

# De una obra digitalizada a una obra digital: reflexiones en torno al *Diccionario Griego-Español* en línea\*

SABINE THULLIER  
ILC-CSIC  
[sabine.thuillier@cchs.csic.es](mailto:sabine.thuillier@cchs.csic.es)

Un amplio repertorio de obras impresas ha pasado a estar accesible en línea. Podría decirse que el proceso ya resulta banal. Sin embargo, el grado de interés y de éxito de estas empresas no está siempre garantizado. ¿Cómo es posible que una obra ocupe su lugar de pleno derecho en el mundo digital? ¿Cómo la digitalización de obras puede desembocar en una genuina obra digital que suponga un salto cualitativo en la forma de acceso y lectura? Ese salto reside, diremos aquí, en dos de sus rasgos más destacados: el dinamismo y la conectividad.

En este artículo, nos proponemos reflexionar sobre las condiciones que han de darse para que una conversión del papel a la pantalla sea en cierto sentido exitosa. Para ello centraremos nuestra atención en un tipo de texto cuyo aprovechamiento y uso en su formato electrónico es de considerable importancia: los diccionarios. La cuestión que nos interesa debatir aquí es la siguiente: ¿Cómo han de abordarse estos parámetros generales, propios de toda obra digital, en el caso específico de los diccionarios? Para dar una respuesta, provisional, a esta pregunta, vamos a partir de la digitalización del *Diccionario Griego-Español*<sup>1</sup> y reflexionar sobre las posibilidades y limitaciones para su transformación en una obra digital de pleno derecho. Mostraremos que en el caso de los diccionarios convertirse en una obra digital de pleno derecho depende de su conectividad con otros recursos y, sobre todo, de que pueda desarrollarse en diálogo con su corpus de modo dinámico, de tal manera que se haga disponible la información lingüística y lexicográfica para sus potenciales usuarios. Sólo de esta manera podrá convertirse en una herramienta filológica nueva. El proceso no carece de riesgos, pues se produce una cierta paradoja:

---

\* Este estudio ha sido realizado en el marco del proyecto de investigación FFI2010-20236. Agradezco a Juan Rodríguez Somolinos sus aportes y muy especialmente a Jesús Vega su inmensa ayuda.

<sup>1</sup> Más adelante [DGE], <<http://dge.cchs.csic.es/xdge/>>.

a mayor dinamismo y conectividad, quizá una mayor tensión en los valores de rigor y de precisión lexicográficas. El énfasis en los diccionarios como herramientas definidas por sus usuarios potenciales y en la necesidad de aumentar su conectividad y dinamismo no debe hacer perder de vista el horizonte desde el cual proceder a la transformación de un diccionario digitalizado en una obra digital: la propia lexicografía.

Procederemos del siguiente modo. Vamos a presentar los resultados de la digitalización del *DGE* en el marco de la clasificación de los diccionarios digitales propuesta por Sven Tarp (2011 y 2021). El *DGE* en línea es lo que Tarp llama un "caballo más rápido". A continuación, evaluaremos cómo y en qué medida puede convertirse en un tipo de diccionario genuinamente digital. Concluiremos con una discusión de las posibilidades y límites de tal tarea y las consecuencias para los diccionarios de griego antiguo.

## 1. UNA NUEVA ETAPA DE LA LEXICOGRAFÍA

Todos estaríamos de acuerdo en que las herramientas informáticas desarrolladas en las últimas décadas están contribuyendo a una renovación de la lingüística. Es cierto igualmente que la lexicografía no escapa a este destino. Sin entrar en un debate sobre las relaciones de la lingüística con la lexicografía, no es menos cierto que esta última ocupa un lugar peculiar entre los trabajos lingüísticos. Las investigaciones que emprende quedan reflejadas en un objeto indisolublemente unido a sus potenciales usuarios. No se puede por menos de destacar su carácter artefactual y funcional. Los diccionarios son primariamente *objetos de consulta*.

Este aspecto esencial cobra, si cabe, una importancia decisiva en la configuración de los diccionarios digitales. Por eso mismo, una de las principales preocupaciones ha de ser la de lograr una mayor adaptación a las necesidades reales de los usuarios o, incluso, a potencialidades que están aún por descubrir. Este es un punto central de algunas teorías lexicográficas recientes: la insistencia en las necesidades de los usuarios constituye el rasgo común al estudio de los diccionarios en línea. Se puede decir incluso que su grado de adaptabilidad condiciona su grado de "digitalidad".

Una de las primeras preguntas que cabe hacerse concierne, pues, a las implicaciones teóricas y prácticas de este cambio de soporte para los diccionarios. ¿Qué significa el paso de una lexicografía de papel a una lexicografía electrónica? En primer lugar, es preciso señalar que la distinción no se ha de hacer de forma puramente cuantitativa; involucra esencialmente cambios cualitativos dependientes del modo en que son aprovechadas las funcionalidades de los ordenadores y las tecnologías de la información. Implica más datos, más flexibilidad y adaptabilidad, una mayor manejabilidad, más accesibilidad y una mayor interconexión con fuentes externas.

Si se recuerda, los inicios de la lexicografía electrónica estaban más bien destinados a producir herramientas que fueran legibles para las propias máquinas. La digitalización del griego antiguo es un buen ejemplo de ello. A mediados de los años 90, el proyecto Perseus<sup>2</sup> emprendió las tareas de digitalización

<sup>2</sup> <<http://www.perseus.tufts.edu/hopper/>>.

de la edición de 1940 del diccionario de referencia *A Greek English Lexicon*<sup>3</sup>, con el fin de enriquecer y controlar la base léxica que necesitaba su analizador morfológico. El proyecto llevó a cabo una versión etiquetada de la obra, que es desde entonces de libre consulta en línea y ampliamente usada por otros proyectos<sup>4</sup>. Perseus ha conseguido formalizar un recurso tradicional que, sin embargo, como señalaremos más adelante, está mal adaptado a un entorno electrónico.

Es en este contexto en el que se va a situar nuestra reflexión, el de los avances que puede significar para un diccionario griego proseguir su conversión en obra genuinamente digital. También aquí es necesario insistir en un cambio de énfasis, que no se limite al aprovechamiento "automático" de las informaciones léxicas sino que tenga en cuenta a los usuarios y sus necesidades. Y aquí las peculiaridades del usuario de los diccionarios del griego antiguo aparecen en primer plano. Los usuarios de estas herramientas léxicas no responden a necesidades de *producción* de texto (salvo quizá en el caso de aquellos lingüistas que se sirvan de los diccionarios para producir elementos ficticios para la contrastación de sus hipótesis), sino a necesidades ligadas a la *recepción*, como por ejemplo la traducción de textos. Los diccionarios son objetos de uso cotidiano entre los helenistas, que les atribuyen una cierta autoridad, quizá mayor que en otros casos.

Por eso, es necesario reflexionar sobre qué tipo de tareas de digitalización se ha de emprender con los diccionarios de griego antiguo y, ante todo, qué tipo de obra digital es recomendable. La experiencia en la digitalización del *DGE* puede ser aquí de ayuda. Por ello, es conveniente en primer lugar situarlo entre las categorías de diccionarios que uno puede encontrar en línea. Para ello, vamos a utilizar la clasificación propuesta por Sven Tarp (2011: 58-62 y 2012: 257-259).

Tarp clasifica los diccionarios en línea en cuatro grandes categorías: copiones, caballos más rápidos, modelo Ford-T y Rolls Royce. A pesar de lo que pudiera parecer, la clasificación no responde únicamente a criterios de mayor velocidad en el tratamiento y en el acceso a la información.

- Los "copiones" son copias fotografiadas o escaneadas de diccionarios impresos ya existentes. No existen muchos ejemplos en la lexicografía griega (por lo menos moderna). Citaremos el Bailly *Abrégé*<sup>5</sup>. Su utilidad es escasa, su dinamismo es nulo y, a menudo, resultan menos prácticos que su versión en papel.
- Los "caballos más rápidos" proporcionan una visualización estática en pantalla de los artículos del diccionario. Están modelados a partir de los artículos correspondientes en los diccionarios impresos. Puede decirse que la mayor parte de los diccionarios actualmente en línea son de este tipo y suponen una mejora sustancial por lo que respecta a un acceso más rápido a los datos mediante el uso de enlaces y cadenas de búsqueda. Las varias versiones electrónicas del LSJ corresponden a esta categoría.

<sup>3</sup> En adelante [LSJ].

<sup>4</sup> Entre otros: <<http://philolog.us/>>, <<http://perseus.uchicago.edu/Reference/LSJ.html>> y <<http://stephanus.tlg.uci.edu/lsj/>>.

<sup>5</sup> <<http://home.scarlet.be/tabularium/bailly/>>.

- Los “modelos Ford-T” ofrecen ya artículos dinámicos con datos dinámicos y adaptados a las distintas funciones que pueden desempeñar los diccionarios. Uno de sus rasgos es el hecho de que ofrecen la capacidad de *reutilizar* datos ya existentes. Logeion, un proyecto en pleno desarrollo que proporciona una búsqueda simultánea en varios diccionarios de latín y griego presentes en la biblioteca digital de Perseus<sup>6</sup>, es quizás el único caso de la lexicografía griega que podría pertenecer en ciertos aspectos a esta categoría.
- Los “Rolls Royce” son concebidos como herramientas lexicográficas avanzadas e individualizadas. En ellas se busca combinar un acceso selectivo a los datos incluidos en una base de datos con la búsqueda en internet (y en corpus existentes). La funcionalidad principal de estos “diccionarios” reside en recrear y re-presentar los datos relevantes en soluciones dinámicas para los usuarios. Tarp comenta que no parece existir aún ninguna herramienta de esta categoría y lo que ofrece la lexicografía digital del griego antiguo no le puede contradecir.

De estos dos últimos tipos de diccionario se puede decir que son obras lexicográficas genuinamente electrónicas, mientras que los dos primeros pertenecen aún a una lexicografía de papel. ¿En qué categoría incluir el *DGE* en línea? Como veremos, su realidad lo sitúa entre los caballos más rápidos; su ambición es avanzar a un modelo Ford-T. Pero ¿por qué no hacer de él un Rolls Royce de los diccionarios en línea? A ello responderemos en la última sección.

## 2. EL *DGE* EN LÍNEA: UN CABALLO MÁS RÁPIDO

### 2.1. El proyecto editorial y su digitalización

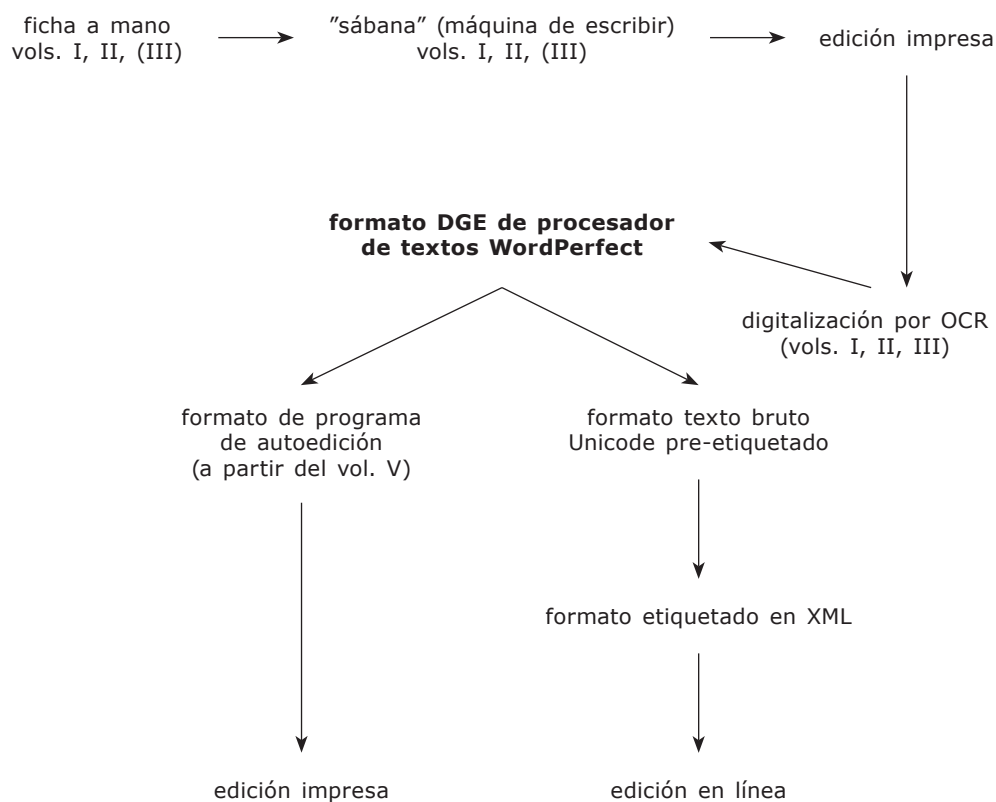
Pocos son los diccionarios del griego antiguo que tengan una versión digital. Recientemente se ha sumado al ya conocido LSJ la parte del *DGE* publicada hasta el momento. A continuación, vamos a presentar este diccionario en el marco de la clasificación de Tarp de los diccionarios electrónicos. Sin duda, por el momento cabe decir que no se trata sino de un “caballo más rápido” que tiene la pretensión de convertirse en un verdadero modelo de Ford-T.

El *Diccionario Griego-Español*, iniciado a finales de los años sesenta y dirigido por el profesor F.R. Adrados, es un diccionario bilingüe unidireccional (griego antiguo – español). Se trata de una obra en curso de elaboración que se publica volumen a volumen desde 1980; los siete volúmenes ya editados cubren la sección alfabética: *ἀ-ἐξαιος* (59.329 entradas). Su ambición es perseguir la larga tradición de la lexicografía griega y contribuir a su renovación, en particular actualizando y enriqueciendo su predecesor inmediato y modelo, el LSJ. El léxico de la lengua griega está analizado de la manera más exhaustiva posible: el corpus despojado comprende todos los textos literarios y documentales transmitidos (incluso el corpus papirológico y epigráfico) y cubre un arco cronológico de más de 1.500 años, desde Homero a los últimos papiros escritos en Egipto en el s. VIII d.C., e incluye la literatura cristiana, los nombres propios, así como el griego micénico. El *DGE*, partiendo lógicamente del LSJ, pretende

<sup>6</sup> <<http://logeion.uchicago.edu/>>.

reelaborar desde cero la organización y estructura interna de los artículos a partir de la lengua de salida, el español. Pretende establecer en qué circunstancias contextuales (lingüísticas o extralingüísticas) una palabra griega puede o debe traducirse por tal o cual palabra española, precisada o matizada por las explicaciones pertinentes y, eventualmente, los contextos griegos de las citas y sus traducciones<sup>7</sup>.

El proyecto de digitalización del *DGE* se emprende con el objetivo de revalorizar un texto regular y fuertemente estructurado. No es este el lugar de describir cómo se ha llevado a cabo todo este proceso. A modo de resumen, el siguiente esquema presenta los distintos soportes y formatos en que ha viajado el *DGE* desde sus inicios hasta llegar a una versión en HTML:



Por su importancia en este proceso, merece la pena recordar que el proyecto ha procedido a llevar a cabo *LMPG*<sup>8</sup>, edición en línea de una publicación aneja al *DGE*, que ha servido como banco de pruebas. El *Léxico de magia* y

<sup>7</sup> Para más información sobre la historia y las características del *DGE* remitimos al Prólogo del vol. I y a la presentación en línea: <<http://dge.cchs.csic.es/hist>>.

<sup>8</sup> <<http://dge.cchs.csic.es/lmpg/>>.

*religión en los papiros mágicos griegos*<sup>9</sup> sólo incluye 2.627 entradas y analiza semánticamente un corpus homogéneo y restringido, el de los papiros mágicos griegos del s. II a.C. hasta el s. V d.C. Su interés para esta empresa residía precisamente en el hecho de que sigue de modo muy estricto la estructura del *DGE*.

Como el primer objetivo de este proceso era “meramente” obtener una presentación HTML del *DGE*, el tipo de informaciones marcadas ha estado codificado por una consecución del objetivo en tiempo razonable. Para ello se ha diseñado un esquema específico conforme a la *Text Encoding Initiative*<sup>10</sup>.

## 2.2. El *DGE* en línea: estado actual

Obtener una presentación en HTML no significa, en principio, más que un cambio de soporte, un cambio que no obstante supone ventajas sustanciales de legibilidad, aunque el acercamiento sea aún muy similar a la consulta de artículos en papel: buscar lemas, leer artículos. Pero el acceso al texto se hace más fluido y empieza ya a romper la rigidez del papel. Señalaremos algunos de los avances que sugieren un mejor aprovechamiento de potencialidades digitales de la tarea ya realizada:

- La navegación en el diccionario. Sin duda, el *DGE* digitalizado es un caballo más rápido, que permite un mejor y más veloz acceso a la información lexicográfica. Para ello se ha implementado una potente herramienta de transliteración (elemento muy útil para el griego) así como una herramienta de construcción de listados inversos, que podría ser de interés para estudios morfológicos, literarios, etcétera.
- La visualización de los artículos mejora sustancialmente en relación al diccionario en papel, pues se ha insistido en una visualización *estructurada*, basada en mayor espacio entre los apartados repartidos en párrafos indentados. Esto ha significado un cierto valor lexicográfico añadido que pone de relieve la composición muy jerarquizada de los artículos.

La descripción de Tarp (2011: 60) de los caballos rápidos se ajusta perfectamente a los resultados obtenidos: “The result of this restrictive use of the technological possibilities is faster lexicographical horses, where the data included are still organized in traditional and static articles, which are completely modelled on the corresponding articles in printed dictionaries”. Cabe por tanto decir que, de momento, abandonar el papel no se ha traducido en la generación de un objeto digital nuevo. De hecho, este no era el primer objetivo de la digitalización iniciada, más preocupada por mejorar la visibilidad y la accesibilidad a datos lexicográficos básicos. Sin embargo, estos primeros pasos traducen a partir de ahora una toma de conciencia de las nuevas posibilidades ofrecidas por el diccionario. Sin duda, la codificación realizada permite ya implementar funciones más avanzadas de consulta y de búsqueda. Pero sigue siendo un reto avanzar en la conectividad entre los propios recursos del diccionario y abrir también la puerta a nuevas conexiones con recursos ajenos.

<sup>9</sup> Luis Muñoz Delgado (2001).

<sup>10</sup> <<http://www.tei-c.org/>>. Para más información sobre la codificación del *DGE*, cf. Glorieux-Thuillier (2010).

### 3. POSIBILIDADES Y LÍMITES PARA UNA MAYOR "DIGITALIDAD"

"If there is one single feature likely to be applicable to all E[lectronic] D[ictionary]s of the future, it is that they will stop functioning as stand-alone products"<sup>11</sup>.

La sección anterior ha descrito las características que hacen del *DGE* en línea, en su versión actualmente accesible, un diccionario electrónico del tipo que Tarp califica de "caballo más rápido". Es claro que, de momento, su salida del papel no ha logrado producir un nuevo objeto. Pero nos podemos preguntar cuáles son las posibilidades de que un diccionario de este tipo pase a engrosar las filas de la categoría de Ford-T, una categoría que, quiero recordar, se definiría por *artículos dinámicos* que contienen a su vez *datos dinámicos*. Más allá de un recurso lexicográfico digitalizado, meramente convertido de un formato a otro, el reto fundamental es perseguir el tratamiento de los datos y lograr así un recurso digital dinámico, enlazable y compartible. Vamos por consiguiente a proponer pistas para desarrollar la conectividad y el dinamismo de varios tipos de datos contenidos en el *DGE*, tanto los datos internos al propio proyecto como otros procedentes de otras fuentes. En un segundo momento, explicaremos los límites previsibles a desarrollos en esta línea y las razones por las que una obra como el *DGE* no puede pretender alcanzar la categoría de Rolls Royce. Al final propondremos algunas reflexiones sobre las consecuencias para la evolución futura del *DGE*.

#### 3.1. Conectarse a sí mismo: dinamizar los datos internos

*Dinamizar los campos.* Ya hemos dicho que las funcionalidades actuales del *DGE* en línea no proporcionan realmente accesos novedosos a los datos. Es, por tanto, preciso conseguir que se puedan llevar a cabo múltiples búsquedas en la base de datos. Este nuevo tipo de navegación, que rompería aún más la linealidad del texto, podría así alcanzar un cierto estadio avanzado de sofisticación según sea el grado de granularidad de la codificación de la fuente y la implementación de un motor de búsquedas adecuado. Esto podría producir una especie de reorganización de los datos y construir así nuevas redes a través de distintos artículos. Los resultados obtenidos en los modos de consulta avanzada del *LMPG* son ya convincentes (búsqueda a partir del corpus citado<sup>12</sup>, traducciones españolas, marcas de uso, etc.<sup>13</sup>) y es preciso adaptar este sistema de búsqueda al *DGE*. Será posible entonces visualizar las informaciones de un modo que es imposible en el LSJ, cuya fuente XML está marcada de modo mucho menos fino.

Del mismo modo, se obtendrá un valor añadido de primera magnitud al extender este dinamismo a otros datos esenciales del diccionario y que están almacenados en otros lugares. He aquí algunos ejemplos:

*Conectar las citas con su información bibliográfica.* Los siete volúmenes en línea del *DGE* contienen 370.000 citas. Recordemos que se trata de una descripción lexicográfica fundada sobre el corpus y que toda acepción, todo matiz semántico, debe estar documentado por una o, muy a menudo, varias

<sup>11</sup> De Schryver (2003: 189).

<sup>12</sup> <<http://dge.cchs.csic.es/lmpg/?textos=>>.

<sup>13</sup> <<http://dge.cchs.csic.es/lmpg/busqueda>>.

referencias al corpus. Más que desempeñar un papel de ilustración, el reparto de las referencias textuales constituye de hecho la base de organización de los artículos. Como es común en los diccionarios de lenguas antiguas impresos, el sistema de abreviaturas empleado para estas referencias es particularmente complejo y denso, y resulta tal vez poco intuitivo para muchos lectores. Dada la importancia de este campo, parece necesario enlazar de manera más fluida estas abreviaturas a su forma completa recogida en las listas bibliográficas informatizadas pero aisladas en otra parte del servidor<sup>14</sup>. Estas listas detallan todas las ediciones de referencia que utiliza el *DGE* y están en permanente proceso de actualización; representan una fuente de primer orden para quien quiere disponer de una información bibliográfica bastante completa sobre el corpus griego editado. Dinamizar la citas del diccionario y conectarlas con estas listas, además de hacerlas más legibles, abrirá la posibilidad, entre otras, de navegar a través el *DGE* a partir del corpus analizado, crear concordancias, etcétera.

*Documentos léxicos.* Existen muchos otros datos, muy valiosos, desarrollados en el proyecto filológico alrededor de la elaboración del *DGE*. Pensamos en primer lugar en el *Repertorio Bibliográfico de la Lexicografía Griega*<sup>15</sup>. Bajo el artículo aparecerían las referencias bibliográficas a estudios léxicos sobre la palabra que hay en RBLG (algunas de ellas además quizá también accesibles en Internet). Sin duda, la literatura secundaria ocupa un lugar muy importante en el trabajo de los lexicógrafos del *DGE*. Esta conectividad sería un modo de hacer visible el trabajo filológico que está en el origen de los artículos y proporcionaría también una ayuda innegable a ulteriores investigaciones<sup>16</sup>.

Se podrá también incorporar al *DGE* en línea los muy numerosos *addenda et corrigenda* que se acumulan en los archivos: palabras nuevas, palabras fantasma, nuevas acepciones, corrección de errores, etc. Los *addenda et corrigenda* se presentarían integrados en el *DGE* pero de un modo que los haga claramente identificables, esto es, manteniendo el original y el suplemento.

### 3.2. Conectarse con datos externos

Los proyectos digitales de la comunidad de estudios clásicos están en plena expansión. La cantidad de información relacionada con datos sobre el mundo clásico disponibles en línea va aumentando día a día, y es identificable un deseo de formular consultas capaces de cruzar estos diferentes almacenes de datos. Con toda seguridad, el *DGE* tiene un papel que desempeñar en esta comunidad digital.

Los diferentes tipos de recursos que se van a proponer aquí se reparten según su orden de proximidad con el *DGE*. Es obvio que esta lista podría quizá extenderse más lejos, a todo lo que contiene internet: "From a sound metalexigraphic point of view this is hard to accept, yet the Internet can definitely be used as the largest available collection of electronic data [...], with many hundreds of dedicated corpora also accessible" (de Schryver 2003: 158)

*Otros recursos lexicográficos.* La conexión con proyectos similares permite sin duda un aprovechamiento conjunto de los recursos y ofrece a los usuarios

<sup>14</sup> <<http://dge.cchs.csic.es/lst/lst-int.htm>>.

<sup>15</sup> RBLG: <<http://dge.cchs.csic.es/blg/blg-s.htm>>.

<sup>16</sup> Cf. Lee (2010: 126 y 134).



la posibilidad de contrastar la información léxica proporcionada por varios diccionarios. Aparte del caso obvio de conectar los artículos del *DGE* con los del LSJ, pensamos igualmente en otros diccionarios de griego antiguo nativamente digitales, que saldrán a la luz próximamente, como el Cambridge Greek Lexicon<sup>17</sup> y el dicciogriego<sup>18</sup>. Otro posible tipo de intercambio tiene más que ver con la lexicografía computacional. La fuente léxica etiquetada del *DGE* puede contribuir al desarrollo de bases de datos léxicos existentes. Esto permitiría actualizar y corregir los lematizadores disponibles (Perseus, TLG). El *DGE* contiene en efecto un número considerable de lemas que faltan en el LSJ. A cambio, el *DGE* debe a su vez implementar estas herramientas lingüísticas, para permitir por ejemplo búsquedas lematizadas de las palabras griegas, función elemental de cualquier diccionario electrónico.

"Re-conectar" un diccionario a su corpus<sup>19</sup>. Conviene aquí insistir sobre la situación excepcional del corpus literario y documental de la lengua griega antigua: una parte considerable de los textos griegos está disponible en formato digital y en gran medida lematizado. Entre el número creciente de proyectos editoriales, nos gustaría destacar el importantísimo *Thesaurus Linguae Graecae*<sup>20</sup> que ha recopilado y digitalizado la mayoría de textos literarios griegos, desde Homero hasta 1453<sup>21</sup> y la biblioteca digital de Perseus, de mucho menor tamaño<sup>22</sup> pero que está considerada en el ámbito de las humanidades digitales como un recurso de referencia por la interactividad de los datos almacenados. Por lo que se refiere al corpus no literario, es preciso citar para los papiros el Duke Databank of Documentary Papyri<sup>23</sup>, y para las inscripciones tanto el Packard Humanities Institute Searchable Greek Inscriptions<sup>24</sup> como el *Supplementum Epigraphicum Graecum*<sup>25</sup>.

Dada esta abundancia de fuentes textuales en formato digital, la lengua griega está en una situación particularmente favorable: casi cualquier hipótesis lingüística se puede comprobar en la casi totalidad de las ocurrencias conservadas del léxico. Así, se puede entender sin dificultad que la interconexión del corpus con su descripción lexicográfica resulte ser una importante clave de la lexicografía digital del griego antiguo. Por eso, la afirmación de Toma Tasovac (2010), "we cannot think of dictionaries any more without thinking about digital libraries and the status which electronic texts have in them", nos parece aplicarse de manera muy oportuna al ámbito helenista. Esta interconexión ya se ha hecho efectiva en Perseus y en el TLG. Ambos integran en su colección diccionarios, entre los que se cuenta el LSJ. Esta interoperabilidad corpus-diccionario constituye el mayor desafío del actual desarrollo digital del *DGE* en línea. Uno de los mayores provechos residiría en que los usuarios pudieran acceder de manera rápida y fácil a los numerosos textos a los cuales *DGE* sólo hace referencia, sin proporcionar contexto.

<sup>17</sup> <[http://www.classics.cam.ac.uk/faculty/research\\_groups\\_and\\_societies/greek\\_lexicon/](http://www.classics.cam.ac.uk/faculty/research_groups_and_societies/greek_lexicon/)>.

<sup>18</sup> <<https://sites.google.com/site/dicciogriego/>>.

<sup>19</sup> Cf. de Schryver (2003: 167-168) y Rundell (2012: 78-77).

<sup>20</sup> <<http://stephanus.tlg.uci.edu/>>.

<sup>21</sup> "105 million words from over 10,000 works associated with 4,000 authors".

<sup>22</sup> Acerca de 10 millones de palabras griegas.

<sup>23</sup> <<http://papyri.info/>>.

<sup>24</sup> <<http://epigraphy.packhum.org/inscriptions/>>.

<sup>25</sup> <<http://hum.leiden.edu/history/research/projects-umw/seg.html>>.

No obstante llevar a cabo un sistema hipertextual adecuado entre el diccionario y los corpus es una empresa delicada que requiere tanto un trabajo colaborativo amplio como gran precaución. Por un lado, existe la dificultad de normalizar los identificadores de las citas, dada la alta variación de formas de citar los textos; alcanzar un marcado operativo al nivel de la referencia exacta de una cita es una tarea costosa. Por otro lado, existe el riesgo de remitir a una edición diferente (a menudo más antigua) de la adoptada por el diccionario. Eso conlleva la posibilidad de tener lecturas diferentes del mismo pasaje y por lo tanto menoscabar el atento análisis del lexicógrafo. Seguramente, se puede encontrar toda una serie de soluciones para cruzar, tecnológica y científicamente, la barrera de estas variaciones textuales y editoriales<sup>26</sup>. Cabe señalar, por otra parte, que no todos los recursos mencionados están en acceso abierto, y esto puede naturalmente dificultar los protocolos de intercambio de datos con algunos proyectos.

*Datos extra-lingüísticos.* El *DGE* no tiene intenciones enciclopédicas propiamente dichas. Sin embargo, a diferencia del LSJ, incluye nombres propios (más de 10.000). Sin seguir los criterios de exhaustividad aplicados al resto del léxico, sus motivaciones quedan todavía a un nivel lingüístico y filológico<sup>27</sup>. En este dominio, la comunidad de estudios clásicos es especialmente activa y se puede contar con una cantidad de recursos que recopilan datos de carácter enciclopédico sobre el mundo antiguo. Son recursos geográficos (el proyecto Pleiades<sup>28</sup> por ejemplo), biográficos y onomásticos (como el gran *Lexicon of Greek Personal Names*<sup>29</sup>), iconográficos (como el *Lexicon Iconographicum Mythologiae Classicae*<sup>30</sup>), sin olvidar todos los metadatos disponibles para los corpus documentales, sobre todo papirológicos<sup>31</sup>. Conectar el *DGE* con recursos de este tipo permitiría facilitar una justa ampliación de su modesta parte extra-lingüística<sup>32</sup>.

### 3.3. Límites y brechas

En lo dicho hasta ahora, hemos intentado proponer diferentes pistas para dinamizar el *DGE* en línea dentro de un marco que le sería, por así decir, favorable y contribuiría a proporcionar accesos a los datos y herramientas digitales novedosos y útiles para la investigación filológica. El grado de factibilidad técnica de esta tarea es variable y depende de los tipos de interconexiones previstas. Pero nos parece que el límite más serio a una dinamización más amplia del

<sup>26</sup> Conviene mencionar los esfuerzos de estandarización desarrollados por el Canonical Text Service, recientemente seguidos por el Perseus Catalog, y también por la comunidad papirológica.

<sup>27</sup> "Los nombres propios son tan signos lingüísticos como los comunes, que a veces conservan o presentan lexemas no documentadas en otras clases de palabras, que un diccionario es un instrumento de trabajo que debe resolver dudas a un traductor (y no siempre es fácil saber ante un nombre propio si se trata de un mes, de persona, o de lugar, lo cual influiría, por ejemplo, en la traducción que rige al nombre propio), etc.", Prólogo del vol. I del *DGE*, p. XXXV.

<sup>28</sup> <<http://pleiades.stoa.org/>>.

<sup>29</sup> <<http://www.lgpn.ox.ac.uk/online/index.html>>.

<sup>30</sup> <<http://www.limc-france.fr/>>.

<sup>31</sup> Entre otros, <<http://www.papyri.info/>>.

<sup>32</sup> Cf. Rundell (2012: 78): "the boundaries between dictionary and encyclopedia (always problematic) is increasingly irrelevant".

contenido de este diccionario procede sobre todo de cuestiones teóricas y científicas. Hemos hablado más arriba de las precauciones que hay que tomar a la hora de vincular las referencias de los textos citados por el *DGE* con los corpus electrónicos que proponen estos "mismos" textos, con el consiguiente riesgo de alterar la precisión de los datos. Por lo tanto, varios puntos contradictorios y delicadas preguntas surgen inevitablemente acerca del grado de adecuación de este tipo de recursos –a pesar de todo diseñados de una forma tradicional– a un ámbito digital "hiperconectado" y altamente modulable y adaptable según las necesidades de los usuarios, con el fin de lograr alcanzar la categoría de Rolls Royce descrita por Tarp.

Evocaremos brevemente varios aspectos que sugieren estas limitaciones; profundizar en el asunto que está en juego desbordaría el propósito de estas páginas. En primer lugar, si bien las capacidades de almacenamiento de datos internos y enlaces a fuentes externas no parecen tener ya límites físicos, no se puede abrumar al usuario con información que podría resultar disparatada. Lo repetimos, el diccionario es (y tal vez aún más en su forma electrónica) una herramienta de consulta puntual, que secunda en la mayoría de los casos otra actividad principal. Existe la necesidad de una mayor claridad, pertinencia, oportunidad, con la menor distracción posible<sup>33</sup>. ¿Hasta qué punto estos requisitos son compatibles con los sueños de una interconexión *ad libitum*? La primera medida consistiría en elaborar un motor de búsqueda capaz de adaptarse a las necesidades del usuario y proponer esta "pos-navegación" como una elección discreta y jerarquizar así los vínculos según su pertinencia, siempre incluyendo una prevención sobre el valor filológico de lo que puede ofrecerse. No se trata sólo de la cuestión trivial de cómo se debe presentar la información sino también de responder a las exigencias de exactitud que esperan de las explicaciones lexicográficas los usuarios, que son en el caso que discutimos aquí, filólogos y más generalmente clasicistas.

Finalmente, ante esa paradoja amenazante, a saber, que al aumentar su conectividad el diccionario corre el riesgo de perder en parte su rigor científico, un proyecto lexicográfico como el *DGE*, procedente del papel, tiene que reasentar sus ambiciones para no sólo soslayar esta paradoja sino también para aprovechar más a fondo el cambio radical que implica la era digital para la lexicografía. El *DGE* se encuentra ahora de cierta manera a caballo entre el papel y la pantalla, tanto por lo que respecta a su parte ya publicada y digitalizada como por lo que respecta a lo que queda por elaborar. Hemos intentado a lo largo de estas páginas sugerir unas pistas para que el contenido actual de *DGE* en línea, a saber siete volúmenes impresos digitalizados, pueda en un futuro ganar un lugar de pleno derecho en el seno de los diccionarios digitales. Estos retos podrán quizás servir para encarar con mayores garantías la necesaria adaptación de sus pretensiones y de sus métodos, para proseguir la redacción de la obra al tiempo que se desembaraza definitivamente del papel.

## BIBLIOGRAFÍA

*A Greek-English Lexicon*, H.G. Liddell, R. Scott, R. y H.S. Jones (eds.), Oxford, Clarendon, 1940 (9.ª ed.).

<sup>33</sup> Cf. de Schryver (2003: 189).

- De Schryver, Gilles-Maurice, "Lexicographers' Dreams in the Electronic-Dictionary Age", *International Journal of Lexicography*, 16 (2) (2003), pp. 143-199.
- Diccionario Griego-Español*, Francisco Rodríguez Adrados et al. (eds.), Madrid, Consejo Superior de Investigaciones Científicas, 1980-.
- Glorieux, Frédéric y Thuillier, Sabine, "Grec ancien, latin médiéval, balisage comparé de deux dictionnaires, vers des ressources linguistiques", *Archivum Latinitatis Medii Aevi*, 68 (2010), pp. 161-181.
- Lee, John A.L., "Releasing Liddell-Scott-Jones from its past", en Christopher Stray (ed.), *Classical Dictionaries: Past, Present and Future*, London, Duckworth, 2010, pp. 119-137.
- Rundell, Michael, "'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical", en R.V. Fjeld y J.M. Torjusen (eds.), *Proceedings of the 15th EURALEX Congress*, Oslo, University of Oslo, 2012, pp. 47-92.
- Tarp, Sven, "Lexicographical and Other e-Tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction", en Pedro A. Fuertes-Olivera y Henning Bergeholtz (eds.), *e-Lexicography: The Internet, Digital Initiatives and Lexicography*, Londres/Nueva York, Continuum, 2011, pp. 54-70.
- Tarp, Sven, "Online dictionaries: today and tomorrow", *Lexicographica*, 28 (2012), pp. 253-268.
- Tasovac, Toma, "Reimagining the Dictionary, or Why Lexicography Needs Digital Humanities", en *Digital Humanities*, 2010, <<http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-883.html>> [29/09/2013].



## RESUMEN

Un número considerable de trabajos filológicos impresos han pasado a ser consultables en línea. No obstante, puede afirmarse que este proceso de digitalización no siempre se traduce en la generación de genuinas obras digitales. Como lo indica Sven Tarp, esto es especialmente significativo en el caso de los diccionarios electrónicos, que han de dejar de ser objetos estáticos de consulta para reconceptualizarse como textos dinámicos y diseñados para conectarse con proyectos colaborativos más amplios. ¿Cuáles pueden ser los desafíos a los que se enfrentan los diccionarios para convertirse en recursos propios del ámbito digital? Para abordar esta cuestión, este estudio pretende describir las estrategias que están siendo implementadas en el *Diccionario Griego-Español* recientemente digitalizado (<http://dge.cchs.csic.es/xdge/>). Sugerimos unos caminos que se pueden seguir para dinamizar su contenido y interactuar con fuentes externas (en particular con los corpus en línea). El cambio de formato ha sumido *de facto* el DGE en un entorno cuyos desarrollos en la filología clásica son ya considerables; su papel activo en las aportaciones de la lexicografía en las humanidades digitales dependerá de que se convierta en una obra digital de pleno derecho.

*Palabras clave:* Lexicografía digital, griego antiguo, corpus.

## ABSTRACT

A high number of printed philological works have become available online. However, it can be argued that this process of digitization is not always the guarantee of the creation of genuine digital works. As indicated by Sven Tarp, this issue is especially relevant in the case of electronic dictionaries that should not be static objects of consultation anymore, but should be reconsidered as dynamic texts featured to connect with broader collaborative projects. What might be the challenges that dictionaries face to convert themselves in resources actually belonging to the digital world? To address this question, this study describes the strategies that are being implemented in the newly digitized *Diccionario*

*Griego-Español* (<http://dge.cchs.csic.es/xdge/>). We suggest some ways to follow in order to make its content more dynamic and to interact with external sources (in particular with online corpora). The change of format has plunged *de facto* the *DGE* in an environment whose developments in classical philology are already considerable, its active role to determine how lexicography can contribute to digital humanities remains in its ability to become a digital work in its own right.

*Keywords:* Digital lexicography, ancient Greek, corpus.