, sino un profesor de otra universidad o un especialista de otra parte del globo terráqueo, ya que nosotros estamos en todas partes. Hoy tenemos de nuevo al profesor Juan Carlos Conde. Durante años ha trabajado en la Real Academia Española. Aún sigue ahí, pero también ahora es profesor de la Universidad Complutense. Y de nuevo nos va a hablar sobre diccionarios, diccionarios históricos, diccionarios... Mejor le dejo a él que nos lo cuente. Muchas gracias, profesor Fradejas.

Es mi intención aprovechar esta oportunidad que tan gentilmente me brinda el profesor José Manuel Fradejas para intentar ofrecer alguna información útil acerca de los recursos y herramientas que hoy tiene a su alcance el estudioso de la historia de la lengua española interesado preferentemente en los problemas que suscita el léxico del español en su proyección diacrónica. A decir verdad no me ocuparé aquí en términos absolutos de todos los recursos de que puede disponer al interesado ya que la mayor parte de ellos son sobradamente conocidos por todos y en algunos casos verdaderos hitos de la lexicografía románica. Piénsese en los casos de dos recursos tan imprescindibles y conocidos como por ejemplo el diccionario de autoridades o el diccionario crítico etimológico castellano e hispánico de Joan Coruminas y José Antonio Pascual. En esta charla prestaré atención a las aportaciones relevantes producidas en este curso. En este campo en los últimos años y sobre todo intentaré vislumbrar cuáles son las perspectivas que aguardan en el futuro próximo haciendo constar al mismo tiempo algunas de las necesidades que más perentoriamente urge cubrir.

en esta parcela de la investigación filológica. Revisaremos el panorama de las aportaciones de interés para nuestros propósitos en varios niveles. En primer lugar, los avances en el ámbito de las aportaciones de índole restringida, con lo que hago referencia a monografías en que se estudian aspectos parciales relacionados con la historia del léxico español, artículos en que se indaga la procedencia, historia y significado de una determinada palabra o familia de palabras, por ejemplo, o contribuciones tales como glosarios o vocabularios que acompañan a la edición de textos literarios, jurídicos o históricos. En segundo lugar, me ocuparé de la publicación de diccionarios de orientación diacrónica no integral, es decir, circunscritos a un área temática, zona geográfica o periodo de tiempo restringidos. En tercer lugar, trataré de la situación de un proyecto cuya necesidad sigue siendo un deseo insatisfecho, la redacción de un diccionario histórico del español. En cuarto y último lugar, se dirá algo sobre las aportaciones realizadas a partir de la utilización

las tecnologías informáticas en la investigación sobre el léxico de la lengua española. Por lo que respecta al primer nivel de los mencionados, es indudable que en los últimos años han sido muy numerosas las aportaciones realizadas en el ámbito de las revistas especializadas, y es preciso mencionar dos referencias tan recientes como necesarias por su dedicación específica al ámbito que nos ocupa, la revista Voces, coedición entre las universidades de Caen y Salamanca, y la revista de lexicografía publicada por la Universidad de La Coruña. Se continúa con ese caudal de monografías una tradición sólidamente establecida en el campo de la historia de la lengua española, en el que brillaron y todavía brillan grandes nombres por suerte aún activos, como los de Rafael Lapesa, Marguerita Morreale, Manuel Alvar, Jacob Malquiel y Germán Colón, por poner algunos ilustres ejemplos, y a los que se suman nombres que continúan dignamente la tradición, como los de Stephen Dworkin, Pedro Álvarez de Miranda, Ray Harris Northall o, y no porque esté a mi lado, José Manuel Fradejas. No debe tampoco olvidar el interesado en el estudio de la historia de la lengua española, el léxico del español, la producción.

intensa en los últimos años de ediciones de textos españoles antiguos, y empleo la expresión en un sentido amplio textos anteriores a 1800, que incorporan glosarios, vocabularios o índices de voces comentadas en las notas. En algunos casos los editores de los textos han llevado a cabo un esfuerzo de investigación léxica del mayor interés. Para el rastreo de este tipo de materiales pertenecientes al ámbito del español medieval es un recurso importantísimo la bibliografía compilada por Stephen Durkin y David Billick, Lexical Studies of Medieval Spanish Texts, Madison 1993, cuya tercera edición se espera con impaciencia. Sería muy deseable una guía bibliográfica de similares características que se ocupara de inventariar exhaustivamente artículos y notas dedicados al estudio monográfico de voces o familias de voces a lo largo de todas las épocas y zonas del español. Es sabido que en ocasiones controlar satisfactoriamente la avalancha de información procedente de todo el mundo resulta sumamente difícil. Por lo que respecta al segundo

de los casos planteados, el de los diccionarios de orientación diacrónica no integral, es decir, circunscritos a un área temática, zona geográfica o periodo de tiempo restringidos, en los últimos años ha habido alguna novedad de gran relieve, pero se siguen sintiendo algunos vacíos notables. En primer lugar, es preciso saludar la publicación del Diccionario Español de Textos Médicos Antiguos, elaborado por un equipo dirigido por la profesora de la Universidad de Salamanca, María Teresa Herrera, y publicado en Madrid por la editorial

Arcolibros en 1996. En sus casi 1.700 páginas en folio, este diccionario ofrece el caudal léxico presente en 32 textos médicos medievales españoles, pulcramente transcritos ad hoc para el proyecto, acompañándolo de definiciones precisas y rigurosas, y lo que es importante, de la documentación textual sobre la que se construye el artículo. Es de notar que este diccionario no recoge solamente el léxico técnico de la medicina o farmacología, sino que también el artículo de la medicina o farmacología, copia medievales, cuyo grado de complejidad por sí mismo justificaría su redacción, sino todo el léxico.

recogido en los 32 textos en los que se basa, por lo que con frecuencia rinde beneficiosos servicios al historiador de la lengua, no directamente interesado por el mundo de los textos médicos. Es destacable señalar que este diccionario se ha llevado a cabo mediante el empleo de la informática, es decir, a partir de la construcción de un corpus textual en soporte informático, luego nos ocuparemos de esto, y que tanto la metodología empleada en la transcripción de los textos utilizados, como el entorno informático de extracción de información, son los desarrollados en el Seminario de Estudios Hispánicos Medievales de la Universidad de Madison, asuntos a los que luego haremos referencia. La publicación de una magnífica reseña de este Diccionario Español de Textos Médicos Antiguos por el profesor Ray Harris Northall, en el número 50 de la revista Romans Philology, me exime de realizar más comentarios, salvo el de consignar la magnitud y la importancia de esta obra para el estudioso de la historia de la lengua española. En segundo lugar, hay que reseñar la publicación relativamente reciente, en 1986, del Diccionario Español de Textos Médicos Antiguos.

del español medieval de Martín Alonso, obra póstuma que es hasta la fecha el mayor de los diccionarios del español medieval existentes. Sin embargo, para el filólogo o el historiador de la lengua dista de ser una herramienta útil por su falta de rigor y de sistematicidad en la recogida de los materiales y por sus carencias en cuanto a técnica lexicográfica. Así lo señaló Stephen Dworkin en un artículo reseña aparecido en el número 47 de Romans Philology, lo que también me exime de realizar más comentarios. Y es una lástima que el de Alonso no resultara ser un diccionario del español medieval aceptable, porque la lengua española no dispone aún de un diccionario de esas características, a diferencia de lo que sucede con la mayor parte de las lenguas románicas, tal y como señala Dworkin en su artículo mencionado, donde reseña los empeños recientes llevados a cabo o en curso de realización para el portugués, el italiano, el francés, el provenzal o el gascón. Sin embargo, es cierto que el español sí dispone de un ambicioso proyecto en curso de redacción de un diccionario del español medieval. Se trata

del así llamado Diccionario del Español Medieval, que dirige en la Universidad de Heidelberg el profesor Bodo Miller. Se trata de un diccionario que, según consta en su prefacio, pretende recolectar el caudal léxico de la lengua española escrita desde los primeros documentos hasta aproximadamente el año 1400. El propósito que guía su realización, y vuelvo a citar el mencionado prefacio, es el de describir en toda su extensión el léxico tradicional fijado en la lengua escrita de la época estudiada, ampliando al máximo el corpus de los materiales analizados. Por este motivo, además de los textos literarios producidos hasta alrededor de 1400, se consideran también los fueros, colecciones diplomáticas, documentos administrativos, escritos técnicos anteriores a 1300, así como el abundante material de la prosa científica del español antiguo y de las traducciones de otras

lenguas. Este planteamiento sólo puede suscitar consenso con la salvedad de la discutible exclusión de los textos correspondientes al periodo 1400-1500.

En cuanto a la estructura del diccionario, sus artículos son complejos y ricos en información. Al lema o palabra definida acompañan las definiciones correspondientes a cada una de sus acepciones, ordenadas según su cronología. Una copiosa documentación textual la respalda, permitiendo a quien consulta la obra acceder a ejemplos de uso de la voz en su contexto. Además de lo dicho, los artículos incorporan referencia a vocabularios, glosarios, concordancias, etc., en que se encuentran más ejemplos de la palabra en cuestión, así como una visión general de su presencia en diccionarios antiguos del español, en diccionarios del español antiguo o en diccionarios históricos. No falta una rica información etimológica con cumplida referencia a las fuentes lexicográficas pertinentes, tanto del español como de otras lenguas. Añádase a esto la especial sensibilidad que Bodo Miller y su equipo de colaboradores muestran hacia el problema de la transmisión de los textos medievales y el de los criterios de edición empleados en la producción. La información de los textos que sirven de base al diccionario.

En lexicografía, y más en lexicografía diacrónica, es obvio que unos materiales textuales de trabajo no fiables traerán consigo un sinfín de deficiencias en el resultado final, por lo que es digno de todo encomio el cuidado que manifiestan Müller y sus colaboradores a estos propósitos. Así las cosas, nos hallamos en presencia de una obra ambiciosa en su planteamiento, acertada en su método e irreprochable en su realización. Pero hay un problema muy serio, el del plazo de terminación de la obra. Los trabajos iniciales del proyecto se remontan al año 1971, y su primer fascículo vio la luz en 1987. En 1994 se publicó el décimo fascículo que remató el primer tomo de la obra. Para realizar una estimativa de plazo de terminación del Diccionario del Español Medieval, piénsese que en este primer tomo se abarca el segmento alfabético, A. Además, y que ocupa 754 páginas.

Hasta la fecha se han publicado seis fascículos más, el último de ellos en 1997, que llega hasta la voz a guardar y completa las primeras 480 páginas del tomo segundo. Es cierto que entre 1996 y 1997 se han publicado cuatro fascículos, pero el ritmo de publicación se antoja lento para todo lo que queda por hacer. Considérese, para tener una idea de lo que falta por hacer, que la voz a guardar, última publicada de este diccionario del español medieval, aparece en la última edición del diccionario de la Real Academia Española en la página 47 y que este diccionario consta de 1513 páginas. Por lo tanto, un proyecto del mayor interés, una calidad científica muy satisfactoria, pero unos plazos de conclusión muy remotos. Ojalá que la utilización de nuevas tecnologías informáticas, ya hace tiempo incorporadas a los trabajos del diccionario del español medieval, y un aumento de los recursos económicos y humanos vinculados a la empresa, permitieran la aceleración de los trabajos de este diccionario.

uso absolutamente imprescindible para el historiador de la lengua, y más teniendo en cuenta la situación de inferioridad en que la tradición lexicográfica española se halla en relación de la de otras lenguas como el francés, el inglés, el alemán, el italiano o el catalán. Por último, haré referencia a la reciente finalización de un proyecto más que centenario, el Diccionario de Construcción y Régimen del filólogo colombiano Rufino José Cuervo, retomado y terminado por el Instituto Caro y Cuervo de Colombia en años todavía muy próximos. Puede ser ocioso caracterizarlo aquí, pero baste decir, siguiendo muy de cerca a

Manuel Seco, que se trata de un diccionario de sintaxis en el cual se consignan únicamente aquellas voces, en su mayoría verbos, afectadas por una casuística peculiar en su funcionamiento sintáctico, pero eso sí, diccionario redactado según un método histórico, esto es, contemplando los fenómenos relativos al régimen de las voces en una perspectiva diacrónica y evolutiva, y adjuntando las características de la tradición de la lengua y, por lo tanto, un gran caudal de textos en los que se documentan.

los usos inventariados. La limitación de su ámbito a unidades léxicas que presentan problemas en su construcción sintáctica hace que este diccionario no sea un verdadero diccionario histórico y, desde luego, condiciona su utilidad. Pero indudablemente el cuervo, como es conocido por los habituales, es un recurso lexicográfico imprescindible convertido al fin en realidad después de muchos años de diversos avatares y desventuras. Cerramos con estas observaciones lo referente a los desarrollos en el ámbito de los diccionarios diacrónicos no históricos. Se ha avanzado, pero indudablemente queda mucho por hacer. Piénsese que tal vez el ya camino de tricentenario diccionario de autoridades publicado por la Academia siga siendo el mejor diccionario existente para la lengua de los siglos de oro. Hablemos ahora del tercer nivel de los mencionados, el de la elaboración de un diccionario histórico del español. Perdona, Juan Carlos, ¿podrías aclarar qué es el diccionario histórico para nuestros alumnos? El diccionario histórico de la lengua española

No, perdón, el diccionario de autoridades, quería decir. Ah, el diccionario de autoridades. El diccionario de autoridades es el primer diccionario realizado por la Academia Española. Es un diccionario que nace de la necesidad de codificar, de establecer, de fijar la lengua del periodo más esplendoroso de las letras españolas, el de los siglos de oro. La Academia nace para preservar esa riqueza lingüística y uno de sus primeros objetivos, único de los que se propone en ese momento, satisfechos, es la creación de un gran diccionario de la lengua española. La Academia construye ese diccionario en un periodo de tiempo muy breve y da lugar, en un periodo de tiempo, como digo, muy escaso, a el que sin duda era en aquel momento el mejor de los diccionarios existentes para cualquiera de las lenguas europeas de cultura. El nombre de diccionario de autoridades viene por... Por el hecho de que cada una de las definiciones o la mayor parte de las definiciones contenidas en el diccionario viene respaldada por una autoridad textual que ejemplifica esa acepción.

definición que se efectúa. Una pregunta más. ¿El diccionario de autoridades es la semilla del diccionario de la academia, del que llamamos DRAE? Bueno, es el punto de partida a partir de un proceso de poda más bien azaroso. La academia buscó realizar una segunda edición del diccionario de autoridades, pero la segunda edición se complicó, no fue capaz la academia de darle encima en el plazo de tiempo previsto, se publicó únicamente la primera edición, perdón, el primer volumen, y entonces la academia decidió para ofrecer un diccionario que era necesario, posponer la elaboración de esa segunda edición del diccionario de autoridades y ofrecer un diccionario despojado de las autoridades, es decir, un diccionario en el que se recogieran únicamente las definiciones. Ese es el origen del llamado diccionario de la Real Academia, que es el que todos conocemos en sus 21 ediciones.

A lo mejor me adelanto un poco, a lo que cuentes a lo mejor dentro de unos minutos, pero ¿por eso es el diccionario de la Real Academia Española tan buen diccionario histórico medieval? Evidentemente muchas cosas que puede encontrar el historiador de la lengua en

el diccionario de la Academia son en cierta medida herencia de ese diccionario de autoridades. Ese es el lado bueno de la inercia que preside la labor lexicográfica de la Academia en buena medida, o que presidía hasta hace bien poco tiempo. La contrapartida de eso es que en el diccionario de la Academia hay todavía, ahora ya menos que antes, pero hay todavía definiciones que están tal cual a la letra tomadas del diccionario de autoridades. Es decir, redactadas en un español dieciochesco y no de forma muy inmediata aprehensible por parte del curioso que se acerca al diccionario de la Academia hoy en día. Y también son definiciones absolutamente enciclopédicas algunas, que no hay cristiano que encuentre a veces el animalito que están definiendo. Hay que reconocer en ese sentido que la Academia está...

haciendo en los últimos tiempos un esfuerzo para hacer de su diccionario usual un diccionario de lengua ajeno a esa especie de lastre o a esa especie de inercia. Se ha redactado recientemente una nueva planta para la redacción del diccionario usual y, sin duda ninguna, esa especie de peso muerto que ha arrastrado a lo largo de todas sus ediciones el diccionario académico se va a perder en breve, dado que por primera vez desde principios del siglo XVIII la academia se enfrenta al reto de redactar su diccionario desde bases enteramente nuevas y desde bases enteramente apegadas a la realidad de los hechos de la lengua y no a la tradición lexicográfica precedente. De acuerdo, si se me ocurre algo más ya te pararé. Perfecto, me parece muy bien. Hablaba entonces de la elaboración de un diccionario histórico del español, recurso del que nuestra lengua carece todavía hoy en día. No hay mucho que se pueda decir acerca del único proyecto actualmente en marcha.

con ese objetivo, el del Diccionario Histórico de la Lengua Española, que elabora la Real Academia Española. Sus últimos fascículos publicados vieron la luz en 1996 y abarcan hasta la palabra Apasanca, fascículo segundo, del tomo tercero, y al mismo tiempo dan inicio a la letra B, primer fascículo del tomo cuarto, lo que obviamente, y contra lo que algunos han pensado, no quiere decir que la redacción de la letra A haya sido terminada. Desde ese año 1996 y hasta la actualidad, la redacción del Diccionario Histórico está detenida, a la espera de que se decida su futuro en el ámbito de las nuevas tecnologías que se introducen en los últimos tiempos en la Real Academia Española. Algo más se dirá más adelante acerca de las posibilidades de viabilidad de este imprescindible recurso. Ocupémonos ahora de las nuevas perspectivas que abre la aplicación de las nuevas tecnologías informáticas al estudio del léxico. Es un hecho que tal vez sea este de los estudios lexicológicos y de la lexicografía, el ámbito de estudios lingüísticos que, de forma más inmediata y obvia, se beneficia de la actualización de la redacción del Diccionario.

de la aplicación y uso de tecnologías informáticas y ya desde hace unos cuantos años, en los que dicho sea de paso, la evolución ha sido vertiginosa. Nos ocuparemos después de forma detenida de los recursos que encarnan de manera dominante esta aplicación de la informática a la investigación lingüística, los corpus lingüísticos, pero ahora tratemos otros aspectos también interesantes que giran todos ellos en torno a un concepto básico, el de texto informatizado. Disponer de un texto cualquiera en formato informático asegura tener la posibilidad de acceder a la información que hay en su interior de forma más rápida y sistemática. Ese acceso, claro está, resulta más accedero si se realiza en el ámbito del léxico que en el de los otros niveles lingüísticos, dado que la presentación gráfica del lenguaje delimita y representa las unidades léxicas con mayor claridad y precisión que las

fonológicas o las sintácticas. Tal circunstancia llevó desde el principio de los desarrollos informáticos aplicados a la filología, a la que se ha convertido en un sistema de la información lingüística, a la que se ha convertido en un sistema de la información lingüística, a la producción de textos electrónicos y a su

su procesamiento mediante programas de generación de concordancias, modo rápido y eficaz de proporcionar al lexicógrafo o al investigador un caudal de materiales que antes sólo le era accesible a través de un penoso e inseguro proceso de despojo manual de textos. Este doble aspecto de texto más sus índices léxicos, es decir, concordancias, pero también índices de frecuencias, inversos, etcétera, etcétera, es el que encontramos en uno de los recursos de investigación más importantes para el ámbito que nos ocupa y también, y desgraciadamente, peor tratados en ocasiones. Estoy pensando en las microfichas producidas por el Seminario de Estudios Hispánicos Medievales de la Universidad de Madison. Sin duda, a este legendario centro de investigación norteamericano le corresponde el privilegio de haber abierto la senda de la aplicación a gran escala de los métodos informáticos a la investigación filológica sobre el español medieval. En los años 60, y encarados a la realización del objetivo del seminario desde su fundación por don Antonio de la Paz, se han realizado muchos estudios de investigación, de investigación, de investigación, de investigación, de investigación, de investigación, de investigación, de investigación, de investigación, de investigación, Antonio García Sololinde, en los años 30.

redacción del diccionario Fould Spanish Language. El director del seminario, el profesor Lloyd Kasten, y su alumno y asistente, el profesor John Nitti, decidieron reemplazar las fichas de papel en que se recogían los ejemplos de uso con materiales informáticos transcritos ex profeso para el proyecto. Y con buen criterio, en un momento dado, el seminario decidió ir publicando todos esos materiales textuales por su interés intrínseco. Esa es la génesis de las famosas ediciones más concordancias en microficha de Madison, muchas veces denostadas por su incomodidad y por los códigos, no tan cabalísticos después de todo, que las pueblan, pero que permiten un grado de conocimiento del léxico rápido y eficaz. Piénsese en las posibilidades que algunas de estas realizaciones, como el Corpus Alfonsi, transcrito por Nitti y Kasten, ofrecen al investigador. En unos minutos se puede saber si una determinada palabra consta en la obra salida del Escriptorium Regio Alfonsi, gracias a las concordancias, en un proceso que de otra manera no se puede hacer. De otro modo, traería consigo la lectura de varios millones de palabras.

de páginas, no siempre accesibles con facilidad, y siempre la sombra de la duda de un error por omisión. Por lo tanto, y como ya puso de relieve hace años el profesor Jerry Craddock, las microfichas del seminario de Madison representan un instrumento de investigación imprescindible, más si se tiene en cuenta que estamos hablando de cerca de 200 textos hispánicos medievales, con sus concordancias a disposición del investigador. Pero el tiempo pasa y las cosas cambian, y más la informática, terreno en que la ancianidad y la obsolescencia avienen apenas en meses. La difusión masiva del CD-ROM en la segunda mitad de los años 80 ofrece un soporte mucho más práctico y versátil que la no demasiado cómoda microficha, y ha permitido al seminario de Madison ofrecer grandes colecciones de textos en ese formato, con la misma estructura y metodología empleadas en el caso de las microfichas, formato de texto más concordancias. Hasta hoy se han publicado cuatro. El que contiene un update de la

venerables microfichas de los textos del Escriptorio Malfoncy, otro que recoge textos medievales navarro-aragoneses, entre los que destaca la inclusión de las obras de Juan Fernández de Heredia, otro que recoge 20 ediciones antiguas de Celestina y otro, por último, en que se encierra el Corpus Médico Español, que tiene su base en los 32 textos producidos para el ya mencionado Diccionario de Textos Médicos Antiguos Españoles. En total, cerca de 200 megabytes de información imprescindible para el investigador sobre el léxico medieval español. La aplicación de estos desarrollos a la producción de obras lexicográficas no ha desembocado todavía en el deseado Dictionary of Old Spanish Language. Ojalá que lo pueda hacer pronto en el desde hace años casi terminado Diccionario de la Prosa Alfonsí. Pero la ejemplificación de las posibilidades y bondades de los desarrollos madisonianos ha encontrado ejemplo práctico en el repetidamente mencionado Diccionario de Textos Médicos Antiguos Españoles, a cuya fisonomía y pronta finalización no ha contribuido poco.

En este ámbito de los recursos informatizados herederos de los planteamientos desarrollados en el seminario de Madison, conviene no dejar de señalar el proyecto ADMITE, Archivo Digital de Manuscritos y Textos Españoles, que ofrece una extensa colección de textos medievales hispánicos en formato informático y codificados según las normas del seminario de Madison. Desgraciadamente, las herramientas de manejo de la información son bastante deficientes y la versatilidad y facilidad del acceso a los datos dejan bastante que desear. Terminaremos este cuarto apartado de la exposición haciendo referencia a los corpus lingüísticos. Sin duda, una de las perspectivas más atractivas generadas por los avances de la informática aplicada al estudio de la lingüística es la construcción de corpus lingüísticos. Un corpus es una colección de textos lo suficientemente amplia y bien seleccionada como para el uso de los textos. Para proporcionar al investigador una muestra representativa de las características de una determinada lengua. Colección de textos, pero de textos en formato electrónico, almacenados en un ordenador de textos.

ordenador y acompañados de los programas informáticos necesarios para la recuperación rápida y sistemática de la información lingüística que contienen esos textos. Los corpus son hijos del desarrollo de la informática, único medio capaz de permitir el almacenamiento de grandes masas de texto en un formato razonable y, sobre todo, de permitir un rápido acceso no secuencial a los textos por medio de herramientas capaces de reconocer, seleccionar y presentar la información requerida. De la amplia tipología de corpus lingüísticos existentes en función de su diseño, el que aquí nos interesa es el corpus diacrónico. Un corpus diacrónico es aquel constituido por textos pertenecientes a un periodo cronológico amplio e, idealmente, pertenecientes a todos los periodos de que conservamos información textual de una determinada lengua. En suma, una colección de textos seleccionada de tal forma que nos permita acceder a datos lingüísticos correspondientes a una etapa significativa de la evolución de una lengua. O, idealmente, a la de todo el desarrollo de su evolución diacrónica.

No abundan en el mundo los proyectos de corpus guiados por esa ambición integral. Actualmente, entre los proyectos que lleva a cabo la Real Academia hay un corpus de esas características, el llamado CORDE, Corpus Diacrónico del Español. Sin duda, cuando esté concluido, facilitará una herramienta de primer orden para el trabajo del historiador de la

lengua y para el interesado en la historia del léxico. Pero desde los intereses de este último grupo de estudiosos, el mayor interés que suscita CORDE es el de poder servir de base material para los trabajos de redacción del Diccionario Histórico de la Lengua Española. El perfil de CORDE como corpus lingüístico asegura una serie de facilidades que no presenta la venerable colección de fichas coleccionadas por la Academia, sobre la que hasta ahora se ha basado la redacción del histórico. Esto permite augurar, sin demasiado riesgo de error, que un diccionario histórico basado en CORDE podría redactarse con mucha facilidad. Y que se pueda hacer con mucha mayor rapidez y sin mengua en la calidad de su fundamento documental y textual, lo que idealmente permitiría solventar la carencia de la redacción del Diccionario Histórico de la Lengua Española.

que hoy sufre la lengua española de un diccionario histórico en un plazo razonable de tiempo. Sin embargo y paradójicamente el desarrollo de corpus diacrónicos dista de ser la fórmula mágica que resuelva el problema y esto a causa de un riesgo perceptible en el nivel metodológico de su diseño y realización. Recapitulemos, un corpus es una colección de textos en formato electrónico más las herramientas necesarias para la extracción de la información lingüística contenida en esos textos. Por lo tanto es evidente que la calidad y utilidad de un corpus dependerá de la calidad de los textos que lo integran, de su ajuste a unos determinados criterios de diseño del corpus, de lo adecuado de los estándares y métodos utilizados para su codificación y marcado y de la bondad de las herramientas empleadas para la extracción de la información lingüística que en él se contiene. Personalmente como filólogo el punto que más me interesa de esa lista de ingredientes son los textos, su calidad, su fidelidad a los textos originales. A las fuentes originales y esto porque los textos son la materia

prima de un corpus como de cualquier trabajo de índole lexicográfica o, en general, filológica. Por buenas que sean las herramientas empleadas en la explotación del corpus y por magnífico que haya sido el trabajo de lematización y marcación morfosintáctica aplicado a los textos, de nada servirán si los textos sobre los que se ha ejercido ese trabajo y que proporcionan a la estructura del corpus los datos que éste aporta al estudioso no resultan fiables, no son fruto de un riguroso trabajo sobre las fuentes primarias. Este problema presenta una variante de menor rango, pero no menos importante, especialmente para los historiadores de la lengua envueltos en estudios grafemáticos, fonéticos, fonológicos o morfológicos referidos a la Edad Media o al Siglo de Oro. Los criterios de edición o transcripción de los textos que configuran el corpus deben ser respetuosos con las peculiaridades gráficas, fonéticas y fonológicas en ellos contenidas. Deben ser especialmente cuidadosos y respetuosos con los textos originales. Por lo tanto, un corpus no es un texto que se puede escribir en un libro, sino que es un texto que es utilizable por todos los historiadores de la lengua. Debe estar basado en transcripciones paleográficas.

figurosas y respetuosas con los manuscritos de que parten. Evidentemente, esto supone renunciar en la construcción del corpus a la utilización de ediciones críticas o, no digamos, modernizadas, ya que en ambos casos tropezaremos con las grafías del editor y no con las del autor o, más bien, las del copista, que son las que interesan al historiador de la lengua. Con lo que los trabajos de construcción de un corpus diacrónico, de por sí faraónicos, pueden verse frustrados por no utilizar la materia prima oportuna, o pueden verse complicados hasta extremos absolutamente utópicos, si quien realiza el corpus decide producir ex profeso los textos adecuados para sus fines. Sin embargo, puede decirse, y es

muy cierto, que quien se emplea en el diseño y la construcción de un corpus diacrónico no puede ejercer verdaderas tareas filológicas, no puede dedicarse a fijar la letra de los textos que integran su corpus, lo imposibilita la enorme cantidad de tiempo y esfuerzo que eso supondría. Además, no sería lógico, ya que lo razonable es utilizar de los logros logrados por una... ...una tradición preestablecida y no hacer...

tabula rasa. Por lo tanto, y como mínimo requisito de calidad científica, es absolutamente imprescindible llevar a cabo un detenido trabajo filológico de evaluación de los materiales textuales accesibles, a fin de basar el corpus en las mejores transcripciones existentes de los textos, todas aquellas que cumplan con los requisitos de fiabilidad antes enunciados. Ahora bien, tampoco puede desestimarse radicalmente esta posibilidad de elaborar materiales ad hoc. Eso es exactamente lo que han hecho, como ya se ha dicho, las autoras del Diccionario Español de Textos Médicos Antiguos, y eso explica en gran medida la calidad final del producto. Y ese mismo propósito guía toda la labor del seminario de Madison. Con todo, no conviene olvidar que, caso de no contar con los ingredientes deseados, lo mejor que puede hacer el investigador es no construir su obra sobre un mal cimiento. El historiador de la lengua no necesita un corpus diacrónico, sino un buen corpus diacrónico, uno que le asegure la fidelidad absoluta respecto de las realidades de la lengua. Y eso es lo que dice el historiador.

Hechas las anteriores salvedades, es evidente que un buen corpus diacrónico de la lengua española, un corpus elaborado con materiales fiables y homogéneos, es una desiderata perentoria para los dedicados a la historia de la lengua y más para los interesados en la historia del léxico de la lengua española. Debemos confiar todos los interesados en esta parcela de la lingüística en que cuando se disponga de productos de esas características, la calidad de las investigaciones será mucho mayor y que determinados instrumentos de los que todavía no disponemos, pienso claramente en un diccionario histórico, podrán ser realidad en un futuro más cercano. Terminemos ya y hagámoslo con la formulación del reto pendiente que tienen los que trabajan en el ámbito de la historia léxica del español. Cuenta el investigador actual con más medios que nunca, con ediciones y transcripciones fiables de textos, con un gran número de monografías sobre palabras o familias de palabras de excelente calidad, con unos desarrollos metodológicos o teóricos cada vez más depurados. Por otro lado, también dispone el filólogo de hoy del auxilio irrenunciable de la lengua española y el de los medios informáticos.

sin duda importantísimo en el ámbito de la historia del léxico. Medios informáticos que, por otra parte, son cada vez más accesibles, más versátiles y más potentes. Así las cosas, todos tenemos derecho a tener la esperanza de que en un futuro experimentarán un notable auge las investigaciones centradas en la historia del léxico del español. Pero todos tenemos la obligación, sobre todo los historiadores de la lengua, aquellos a los que al comienzo aludía como usuarios de una serie de herramientas de investigación, pero a los que ahora apelo como protagonistas de su desarrollo, de que ese auge no se difumine únicamente en una superproducción de trabajos micrográficos dispersos, sino que también y fundamentalmente cristalice en un trabajo colectivo a medio o largo plazo que conduzca a la creación de los grandes recursos imprescindibles para el estudio científico de una lengua en su evolución diacrónica. Muchas gracias Juan Carlos. Muchísimas gracias por esta pequeña lección sobre lexicografía histórica y instrumentos para estudiarla y hacer los diccionarios. Mi intención, al llamarlo, es llamarte a ti este año y otros años.

años y los profesores que iban a haber venido era ofrecerles a los estudiantes una pequeña lección de semántica de psicografía histórica, porque es una parcela que ustedes no van a ver en los cursos normales de Historia de la Lengua. No solamente en la universidad a distancia, sino en cualquier otra universidad, puesto que no hay tiempo material en los planes de estudios para embutir toda la Historia de la Lengua. Y la de psicografía histórica es una parte de ella. Y en muchos sitios es un simple curso de doctorado, al que no todos ustedes llegarán, porque no lo querrán. Buenas noches y buena suerte en los exámenes. Bien, agradecemos desde aquí la participación en este espacio a Juan Carlos Conde, investigador de la Real Academia Española y profesor de la Universidad Complutense que aceptó la invitación de José Manuel.

Manuel Fradejas, profesor titular de Historia de la Lengua Española I en la UNED. Entre ambos nos proporcionaron unas notas para acercarnos al conocimiento de la lexicografía histórica. Por nuestra parte, nada más. Aula de Filología volverá a estar de nuevo con ustedes el próximo lunes a partir de las ocho y media de la tarde, hora peninsular. Gracias por la atención que nos han prestado hoy hasta entonces. En el tiempo de la UNED es el turno ahora de nuestra compañera Pilar Carrasco y Revista de Geografía e Historia. Gracias.