

## «LEXIS», DE RICHARD GOULET: UN PROGRAMA PARA LA ELABORACIÓN AUTOMÁTICA DE LÉXICOS DE TEXTOS GRIEGOS Y LATINOS EN MACINTOSH™ (APPLE)

Con el presente trabajo nos proponemos participar a los lexicógrafos españoles, particularmente a los estudiosos de los textos griegos y latinos, la existencia y el funcionamiento básico de un instrumento informático que acomete, con gran automatismo y rapidez, las principales tareas ligadas a la creación y manejo de léxicos. Se trata del programa «Lexis» (versión 2.0), creado por R. Goulet<sup>1</sup> en el marco de la «library»<sup>2</sup> *MacApp (Apple)*. La implantación cada vez mayor entre nosotros del sistema *Macintosh* nos permite suponer que no serán pocos los que podrán beneficiarse del empleo de dicho programa, al tiempo que contribuir a su posible en próximas versiones. Dejando a un lado la sorprendente utilidad y el grado de perfección —como intentaremos mostrar— de esta herramienta, nos resulta gratísima la tarea de su presentación no sólo por la amistad que nos une a la persona del autor sino por tratarse del trabajo de un «amateur» (¡de qué amateur!) a quien no mueve otro interés que el de facilitar los quehaceres del filólogo, partiendo, por supuesto, muchas veces de las exigencias ligadas a sus propias investigaciones.

<sup>1</sup> Richard Goulet es investigador del C.N.R.S. de París, miembro del Equipo 76 («Histoire des doctrines de la fin de l'Antiquité et du Haut Moyen Age»). Realizó su tesis sobre la tradición de la alegoría filosófica judía (*La philosophie de Moïse. Essai de reconstitution d'un commentaire préphilonien du Pentateuque*, colección «Histoire des doctrines de l'Antiquité classique» 11, París 1987), ha editado, traducido y comentado *Cléomède. Théorie élémentaire («De motu circulari corporum caelestium»)*, colección «Histoire des doctrines de l'Antiquité classique» 3, París 1980, y ha publicado numerosos artículos sobre filosofía griega de época tardía. El impresionante proyecto del *Dictionnaire des Philosophes Antiques* que dirige y cuyo primer volumen vio la luz en 1989 (París), siendo de inminente aparición el segundo, hace en realidad poco necesaria su presentación. Este proyecto, por lo demás, ha sido puntualmente reseñado en J. Rodríguez Somolinos, «La lexicografía griega en los últimos años», *E. Clás.* 100, 1991, pp. 83-118.

<sup>2</sup> Una «library» en informática es un conjunto elaborado de funciones y procedimientos a partir de los cuales se pueden crear aplicaciones concretas.

De hecho, R. Goulet comenzó a gestar «Lexis» ante la necesidad de resolver un problema propio muy concreto: la elaboración de un índice de las 51.000 palabras del *Monogenes* de Macario de Magnesia, cuya edición prepara para la colección «Sources Chrétiennes». Este caso de un filólogo clásico metido, de un modo tan absolutamente autodidacta como honesto, en el mundo de la informática de programación es sin duda indicador de un hecho que merece reflexión: que no parece adecuado esperar que los profesionales del ámbito informático se planteen la creación de programas que respondan a las necesidades y a los intereses más específicos de los filólogos, sobre todo cuando entran en juego peculiaridades lingüísticas y metodológicas ineludibles que no están al alcance de los profanos.

Ante tal situación no es de extrañar que quien se ha impuesto a sí mismo esta doble actividad –como Goulet, enfrentándose al reto que suponen el programa de léxico que aquí nos ocupa o un esperado corrector de textos griegos, o el diseño de una variada gama de fuentes de caracteres griegos, ideadas desde el punto de vista de su propio alfabeto, armónicas y fáciles de utilizar– intente transmitir a otros el mismo compromiso e inspirarles el mismo entusiasmo, y quien esto escribe se confiesa tentado al respecto, aunque no muy seguro de ser verdaderamente apto para un eventual servicio de esta índole. En cualquier caso, vale la pena invitar a los colegas, en la medida de sus inquietudes y de sus posibilidades, a contribuir a la creación de una serie de útiles informáticos de calidad específicos para el filólogo clásico, así para la investigación como para la pedagogía.

El trabajo publicado en esta misma revista por un latinista, Manuel Martínez Quintana<sup>3</sup>, nos da buen testimonio de la actualidad del interés que suscita la obtención de un método para la elaboración de concordancias y léxicos por ordenador<sup>4</sup>. En este caso era el autor mismo quien presentaba y describía su obra, un programa para *WordPerfect (IBM)* en cuatro fases: extracción de contextos o «despojo» del texto, concordancias, lematización y ordenación del artículo. En nuestro caso, lo que podemos ofrecer es

<sup>3</sup> «Léxicos con el WordPerfect», *E. Clás.* 100, 1991, pp. 143-165. Véase también del mismo autor «La lematización con el WordPerfect», *E. Clás.* 103, 1993, pp. 91-104.

<sup>4</sup> En general, para la importancia creciente de la informática en la lexicografía griega remitimos a los trabajos de J. López Facal, «Ordenadores y lexicografía griega. El banco de datos», *Introducción a la lexicografía griega*, Madrid 1977, pp. 209-215; F. R. Adrados, «The use of computers in the *DGB*», en *Standardization in computerized Lexicography*, Saarbrücken 1986, pp. 161-170; J. Rodríguez - I. Álvarez, «Informática y lexicografía: la experiencia del Diccionario Griego-Español», *Emerita* 49, 1991, pp. 81-99; y J.A. Ochoa, «Informatic Support to General Greek Lexicography», *Revue, Informatique et Statistique dans les sciences humaines*, 27, 1991, pp. 161-184.

sólo la visión mucho menos experta del usuario, aunque contamos con la inestimable ayuda del autor, y en concreto con un trabajo publicado por él mismo en una revista especializada<sup>5</sup>, aparte de con el Manual detallado de «Lexis» (150 páginas).

En primer lugar, parece indicado comenzar por una exposición sumaria de las capacidades del programa, de las funciones para las que ha sido ideado; a saber: normalización tipográfica y morfológica del texto que se desea tratar (en principio en griego o en latín), filtración (opcional) de las palabras vacías, indicación precisa de la referencia (triple), lematización automática de la mayor parte del vocabulario mediante un diccionario de formas léxicas y una lista de desinencias, lematización manual asistida para el vocabulario restante, incorporación progresiva de los nuevos lemas en el mencionado diccionario, generación de un léxico susceptible de ser impreso de modo inmediato y según diversas opciones de preferencia (posibilidad del formato de intercambio *RTF*: «Rich Text Format»), discriminación mediante símbolos de los términos pertenecientes a citas o producto de conjetura, creación de archivos de frecuencias, de concordancias (incluso de dos términos simultáneamente), búsqueda de cadenas de caracteres, etc.

Como puede verse, «Lexis» intenta conjugar las diversas tendencias y las múltiples exigencias de la lexicografía moderna. La versión que describimos (2.0) permite asimismo una función extra muy de agradecer: la de transformar de modo automático los textos del *Thesaurus Linguae Graecae* (Universidad de California, Irvine) en archivos directamente tratables por «Lexis», con lo que se lleva al extremo la preocupación por simplificar y agilizar el trabajo del usuario<sup>6</sup>.

«Lexis» (unos 447 K) debe ser utilizado, como hemos dicho, con ordenadores *Macintosh*, y funciona en principio en toda la gama. Se recomienda, eso sí, el uso de un disco duro, en la medida en que lo exija la propia naturaleza y envergadura de los trabajos que se deban realizar, la rapidez de cuya ejecución dependerá por

<sup>5</sup> R. Goulet, «Lexis: Un logiciel d'indexation et de lexicalisation automatiques pour Macintosh™ (Apple)», *Revue, Informatique et Statistique dans les sciences humaines*, 26, 1990, pp. 93-110.

<sup>6</sup> Dado que no siempre se puede disponer del *Thesaurus* de Irvine, contando eventualmente, por supuesto, con la correspondiente autorización de explotación, y dado que, por otra parte, éste no ha sido aún completado, sería de gran utilidad si se pudiesen recuperar textos griegos mediante reconocimiento óptico de los caracteres (*scanner*). Según nos comunica Goulet, un programa como *OmniPage* presenta dificultades en el caso de ciertas letras: por ejemplo, siempre que encuentra una ómicron, la convierte en una o romana, cuando, en realidad, ambas letras no están localizadas en el mismo código informático.

supuesto de las propias características del aparato. Goulet explica, por otro lado, en el citado artículo (p. 95) que «Lexis» procura efectuar todas las operaciones importantes en la memoria interna, recurriendo al disco sólo cuando no puede operar de otro modo, de tal suerte que disponer de una memoria viva de 2, 4 u 8 Mb es un factor de reducción del tiempo en la ejecución de las mismas. En fin, el programa resulta del todo compatible con *MultiFinder* y con las versiones del Sistema posteriores a 4.1, incluida la versión 7. En cambio, la posibilidad de hacerlo compatible con un universo informático diferente (*IBM*) parece por el momento difícil, dada su estrecha dependencia con respecto a los recursos de la *ToolBox* del *Macintosh*. Ahora bien, por lo mismo, quienes estén habituados al *interfaz* del usuario característico de este tipo de ordenador (la metáfora del escritorio) pueden tener la certeza de que el programa se maneja con extraordinaria facilidad.

Pasamos por alto precisar los instrumentos concretos de programación que han hecho posible «Lexis», lo que sólo podría interesar a determinados lectores más iniciados. Sin embargo, antes de abordar la descripción propiamente dicha del programa, si son necesarias unas palabras sobre los caracteres griegos que utiliza. En efecto, Goulet, como adelantábamos, ha creado para «Lexis» fuentes de caracteres específicas de formato *PostScript* (una primera llamada *Corinthe* ha sido luego reemplazada por otra de nombre *Eleusis* que se armoniza perfectamente con la fuente *Times*, y a ellas se une una fuente para los caracteres y signos más raros llamada *Critica*), que tienen todas ellas una original peculiaridad frente a las fuentes griegas corrientes: la renuncia a la caprichosa e ingenua correspondencia entre carácter griego y carácter romano, según la cual la *gamma* aparece clasificada como la séptima letra del alfabeto. El hecho de ordenar los diferentes signos griegos en el orden alfabético del griego no responde a un gratuito prurito de filólogo sino que, como veremos, esta lógica redundante en la rapidez de las clasificaciones y de las búsquedas. Y ello no supone tampoco pérdida de nuestros hábitos basados en las analogías sobre el teclado romano. En efecto, es ahora una configuración especial de teclado la que gobierna la correspondencia entre las teclas familiares del teclado material y el código interno de la fuente. En consecuencia, el empleo de estas fuentes va ligado al empleo de una configuración de teclado específica. Todo ello parece complicado, pero *Eleusis* pertenece a un sistema de escritura (*script*) que, trabajando con la versión 7.1 del Sistema, se

encarga de garantizar automáticamente la correspondencia de la fuente seleccionada con su teclado. Cuando no se dispone de dicha versión o cuando se utilizan programas que, como *Word*, no son todavía compatibles con el *Script Manager* seguido por *Éleusis*, hay que recurrir a ciertos artificios, pero que no suponen tampoco una gran complicación. Entonces, en efecto, el paso del teclado estándar (romano) al teclado griego se puede realizar con la mayor rapidez y facilidad por medio de una tecla de función (*FKEY*, normalmente la combinación Comando+Mayúsculas+5), y, más aún, este paso unido al paso de la fuente romana a la fuente griega puede lograrse por medio de una simple macro (activada con las teclas Control+G, por ejemplo). Al respecto, debemos decir que Goulet ha incorporado al programa un útil (*Multiclaviers*) que, desde el marco de los paneles de control, se ocupa del problema de simular diversos teclados, permitiendo seleccionar el teclado especial que sustituirá mediante la tecla de función al teclado estándar. En fin, aunque «Lexis» utiliza hoy la fuente *Éleusis* (+ *Critica*), el usuario no está obligado a trabajar directamente con ella, por cuanto el programa puede importar textos registrados mediante otras fuentes griegas del tipo de *Sparta*.

Y dicho esto damos ya comienzo a la descripción detallada de «Lexis». El punto de partida es, evidentemente, un texto presentado en soporte informático. «Lexis» ha sido programado para el tratamiento tanto de textos latinos como griegos<sup>7</sup>. El programa dispone de un tratamiento de texto propio muy básico, que ignora, por ejemplo, los detalles de los diversos tipos de fuentes de caracteres, tamaños y estilos<sup>8</sup>, detalles todos irrelevantes para sus fines. El documento puede crearse, pues, directamente sobre «Lexis», en la carpeta correspondiente (latín o griego), seleccionando la lengua en el menú de opciones («Options»)<sup>9</sup> y abriendo un nuevo texto en el

<sup>7</sup> A los estudiosos de las lenguas modernas, como el francés, el inglés o el español, les interesará tal vez saber que «Lexis» podría trabajar sobre ellas con sólo ser suprimidas ciertas reglas de normalización morfológica, pero, dada su propia concepción para las lenguas de flexión, quizá resultaría el programa demasiado pesado para las lenguas no sometidas a flexión. Goulet nos indica que parece más lógico pensar en ampliar sus posibilidades con lenguas de flexión, como, por ejemplo, el árabe, lo que exigirla, por supuesto, la colaboración con especialistas de cada una de esas lenguas.

<sup>8</sup> Este procesador de texto interno puede sólo tratar a la vez un máximo de 32.000 caracteres; para pasar a secciones anteriores o posteriores a una dada se disponen de cuatro flechas: inicio, adelante, atrás, fin. Por otro lado, si se prevé tener que añadir ulteriormente texto a un archivo, es posible darle una extensión inferior en principio a los 32.000 caracteres.

<sup>9</sup> Si hemos elegido el griego, a la hora de cerrar el documento ya redactado, habrá que volver al teclado estándar mediante la tecla de función (*FKEY*) para darle nombre.

menú de archivo («Fichier»); o bien, si ya tenemos creado el documento (por ejemplo, en un tratamiento *Word* o *WordPerfect*)<sup>10</sup>, puede ser trasladado, en formato «text only»<sup>11</sup>, a la ventana correspondiente del programa. Si el texto procede del *Thesaurus* presentará ya las líneas de una edición estándar, así como una referencia triple (tomo, página y línea, por ejemplo). De no ser así, habremos de procurar nosotros mismos esta presentación, utilizando el retorno manual para las líneas<sup>12</sup> e introduciendo los números de referencia separados por un signo convenido. Sólo será necesario precisar la referencia en el paso a una nueva página (indicando siempre línea 1), parágrafo o tomo. Se aconseja reservar la mayúscula para los nombres propios. Por otro lado, como adelantamos, es posible dejar constancia de que determinadas palabras o partes del texto proceden en realidad de una cita o del aparato crítico o simplemente que son formas raras o anómalas: se utilizarán para ello signos específicos de apertura y de cierre seguidos respectivamente de un carácter alfanumérico (por ejemplo, el número 9) y del código 0 (valor por defecto). En el léxico final estas palabras serán recogidas con un código 9 entre paréntesis después de la referencia. Si se quiere precisar más el caso (por ejemplo, Homero, ap. crit., conj., etc.)<sup>13</sup>, se podrán utilizar otros códigos aparte del 9 citado<sup>14</sup>, creándose para cada documento un archivo que recoja estas particularidades (extensión «.QO», «Qualificateurs d'occurrence»): en una línea se indicará el carácter alfanumérico elegido y en la siguiente su equivalencia (Homero...)<sup>15</sup>. Como veremos, a la hora de imprimir el léxico final, si se quiere que ta-

<sup>10</sup> Al mencionar el programa *WordPerfect*, debemos precisar que nos referimos propiamente a los textos en latín, puesto que en el caso del griego no nos consta que dicho programa haya incorporado todavía en alguna de sus versiones *IBM* fuentes del tipo *Sparta*, que pudieran ser convertidas en *Éleusis*. Ahora bien, las versiones *Macintosh* de *WordPerfect* pueden utilizar sin dificultad, lo mismo que *Word*, tanto una como otra fuente griega.

<sup>11</sup> Como es sabido, este formato se limita a registrar el documento como archivo de texto *ASCII*, sin otros detalles formales que afecten a los caracteres, a todo el documento, etc.

<sup>12</sup> Éstas no podrán sobrepasar nunca los 128 caracteres; de ser así, lo que sólo sucederá raramente, se deberá desdoblarse la línea con la misma referencia.

<sup>13</sup> Cada precisión está limitada a un máximo de 10 caracteres.

<sup>14</sup> *ASCII* números 33 a 255. Se recomienda utilizar primero las cifras de 0 a 9, luego las letras mayúsculas y minúsculas. Se exceptúan, sin embargo, cuatro caracteres que ya tienen un valor definido en «Lexis».

<sup>15</sup> Puesto que el programa no permite escribir a la vez en caracteres griegos y romanos, contendría, cuando se trate de texto griego, escribir las correspondencias de los códigos en un tratamiento de texto usual, teniendo el cuidado de utilizar siempre la fuente y el teclado de *Éleusis* para los códigos. Con todo, en el caso de que se prefiera utilizar el tratamiento interno de «Lexis», bastará con actuar como si se tratase de un documento en latín, utilizando la *FKEY* para poner en funcionamiento el teclado estándar y tecleando a ciegas la correspondencia del código.

les precisiones aparezcan en el mismo, se podrán sustituir automáticamente los códigos por su equivalencia activando la casilla correspondiente («développer qualificateurs d'occurrence») en el cuadro de preferencias de presentación del léxico, accesible desde el menú «Options». En fin, existen también signos especiales de apertura y de cierre si se quieren añadir al texto títulos, comentarios o notas que no deberán aparecer en el léxico.

En nuestra descripción nos vamos a ceñir de modo particular al caso de los textos griegos, más complicado y que suponemos reviste mayor interés. El programa, por otro lado, ha sido ya suficientemente experimentado con textos de esta lengua, hasta el punto de que «Lexis» se distribuye hoy con un diccionario de alrededor de 30.000 formas léxicas griegas.

Pues bien, una vez dispuesto el archivo de texto en sus líneas (conviene darle una extensión como «.TO»), «Lexis» seguirá para la obtención del léxico correspondiente un proceso que utiliza básicamente dos instrumentos: 1) un archivo principal que comporta las líneas (normalizadas) con su referencia así como una lista de las formas con sus lemas y, en su caso, sus códigos especiales; 2) un diccionario que hace posible la lematización automática e incorpora los resultados de la lematización manual. El final del proceso es un archivo con el léxico resultante (extensión «.lexique»), que se guardará por defecto en un formato *RTF*.

La primera operación ejecutada por «Lexis» sobre el documento .TO es la creación de un «índice» («indexation») que engloba el conjunto de los datos (lemas, formas, referencias, etc.) a partir de los cuales se generará el léxico o la concordancia. Basta abrir un nuevo archivo (menú «Fichier») al que se le dará el nombre del documento en cuestión sin la extensión «.TO»<sup>16</sup>, elegir «Indexer» en el menú de trabajo («Travail») y seleccionar ahora el documento .TO. Esta operación conlleva en primer lugar una depuración y una normalización tipográfica del texto: detecta las líneas con más de 128 caracteres, suprime los dobles espacios en blanco, ignora los comentarios introducidos por el usuario, separa con un espacio las formas elididas, elimina los acentos de enclisis, uniforma los dobles (u y v en latín, β y β, φ y φ, θ y θ en griego), reemplaza las marcas de tabulación por espacios en blanco, restablece el apóstrofo allí donde se ha teclado erróneamente el espíritu suave, etc. Después «Lexis» anali-

<sup>16</sup> Este archivo servirá de guía y de balance de las sucesivas operaciones.

za las líneas: siempre que localiza una forma léxica, transforma los barítonos en oxítonos, registra el código correspondiente si la palabra forma parte de una sección marcada por un código («qualificateur d'occurrence»), y, por último, indica el cómputo de las líneas registradas, así como de las palabras retenidas y eliminadas en la operación. En efecto, uno de los puntos del cuadro de preferencias (menú «Options») de «Lexis» permite justamente la elección de un índice selectivo que ignore de entrada, por ejemplo, las diversas formas del artículo, así como las formas *καί*, *μέν* y *δέ*. Esta lista puede modificarse o incrementarse fácilmente, según el grado de exhaustividad con que se desee presentar el léxico.

Toda la operación que acabamos de describir, pese a su apariencia complicada, se realiza en un tiempo verdaderamente breve: por poner un ejemplo extremo, en un ordenador SE 30, el tratamiento de unas 6.000 líneas (o unas 50.000 palabras) requiere menos de 10 minutos. La marcha de la operación se sigue a través de un diálogo en la forma habitual de «termómetro» y su culminación viene indicada por un mensaje acompañado de un acorde musical.

Antes de proceder a la lematización del archivo de las formas léxicas registradas conviene clasificarlo alfabéticamente (operación «trier formes» en el menú de trabajo), lo que en el caso del griego supone enfrentarse con un problema metodológico<sup>17</sup>: el cierto desorden que ocasionan los acentos en la clasificación de las palabras, al no ser ordenadas igual las mismas secuencias cuando intervienen acentos que cuando no<sup>18</sup>. En este sentido, Goulet, partiendo del principio de sólo tomar en consideración los acentos de las palabras cuando se ha ensayado primero su clasificación con independencia de la acentuación, ha ideado un sistema que baraja, en su versión más compleja, cinco claves: lema sin acento, lema acentuado, forma sin acento, forma acentuada y número de la línea. Así, «Lexis» clasificará de entrada los lemas sin tener en cuenta los acentos; si dos lemas son idénticos, los clasificará teniendo en cuenta los acentos; si siguen siendo idénticos, los clasificará según la forma léxica; por último, si los lemas y las formas léxicas son idénticos, recurrirá a la quinta clave: el número de orden de la referencia en el texto. El al-

<sup>17</sup> El hecho de que en la fuente *Éleusis* la *gamma* sea la tercera y no la séptima letra del alfabeto evita ya el inconveniente que supondría esto último a la hora de la clasificación alfabética de las palabras griegas.

<sup>18</sup> Así tendríamos en principio, por ejemplo, la serie *αὐταί*, *αὐταῖς*, *αὐτά*, *αὐται*, etc.; por cuanto todas las *α* pasarían delante de las *ά*, todas las *ὀ* delante de las *ϋ*.

goritmo es complicado, pero una vez más el usuario sólo es consciente de la rapidez de la operación: por seguir con un ejemplo extremo, para la clasificación alfabética de unas 30.000 palabras, «Lexis» apenas requiere un minuto en un ordenador SE 30.

El paso siguiente consiste ya en un intento de lematización automática: (apartado «lemmatiser auto» en el menú de trabajo). Se trata de la parte mecánica del proceso, que luego será corregida y completada manualmente. En este punto es donde el programa recurre al diccionario (ampliable) del que ya hemos hablado. «Lexis» actúa como sigue: mediante un procedimiento de búsqueda dicotómica, busca en este diccionario cada forma registrada y, si la encuentra, introduce directamente en el archivo de las formas el lema que para esa forma da el diccionario. Si, gracias a un indicador de homonimia, el programa reconoce que la forma puede remitir a varios lemas, se limita, prudentemente, a indicar como lema una cruz (†). Por otro lado, si, según el diccionario, la forma es desconocida, «Lexis» recurre a otro instrumento: una tabla que empareja una serie de terminaciones básicas con las desinencias que les corresponderían en el lema (nominativo o infinitivo)<sup>19</sup>. El programa compara las diferentes desinencias con la terminación de la forma en cuestión y, si hay coincidencia, sustituye la terminación de la forma por la correspondiente al lema que da la tabla de desinencias. Para facilitar aquí también la verificación manual, «Lexis» marcará el lema así resultante con un asterisco (\*), como llamada de atención sobre su corrección. He aquí algún ejemplo de lematización errónea a partir de la tabla de desinencias: δοκοῦσις\* como lema de δοκοῦσιν ο πράττουσις\* de πράττουσιν. En este caso, el usuario deberá, pues, rectificar los lemas en δοκεῖν y πράττειν, respectivamente, mientras que para confirmar una lematización dada bastará eliminar el asterisco. En fin, si una forma registrada es idéntica a la precedente, entonces el programa, lejos de reiniciar la búsqueda, se limitará a repetir el lema ya encontrado. De ahí el interés de la previa ordenación alfabética de las formas.

Haciendo de nuevo balance, el diccionario de griego que dispone ya «Lexis», con unas 30.000 palabras, debería poder asegurar la lematización automática del 70 al 80 % de cualquier texto en

<sup>19</sup> Esta tabla es susceptible de ser modificada o ampliada del mismo modo que la lista de palabras eliminables o el diccionario.

prosa. El resto del trabajo exige, por supuesto, la colaboración experta del filólogo, quien tendrá que comprobar la corrección de los lemas propuestos por «Lexis» de modo automático así como encontrar él mismo el lema de las formas no lematizadas. Pues bien, esta lematización manual se realiza en una ventana de trabajo<sup>20</sup> que recoge en diversos compartimentos todos los componentes del proceso (véase la figura I del anexo): arriba presenta una zona de edición donde se inscribe la forma que en cada momento se está tratando y donde se puede también proponer la búsqueda de una forma determinada a lo largo del texto; en el nivel siguiente, la ventana presenta tres columnas, que contienen en celdas, respectivamente, los lemas, las formas y los códigos específicos («qualificateurs d'occurrence»); por último, en el tercer nivel, presenta una primera columna con la triple referencia numérica y una segunda con varias líneas del contexto correspondiente a la forma en cuestión<sup>21</sup>. Desde esta ventana son accesibles las funciones de edición típicas del *Macintosh*: anular, copiar, cortar, pegar, borrar... El paso de una celda a otra (mediante las flechas de desplazamiento, el ratón, el retorno de línea o *Enter*) confirma inmediatamente el contenido de aquélla. Por último, desde la misma ventana se pueden realizar otras tres funciones: 1) insertar en el diccionario una o varias correspondencias lema-forma<sup>22</sup>; 2) añadir una determinada forma al archivo de formas eliminables; 3) abrir el diccionario para su corrección (véase la figura II del anexo).

Una vez finalizada la lematización manual, conviene proceder a una nueva clasificación alfabética del documento: «trier lemmes et formes» (o «trier lemmes sans formes», según la preferencia elegida) en el menú de trabajo. Hecho esto, sólo resta ya crear el léxico final («créer lexique» en el mismo menú). El archivo, como dijimos, aparecerá por defecto con un formato *RTF*, que permitirá un uso inmediato sobre tratamientos de texto profesionales como *Word* o *WordPerfect*. Ciertas peculiaridades de presentación podrán ser predeterminadas en el cuadro de preferencias («Préférences» en el menú «Options»), y ante todo se podrá elegir si se desea un léxico completo, de lemas y formas, o sólo de lemas. Destaca-

<sup>20</sup> Para hacerla aparecer basta seleccionar «éditer formes» en el menú de trabajo.

<sup>21</sup> El número de líneas dependerá de la altura de la pantalla: pueden ser seis o incluso más.

<sup>22</sup> Si ya está presente en el diccionario una determinada correspondencia, ésta no se grabará; si está presente la forma pero con otro lema, un indicador de homonimia pasará de 0 a 1 y el lema no será propuesto en la lematización automática para evitar posibles errores (se indicará con † la homonimia).

mos, en particular, asimismo la posibilidad de reagrupar los nombres propios, o de dejar constancia con mayor o menor precisión de los casos de cita (Homero...), aparato crítico (conjetura...), «háπαξ», etc., así como la posibilidad de añadir al final de cada lema las formas compuestas en las que éste reaparece (por ejemplo, γράφειν -> ἀναγράφειν, esse -> adesse) e, inversamente, en el caso de un lema precedido de preposición, la forma simple y la lista de los compuestos de la misma forma simple (por ejemplo, ἀναγράφειν -> γράφειν, ἐπιγράφειν, adesse -> esse, prod-esse). Al respecto, dado que la lematización automática puede hacer que se deslicen falsos compuestos (por ejemplo, *laudare* bajo *dare*), habrá de nuevo que estar atento a ello en la fase manual. Con todo, para evitar la proliferación de este tipo de gazapos, el programa sólo aplicará el procedimiento en el caso de lemas de más de tres letras (así, no encontraremos nunca γράφειν bajo εἶ). Ofrecemos, en fin, un ejemplo de léxico resultante de «Lexis» en el anexo.

Hasta aquí lo referente propiamente a «Lexis» como instrumento para la elaboración de un léxico, pero el programa ofrece otras posibilidades, de entre las cuales no podemos dejar de reseñar algunas<sup>23</sup>. En primer lugar, «Lexis» puede realizar determinadas búsquedas en un archivo (incluido el diccionario): puede, por ejemplo, buscar todas las palabras que terminan por un grupo de letras o incluso todas las palabras que presentan una secuencia determinada de letras, mediante el uso de una serie de símbolos que permiten, por ejemplo, marcar como facultativos ciertos caracteres. Como Goulet señala en su manual (p. 84), esta función se revela todavía más útil que un índice inverso, por cuanto no limita su búsqueda a la terminación de las palabras: el autor imagina, con toda razón, la rentabilidad que un epigrafista o un papirólogo pueden extraer de este instrumento, siempre, por supuesto, que se vayan grabando en el diccionario todas las nuevas formas léxicas halladas.

El programa es capaz, por otro lado, de generar una concordancia («Concordance» en el menú de trabajo): se reunirán así todas las apariciones a lo largo del texto, con un contexto de tres líneas, de un lema elegido o, mejor incluso, de dos lemas simultáneamente. El archivo resultante puede ser modificado, grabado y,

<sup>23</sup> Apuntemos además que «Lexis» puede generar un archivo con todos los lemas de un texto en orden alfabético, fusionar el contenido de dos archivos, así como de dos diccionarios —no de otro modo se ha ido ampliando el ya importante diccionario de griego— o comprimir (y luego, por supuesto, des-comprimir) los archivos (incluidos los diccionarios) hasta un tercio de sus dimensiones reales.

en consecuencia, impreso de forma inmediata siempre que contenga 80 entradas o menos. Si el programa ha localizado más de 80 pasajes que responden a los criterios de la búsqueda, los límites del tratamiento de texto interno no permitirán su visualización inmediata y entonces «Lexis» propondrá la creación de la concordancia en un archivo de texto que luego podrá ser abierto desde cualquier tratamiento de texto, incluido el interno del programa.

Por último, «Lexis» está preparado para la creación de un archivo de frecuencias («créer fichier de fréquences» en el menú de trabajo), que recogerá en una lista estándar todos los lemas del texto según su frecuencia de aparición. Esta lista puede ser grabada (extensión «.fréquences») e impresa. Asimismo, si así se indica en el menú de preferencias, los índices de frecuencia pueden ser añadidos al léxico final, donde aparecerán entre corchetes: por ejemplo, πολύς [4 : 0.0444]<sup>24</sup>.

En fin, hemos intentado a lo largo de este trabajo describir con cierto detalle las capacidades y el funcionamiento de lo que hemos considerado más interesante de «Lexis», conscientes, por supuesto, de que sólo el uso y la práctica podrán hacer el resto. Esperamos al menos haber conseguido llamar la atención de los estudiosos españoles de las lenguas griega y latina sobre este instrumento filológico que en verdad estimamos utilísimo. Y subrayamos lo de instrumento, por cuanto creemos haber puesto también de relieve que la rentabilidad del programa y la perfección de sus productos no podrá sino ser siempre obra de la pericia y erudición del filólogo. Nos consta, por otro lado, la importancia y el valor que Richard Goulet concede a la experiencia y a la crítica de los usuarios con vistas a la mejora del programa en versiones sucesivas.

PEDRO PABLO FUENTES GONZÁLEZ  
*Universidad de Granada*

<sup>24</sup> Según la técnica común entre los lingüistas, el primer número representa las apariciones de un lema en un texto dado; el segundo, el número medio de apariciones del lema por 10.000 palabras.

## Anexo

Figura I: Ventana para la corrección de la lematización automática

Eunape			
Αἴγυπτος			
D: 0	Lemmes	Formes	
Αἴγυπτος	+	Αἴγυπτος	0
Αἴγυπτος		Αἴγυπτος	0
Λυκώ		Λυκώ	0
Πλατωνικός		Πλατωνικός	0
Πλωτίνος		Πλωτίνος	0
Πλωτίνος		Πλωτίνου	9
Πορφύριος		Πορφύριος	0
Πορφύριος		Πορφύριος	0
3, 5, 18	Πλωτίνος ἦν ἐξ Αἴγυπτος φιλόσοφος, τὸ		
3, 5, 19	ἐξ Αἴγυπτος νῦν γράφων, καὶ τὴν πατρίδα προσθήσω.		
3, 5, 20	Λυκῶ ταύτην ὀνομάζουσιν· καίτοι γε ὁ θεσπέσιος φιλόσοφος		
3, 5, 21	Πορφύριος τοῦτο οὐκ ἀνέγραψε, μαθητῆς τε αὐτοῦ γεγενησθῶ		
3, 5, 22	λέγων, καὶ συνεσοχολακῆναι τὸν βίον ἅπαντα ἢ τὸν		
3, 5, 23	πλείστον. Ἔσθουτου Πλωτίνου θερμοὶ βωμοὶ νῦν, ἴθαι τὰ		

Figura II: Ventana de trabajo y listas de inserción.

 Fichier Edition Travail Dictionnaire Outils Options

Eunape	
γε	
D: 5	Formes
βιβλίον	βιβλία
βιβλίον	βιβλίων
βίος	βίον
βίος	βίον

  

Insérer dans le dictionnaire	
Αἴγυπτος - Αἴγυπτος - 0	
Λυκῶ - Λυκῶ - 0	
Πλατωνικός - Πλατωνικός - 0	
Πλωτίνος - Πλωτίνος - 0	
Πλωτίνος - Πλωτίνου - 9	

  

A éliminer	
αἰ	
αὐτά	
αὐτοῦ	+
γε	
δ'	
δέ	

Ejemplo de léxico resultante de «Lexis»:  
Eunapio, *Vida de Plotino*

<b>Αἴγυπτος</b>		<b>βίος</b>	
Αιγύπτου	3, 5, 18; 3, 5, 19	βιον	3, 5, 22; 3, 6, 3
[2 : 0.0222].		[2 : 0.0222].	
<b>Λυκώ</b>	3, 5, 20	<b>βωμός</b>	
[1 : 0.0111].		βωμοί	3, 5, 23 (Cita)
<b>Πλατωνικός</b>		[1 : 0.0111].	
Πλατωνικούς	3, 6, 2	<b>γε</b>	3, 5, 20
[1 : 0.0111].		[1 : 0.0111].	
<b>Πλωτίνος</b>	3, 5, 18	<b>γίγνεσθαι</b>	
Πλωτίνου	3, 5, 23 (Cita)	γεγενῆσθαι	3, 5, 21
[2 : 0.0222].		[1 : 0.0111].	
<b>Πορφύριος</b>	3, 5, 21; 3, 6, 4	<b>γράφειν</b>	
[2 : 0.0222].		γράφων	3, 5, 19
<b>ἀλλά</b>	3, 6, 2; 3, 6, 5	[1 : 0.0111] -> ἀναγράφειν.	
[2 : 0.0222].		<b>διά</b>	3, 6, 1
<b>ἀναγράφειν</b>		[1 : 0.0111].	
ἀνέγραφε	3, 5, 21	<b>δόγμα</b>	
[1 : 0.0111]-> γράφειν.		δογμάτων	3, 6, 3
<b>ἅπας</b>		[1 : 0.0111].	
ἅπαντα	3, 5, 22	<b>ἐάν</b>	3, 6, 2
[1 : 0.0111].		[1 : 0.0111].	
<b>αὐτός</b>		<b>εἶναι</b>	
αὐτά	3, 6, 3	ἦν	3, 5, 18; 3, 6, 4
αὐτοῦ	3, 5, 21; 3, 6, 3; 3, 6, 5	[2 : 0.0222].	
[4 : 0.0444].		<b>εἰσφέρειν</b>	3, 6, 4
<b>βιβλίον</b>		[1 : 0.0111].	
βιβλία	3, 6, 1	<b>ἐκφέρειν</b>	
βιβλίων	3, 6, 5	ἐξήνεγκεν	3, 6, 4
[2 : 0.0222].		[1 : 0.0111].	
		(...)	