

Delenda est machina? Informática y Filología Latina

MANUEL LÓPEZ-MUÑOZ¹
Universidad de Almería

A la Collection Budé, resurgida de las cenizas gracias a la informática

Resumen: En el presente artículo, se intenta mostrar que la interacción de la informática y la Filología Latina va más allá de la mera utilización de *software* ofimático. La Lingüística computacional podrá beneficiarse de la Filología Latina, que le suministra todo un *corpus* cerrado. De la Lingüística computacional puede esperarse sobre todo que proporcione datos para comprender mejor algunos problemas concretos. Las técnicas de trabajo colaborativo son, sin duda, el campo más prometedor, y su uso puede constituir un punto de inflexión en los métodos tradicionales de nuestra materia. Vinculada a la idea de colaboración está la de trabajo en red, con Internet como ejemplo. También se trata la cuestión de la utilización de la informática para la docencia cotidiana. Por último, se proponen algunos campos de desarrollo futuro que integren lo visto.

Palabras clave: *Filología; Latín; Informática.*

Summary: This article intends to prove that the interaction of informatics and Latin philology goes far ahead from the simple use of ofimatic software suites. Computational linguistics may be expected to benefit from Latin Philology, which offers a whole closed *corpus*. Not so many things are to be expected from Computational linguistics apart from empirical data for a better understanding of some punctual issues. The collaborative work techniques are no doubt the most promising area insofar their adoption would become a turning point in the traditional methodology of our studies. Close to the idea of collaboration is that other of networking, with Internet as a paradigm. Another aspect to be considered is the use of software for daily teaching. The last part of the paper is some kind of proposal of future developments combining both informatics and Latin Philology.

Key words: *Classics; Latin; Informatics.*

¹ Se hacen constar expresamente agradecimientos al Ministerio de Ciencia y Tecnología y al Fondo Europeo de Desarrollo Regional (FEDER), financiadores del Proyecto «CVMAS» (Catálogo Unificado de Materiales de Filología Latina y Derecho Romano en Internet), referencia BFF2002-00028, en el que

0. CONSIDERACIONES GENERALES

Tratar las relaciones entre Informática y Filología Latina es un asunto espinoso. No procede entrar en la cuestión de si es o no aceptable identificar unívocamente las disciplinas humanísticas y sus profesionales con el alejamiento respecto de los desarrollos de la informática, cosa que ya en otras ocasiones ha sido abordada², máxime si se tiene en cuenta la progresiva generalización del uso de los ordenadores como herramienta de nuestro trabajo, así como que los nuevos desarrollos tecnológicos no son de obligada contemplación. En cualquier caso, nos encontramos ante un asunto intrínsecamente complejo, muy probablemente provocado por la difícil relación que, para el profesional de la Filología Latina, se establece entre un objeto de estudio tan estático como dinámicas son estas herramientas de trabajo. Y a buen seguro que de aquí procede algo de esa alergia tecnológica a la que antes hacía alusión.

Una catalogación de recursos y software sería un modo bastante sencillo de abordar las relaciones de la Filología Latina y la Informática, pero su contrapartida es la desactualización de los contenidos de este artículo en un plazo de tiempo excesivamente corto. Las tecnologías informáticas cambian con un ritmo tan acelerado que el programa que hoy empezamos a manejar será una curiosidad de museo en el próximo lustro³. Tampoco es procedente abordar la cuestión desde el punto de vista de la confrontación de modelos sociales (la sociedad tradicional frente a la pomposamente llamada sociedad del conocimiento) o incluso de estructura de la propiedad de las herramientas informáticas: las discusiones entre partidarios del *software* propietario y defensores del *software* libre, aun siendo relevantes para la configuración próxima de las industrias tecnológicas, son más propias ahora mismo de informáticos que de usuarios de la informática. Entre un enfoque y otro, se ve más adecuado proceder por la vía in-

se inscribe el artículo y del que es investigador responsable su autor. Igualmente, se hace constar el agradecimiento al Grupo de Investigación «El legado de la Antigüedad», de la Universidad de Almería (Consejería de Educación y Ciencia, Junta de Andalucía, HUM800), en cuyo seno están el mencionado Proyecto y el autor.

² Cf. R. MORDENTI, «Filologia e computer», C. Leonardi - M. Morelli - F. Santi (eds.), *Macchine per leggere. Tradizioni e nuove tecnologie per comprendere i testi*, Spoleto, Centro Italiano di Studi sull'Alto Medioevo, 1994, p. 54: «...non si può immaginare niente di più sgradevole per la mentalità di qualsiasi filologo (in senso lato: cioè studioso di materie cosiddette umanistiche), e non solo per (pure legittimi) aspetti di conservatorismo psicologico, ma per motivi assai più profondi e radicali che ineriscono al cuore stesso della nostra formazione disciplinare: essa consiste essenzialmente nell'abitudine a *sedimentare* conoscenze, a *raccogliere* informazioni, ad *addestrarci* a procedure di ricerca, a *collezione* dati, ed a *collazionarli* (cioè a metterli in rapporto fra loro) così che tutto ciò che interviene e modificare, anzi a sconvolgere, il *criterio* di organizzazione dei dati su cui lavoramo e gli stessi procedimenti del nostro lavoro, ci turba profondamente, ci colpisce in un punto vitale e delicatissimo».

³ Cf. L. C. PÉREZ CASTRO, *Informática y Filología (ordenadores PC)*, Madrid, Ediciones Clásicas, 1991. Se trata de un trabajo interesante, sobre todo cuando habla de cómo preparar los textos para su manejo a modo de (o en combinación con) bases de datos. No obstante, las páginas que dedica a la descripción del ordenador o el sistema operativo quedaron anticuadas prácticamente en el momento de su publicación.

termedia, verbigracia, analizando sobre todo el valor que las herramientas informáticas pueden aportar al trabajo filológico. En este sentido, resulta ser especialmente afortunada la reflexión de Iso Echeгойen cuando afirma que los ordenadores no suponen un aumento de la calidad de los resultados de la investigación filológica, sino más bien la oportunidad de que el filólogo se libere de tareas ingratas de mecanización y adquisición de datos y dedique sus esfuerzos al análisis de esos mismos datos⁴.

Afirmar que la cuestión que se trata aquí constituye una cierta novedad en nuestra disciplina no es un fácil mecanismo de *captatio benevolentiae*, pero tampoco una verdad incontestable. Ya en 1976 se empezaron a publicar algunos trabajos sobre Informática y Filología (no sólo latina), aun cuando encontramos que, de un lado, en muchas ocasiones se trata de descripciones de aplicaciones específicas y, de otro lado, no se encuentran actualizaciones ni revisiones críticas del panorama general. A modo de presentación del estado de la cuestión, es fuerza mencionar a autores como Fernández Galiano y López Facal⁵, Mariner⁶, Marinone⁷, Martínez Quintana⁸, o Rodríguez Adrados y Rodríguez Somolinos⁹; frente a esto, encontramos que en obras tan recientes como la *Introducción en la Filología Latina* coordinada por Fritz Graf¹⁰ ni se menciona la posibilidad de que la informática pueda tener nada que ver con nuestros estudios. Lo más cercano a estado de la cuestión viene a ser la obra colectiva *Macchine per leggere*¹¹..., coordinada por Leonardi, Morelli y Santi, y aquejada de un desigual balanceo de carga entre capítulos interesantes y otros simplemente ajenos al tema.

⁴ Cf. J. J. ISO ECHEGOYEN, «Sobre una base de datos métricos automatizada», en J. Luque Moreno - P. R. Díaz y Díaz, *Estudios de Métrica latina*, Granada, Servicio de Publicaciones, 1999, 2 vols.: «... quienes se mantienen distantes de la informática como auxiliar de la investigación aduciendo la complejidad de las ciencias humanas, la abismal distancia que éstas mantienen respecto al pensamiento y método binario, no se dan cuenta de que, aun hoy en día, no sólo la informática no impide el pensamiento científico en nuestro ámbito, sino que nos evita ese tedioso manejo del acarreo y clasificación de los datos; dicho de otro modo, la facilidad en la obtención de datos jubila inexcusablemente a esos practicones, a esos *operarii* de la filología y nos hominiza aún más, nos obliga a utilizar nuestra capacidad mental, no en sus niveles más bajos, los computadorios, sino en esa difícil tarea de combinar la imaginación y creatividad con la tozudez de los hechos, de establecer un fecundo y complejo diálogo entre deducción e inducción lógicas».

⁵ E. FERNÁNDEZ GALIANO - J. LÓPEZ FACAL, «Panorama general de los tratamientos por ordenador en Filología y Lingüística griega y latina», *Revista de la Universidad Complutense* 25 (1976), pp. 101-112.

⁶ S. MARINER BIGORRA, «El estudio de la métrica latina con ayuda de ordenador», *Revista de la Universidad Complutense* 25 (1976), pp. 179-189.

⁷ N. MARINONE, «Lessico latino e analisi elettronica», *Voces* 1 (1990), pp. 23-28.

⁸ M. MARTÍNEZ QUINTANA, «Procedimiento técnico para la realización de 'índices verborum' en microordenador», *EClas.* 30 (1988), pp. 93-107; «La lematización con el WordPerfect», *EClas.* 35 (1994), pp. 91-104. Al respecto de este último trabajo citado, fuerza es señalar que, bien por errata, bien y probablemente por impericia propia, el procedimiento descrito en este artículo no funcionó.

⁹ F. RODRÍGUEZ ADRADOS - J. RODRÍGUEZ SOMOLINOS, «The TLG data bank, the DGE and Greek Lexicography», *Emerita* 62 (1994), pp. 241-251.

¹⁰ F. GRAF (ed.), *Einleitung in die lateinische Philologie*, Stuttgart, Teubner, 1997.

¹¹ *Op. cit.*

1. FILOLOGÍA LATINA Y LINGÜÍSTICA COMPUTACIONAL

Una cuestión de obligado tratamiento al abordar las relaciones de Informática y Filología Latina se suscita al recurrir a los trabajos de la Lingüística computacional, disciplina relativamente nueva en la que¹²: «...es posible agrupar un conjunto relativamente heterogéneo de teorías, métodos, herramientas, aplicaciones y productos que tienen en común la consideración de la lengua como un objeto susceptible de ser tratado mediante procedimientos informáticos. Adoptando una perspectiva muy amplia, entrarán dentro de este ámbito desde las ayudas proporcionadas por los ordenadores a la investigación lingüística en disciplinas como la filología, la estilística, la estadística lingüística o la lexicografía hasta los sistemas capaces de llevar a cabo automáticamente la traducción de un texto o de ofrecer servicios telefónicos sin la intervención de un operador humano, pasando por los correctores que habitualmente se integran en los programas de tratamiento de texto o por los sistemas de recuperación automática de la información que hacen uso de datos lingüísticos».

Un somero recorrido de los campos mencionados como propios de la Lingüística computacional hace posible afirmar que gran parte de ellos queda lejos de ser aplicable a la Filología Latina, excepción hecha de aquellos desarrollos que permiten liberarse de las servidumbres del acopio de datos y su tratamiento estadístico. Procedimientos y herramientas como los que permiten desarrollar análisis estilísticos computerizados, escanores, o recogidas de testimonios mediante técnicas de minería de datos (*data mining*) son interesantes, y tanto más cuanto más se automatice toda la tarea previa a la reflexión del filólogo; los sistemas de traducción automática o de generación de textos son de poco interés en su estado actual de evolución; los correctores ortográficos y gramaticales tienen un valor limitado, salvo en las fases iniciales de un trabajo de edición textual.

Ahora bien, al tomar en consideración, no ya lo que aporta la Lingüística Computacional, sino lo que la Filología Latina puede proporcionarle, es posible afirmar que existe una vía de transmisión y de colaboración muy interesante. El carácter eminentemente acabado de la lengua latina permite trabajar con ella como un banco de pruebas perfecto para los lingüistas computacionales, cuyos esfuerzos actualmente se ven frenados por la difícil disponibilidad de *corpora* cerrados, que todavía están incompletos para la mayor parte (si no para todas) de las lenguas habladas. Los textos latinos, en tanto que *corpus*, están recopilados, depurados de errores, sistematizados y en buena medida automatizados, aunque todavía no pueda afirmarse lo mismo de los periodos medio y neo-latino.

Por su parte, la estructura morfosintáctica del latín es un factor que facilita los trabajos de los lingüistas computacionales: existen repertorios léxicos y, sobre

¹² T. MOURE - J. LLISTERRI, «Lenguaje y nuevas tecnologías. El campo de la lingüística computacional», M. Fernández Pérez (coord.), *Avances en Lingüística aplicada*, Santiago de Compostela, Servicio de Publicaciones e Intercambio Científico, 1996, p. 147.

todo, repertorios de los morfemas del latín. La tarea de diseñar analizadores morfosintácticos (*parsers*¹³) ya ha comenzado, y con una cierta fiabilidad en sus resultados. Sólo queda experimentar en profundidad la posibilidad de incorporar los elementos de descripción semántica¹⁴ y de conocimiento del mundo¹⁵ que hacen falta para el diseño de sistemas expertos capaces de producir análisis computacionales del latín. Nuevamente, el carácter cerrado de los textos latinos es una sustancial ventaja: la práctica imposibilidad de aparición de significados nuevos hace factible el establecimiento de redes cognitivas depurando y actualizando las tareas de delimitación y lematización que ya desde hace mucho tiempo se vienen aplicando en la elaboración de herramientas tan típicamente filológicas como los diccionarios, léxicos y concordancias.

Corpus estable, ausencia de evolución lingüística, tradición en la descripción gramatical, elementos ya informatizados casi por completo... La lengua latina cubre prácticamente todos los requisitos necesarios para que el lingüista computacional logre avanzar en su trabajo y comprobar si sus teorías se validan o no por recurso a una lengua completa.

Sensu contrario, si la Lingüística computacional es capaz de producir modelos que expliquen el latín, modelos epistemológica y heurísticamente adecuados¹⁶, podrá también ayudar en la verificación de perspectivas, descripciones o

¹³ Algunos ejemplos disponibles en www.levity.com/alchemy/latin/latintrans.html, citeseer.nj.nec.com/covington90dependency.html, y en community.middlebury.edu/~harris/SubIndex/classics.sub.html. Por simplificar la grafía, cada vez que citemos una URL, lo haremos eliminando el recurrente <http://> que siempre encabeza la dirección.

¹⁴ Cf. T. MOURE - J. LLISTERRI, *op. cit.*, pp. 190-191: «Una vez conseguido el análisis sintáctico, el tratamiento automático de las lenguas exige un análisis semántico, que asocie cada secuencia a una representación interna de su contenido. El procedimiento más habitual en la construcción de sistemas de este tipo pasa por separar la información gramatical de la meramente semántica en bases de datos independientes y, a la vez, independientes del algoritmo de análisis, para que sea más fácil hacer correcciones o añadir información nueva. Sin embargo, las distintas fuentes de información deben interactuar cuando el programa está corriendo.» Una interesante panorámica de la cuestión en M. ROSNER - R. JOHNSON (eds.), *Computational Linguistics and formal Semantics*, Cambridge, Cambridge University Press, 1992, así como en el recentísimo trabajo de R. Mitkov, *The Oxford Handbook of Computational Linguistics*, Oxford, Oxford University Press, 2003, aparecido mientras este artículo se encontraba en prensa casi.

¹⁵ Cf. R. GRISHMAN, *Introducción a la lingüística computacional*, Madrid, Visor, 1991, (= *Computational linguistics*, Cambridge University Press, 1986, trad. A. Moreno Sandoval), p. 167: «Una vez que hemos elegido un formalismo lógico ... podemos tomar todo nuestro conocimiento del mundo y codificarlo en este formalismo como una amplia relación de teoremas. A medida que se van recibiendo las oraciones de un texto, es posible traducirlas a este formalismo ... Podemos usar entonces algún comprobador de teoremas poderoso para extraer las inferencias necesarias (por ejemplo para eliminar algunas interpretaciones semánticas mostrando que contradicen nuestro conocimiento del mundo) ... Es difícil imaginarse un sistema de inferencias (o un cerebro humano) abriéndose camino entre cientos de miles de teoremas no organizados. Por tanto, el principal desafío del análisis del texto, tal y como se ha concebido desde principios de los 70, es identificar y organizar el conocimiento del mundo necesario para que pueda resolver con éxito los problemas del análisis del texto.»

¹⁶ Y. WILKS, «Form and content in semantics», M. ROSNER - R. JOHNSON, *op. cit.*, p. 259: «An *epistemologically adequate* model is one such that a solution to a given problem follows from the model. A *heuristically adequate* model is one which provides, in addition, a method for finding that solution».

teorías. Es pensable, por ejemplo, el diseño de un sistema que, aplicando la teoría laringal en forma de algoritmos y cruzándola con los datos léxicos de *toda* la lengua latina (la identificación del autor ya permitiría incluso establecer la prelación cronológica de las formas de una misma palabra), nos indique si la teoría es correcta o incorrecta, si hay o no elementos no válidos, los límites de su aplicabilidad... y no sólo con razonamientos *ad hoc* y series limitadas de ejemplos y contraejemplos, sino gracias al manejo de datos cuyo volumen sea en sí mismo ya un criterio de pre-validación de enunciados.

Dejando a un lado perspectivas realizables de futuro y centrando la discusión en los desarrollos actuales, bien puede afirmarse sin grandes problemas que la aplicación de la potencia de cálculo de los ordenadores al tratamiento de la información textual ya produce resultados apreciables. Sin pasar por alto los programas de elaboración automática de índices, léxicos y concordancias, justo es señalar que bases de datos métricos como las que mencionan Iso¹⁷ y Castrillo¹⁸ permiten (o permitirán) completar o refutar las teorías asentadas y las innovadoras, al igual que se ha podido ver cómo un tratamiento estadístico simple de la mayor parte de la poesía latina permite, por ejemplo, proponer que los conectores de causa y consecuencia tienen una distribución que correlaciona bastante directamente con los géneros poéticos¹⁹. Esta zona de confluencia de las técnicas informáticas, los análisis estadístico-lingüísticos y la literatura puede resultar ciertamente prometedora como forma de encarar el clásico problema de las atribuciones de autoría: no parece previsible que el *output* de un algoritmo lleve a la rotunda afirmación de si la *Octavia* es o no obra de Séneca, pero entra en el horizonte de expectativas la generación de un volumen suficiente de datos que ayuden a ejercer la reflexión filológica con algo más que la mera intuición del buen filólogo.

Más discutible es, en principio, la relación posible entre Filología Latina y Lingüística computacional en el campo de la inteligencia artificial, aun cuando ya se ha intentado algo²⁰. De un lado, puede rendir beneficios en el corto plazo el diseño de sistemas de reconocimiento óptico de caracteres (OCR) que nos ayuden, no ya a digitalizar como simple imagen ediciones humanísticas (es lo que se obtiene, y no es poco, cuando se solicita algún documento disponible en la web de la Biblioteca Nacional de Francia, <http://gallica.bnf.fr>) o manuscritos medievales, por ejemplo, sino a realmente leerlos y ser capaces de almacenar el texto.

En nuestros días, la técnica permite la reproducción de un texto facsímil, no ya mediante la reimpresión de una imagen fotográfica, sino previa una tarea de

¹⁷ *Op. cit.*

¹⁸ N. CASTRILLO BENITO, «Automatismo y cómputo de las características métricas del hexámetro latino: aportaciones del Prof. Wilhelm Ott de la Universidad de Tubinga», J. LUQUE MORENO - P. R. DÍAZ Y DÍAZ (eds.), *op. cit.*, pp. 137-145.

¹⁹ Cf. M. LÓPEZ MUÑOZ, «Causa y consecuencia en la poesía romana: estudio frecuencial de conectores», *Flor.Ilib.* 10 (1999), pp. 187-199.

²⁰ Vid. A. VIÚDAS CAMARASA, «Inteligencia artificial en Filología», *Anuario de Estudios Filológicos* 13 (1990), pp. 403-409, para un tratamiento necesitado de una revisión crítica.

limpieza de la imagen que hace mucho más legible el texto mismo. En este sentido, los *scanners* son de buena utilidad para el editor, siempre que éste sea consciente de sus limitaciones. De entre ellas, la principal atañe al tipo de letra mismo: tanto el de los manuscritos como el de las imprentas hasta principios de nuestro siglo son difícilmente identificables por los programas actuales de reconocimiento óptico de caracteres (OCR), de donde que el *scanner* sirva, básicamente, para interpretar cada página a modo de *imagen*, no de *texto*. Hoy en día, este aparato permite editar facsímiles, pero no hacer ediciones, aunque ya van apareciendo tímidos intentos de adaptar las tecnologías de reconocimiento óptico a las peculiaridades de nuestros textos. Téngase en cuenta que los primeros reconocedores de caracteres operaban por comparación de las formas observadas con las almacenadas en una base de datos propia, mientras que los desarrollados con las nuevas tecnologías prefieren utilizar una tecnología más compleja, que intenta abstraer una grafía general a partir de los puntos de reconocimiento del original. Es un sistema en desarrollo, pero que permite mejores resultados y, sobre todo, más flexibilidad en el reconocimiento, ya que le deja más terreno a las capacidades heurísticas del OCR. Lo que falta es adaptar esos algoritmos de reconocimiento a las peculiaridades gráficas de las ediciones propias de la disciplina.

De otro lado, y en paralelo con la afirmación anterior de la utilidad de la Filología Latina para la evaluación y ajuste de los mecanismos de estudio de *corpora* como los entiende la Lingüística Computacional, puede también afirmarse que la amplia tradición de estudios semánticos de la lengua latina es una buena vía para el desarrollo de sistemas expertos que, de acuerdo con un procedimiento de gestión de bases de datos de conocimiento del mundo²¹, ayuden al diseño de *software* específico para la traducción. No es la necesidad de traducciones automatizadas lo que aquí se defiende sino, más bien, la capacidad de la Filología Latina de proporcionar datos suficientemente filtrados y verificados como para que las tareas de investigación y desarrollo, al menos en este aspecto, se vean facilitadas y aceleradas.

2. INFORMÁTICA Y CRÍTICA TEXTUAL

Parece bastante claro que los modernos avances de la informática le permiten al filólogo acometer una serie de tareas hasta hace poco más bien ideadas que

²¹ T. MOURE - J. LLISTERRI, *op. cit.*, p. 193: «...el mundo de la lingüística computacional y el de la inteligencia artificial permanecieron distantes durante bastante tiempo pues sus marcos teóricos eran bien diferentes, y sus objetivos no llegaban a coincidir. No obstante, en los últimos años han ido experimentando un tímido acercamiento. La idea de fondo es integrar en los programas de comprensión automática sistemas expertos que permitan gestionar grandes bases de datos de conocimiento del mundo y efectuar las inferencias necesarias para establecer el significado de un texto. Este conocimiento del mundo no se utiliza tanto para simular una comprensión total del texto, cuanto para lograr determinados efectos prácticos.»

realizadas. Para la utilidad de los *scanners* y sus perspectivas futuras, es mejor remitirse a lo afirmado en el apartado anterior; en cuanto al planteamiento mismo de la edición crítica, parece lógico afirmar que, al menos por ahora, sigue siendo lo más práctico editar uno mismo el texto, aun cuando lo ideal sería poder, como dice Iso, dejar de ser *operarii* y utilizar nuestros recursos mentales en tareas más productivas o, si lo preferimos, más propiamente filológicas.

En cuanto a la introducción del texto, es muy interesante intentar ceñirse a sistemas que permitan su posterior tratamiento informatizado superando la actual costumbre de utilización del procesador de textos como programa de autoedición, que cuenta además con el inconveniente de la difícil traducción de los formatos de archivo propios de cada *suite* ofimática: OpenOffice.org no es reconocido por Microsoft Word, los antiguos formatos de WordPerfect no encajan bien en éste... La falta de una estandarización en el mecanismo de presentación de las ediciones críticas es un problema serio, máxime si se tiene en cuenta que, en el panorama previsible de las ediciones electrónicas, el texto como tal es sólo una de las posibilidades que ofrece la edición. Así, la disponibilidad de un formato editorial fácilmente convertible al lenguaje de bases de datos supondrá una innegable facilitación de la capacidad de abstraer datos concretos del conjunto de las distintas ediciones. La cuestión central es la de encontrar el sistema que permita trasladar el trabajo filológico de la crítica textual a su interconexión con la elaboración prácticamente automatizada de *corpora*.

Un texto almacenado con una serie de marcas específicas (no necesariamente visibles en su versión impresa, pero sí utilizables por un *software* específico de gestión de datos, contiene una serie de informaciones que luego le van a permitir al procesador interpretar la información para presentarnos una página (o para presentársela a la impresora). Tal es el motivo por el que se está recurriendo cada vez con mayor amplitud a aceptar el lenguaje de marcas textuales de la *Text Encoding Initiative*²², que busca contar con un formato que muestra el texto, pero que lo prepara para su tratamiento como fuente de datos. Para esto, se sirve de lenguajes de representación y marcas (el SGML, Standard General Markup Language, y el XML, Extended Markup Language) genéticamente emparentado con el código HTML (HyperText Markup Language), el estándar de las páginas web²³.

El proceso de informatización de la crítica textual no es excesivamente nuevo, y arranca de los años sesenta, cuando una serie de programas desarrollados

²² Un ejemplo de los resultados de TEI puede verse en el *Emblem Project* de la Universidad de Utrecht (www2.let.uu.nl/emblems/html/index.html), donde se nos manifiesta: «At this moment, five books have been digitised. For these we have full transcription, page facsimiles, indexes, and links to sources and parallels, translations and annotation.» Un ejemplo del código fuente en torno al que se estructuran los *Emblemata amatoria* de Heinsius, parte del proyecto mismo, puede consultarse en www.tei-c.org/Applications/Samples/ep02-he1608v1.xml.

²³ www.tei-c.org/: «Initially launched in 1987, the TEI is an international and interdisciplinary standard that helps libraries, museums, publishers, and individual scholars represent all kinds of literary and linguistic texts for online research and teaching, using an encoding scheme that is maximally expressive and minimally obsolescent.»

iban comparando dos textos línea a línea²⁴. En 1970, aparecieron programas como OCCULT (*Ordered Computer Collation of Unprepared Literary Text*), que algo más hacían, pero que seguían requiriendo una colación manual previa y arrojaban un porcentaje de error excesivo, en torno al diez por ciento. En 1972, comienzan a aparecer programas que separan la introducción del texto (procesadores de textos) de la comparación. A principios de los años ochenta, TUSTEP, URICA! II y UNITE van a tomar el relevo de aquellos pioneros, con la secuenciación de edición de textos, procesamiento de palabras y composición tipográfica²⁵.

En la actualidad, existen programas muy claramente enfocados, si no a las tareas propias de la crítica textual, sí al menos a facilitarle al editor la disposición gráfica de la edición misma. El más avanzado es el Classical Text Editor (CTE), desarrollado por Stephan Hagel, que declara haberlo pensado para manejar distintos aparatos que se actualizan automáticamente cuando se cambia el texto, así como para crear y (re)definir siglas, manejar constelaciones de documentos y servir de ayuda a la hora de detectar parentescos entre manuscritos²⁶.

Puede también considerarse en este grupo el Critical Edition Typesetter (CET), que viene a hacer prácticamente lo mismo. Las diferencias más sustanciales entre uno y otro estriban en que CTE está desarrollado de forma propietaria (hay que pagar por su usufructo, y no es un precio especialmente bajo), mientras que CET no es de pago (o no lo advierte en su página web²⁷) y puede, por demás, integrarse con Microsoft Word.

²⁴ Actualmente, dentro del sistema operativo Linux (derivación amistosa y gratuita de Unix) existen utilidades como *diff* y su versión en editor de textos, *vimdiff*, que permiten presentar en pantalla los resultados de la comparación de dos o tres documentos y, más interesante aún, crear uno nuevo mediante la selección intencionada de variantes. No son utilidades desarrolladas para la crítica textual, ni siquiera con intención filológica, sino para facilitar el trabajo de seguimiento de versiones diferentes de un mismo documento; con todo y con eso, tienen una clara aplicación a una parte del trabajo de los editores. Resulta interesante, a modo de ejemplo, ver funcionar Kompere, no otra cosa que la GUI (Interfaz Gráfica de Usuario) de esos programas, desarrollada para distribuciones como Linux Mandrake 9.0.

²⁵ Vid. F. MARCOS MARÍN - M. P. SALAMANCA FERNÁNDEZ, «Programas informáticos para la crítica textual», *Telos* 11 (1987), pp. 105-111. Como en otros trabajos de Marcos Marín, se hace un elogio de la utilidad de UNITE para la crítica textual y se pone un ejemplo de su uso.

²⁶ www.oeaw.ac.at/kvk/cte/main.htm: «The Classical Text Editor was designed to enable the scholar working on a critical edition or on a text with commentary or translation to prepare a camera-ready copy or an electronic publication without bothering much about making up and page-proofs. Its features, formed in continuous discussion with editors actually using the program, meet the practical needs of the scholar concerning text constitution, entries to different apparatus and updating them when the text has been changed, as well as creating and redefining sigla. The possibility to search for manuscript constellations may be of considerable help in detecting affiliations between manuscripts. It is the primary purpose of the Classical Text Editor to do the automatable work which consumes so much time and energy, and let the scholar concentrate on scientific issues.» Vinculado a este programa está una colección de macros desarrollada por el propio Hagel, se supone que para facilitarle al editor que trabaje en Microsoft Word la tarea de maquetación con doble nivel de notas a pie de página; instalada en la versión 2000 de este procesador de textos, la colección de macros no funciona.

²⁷ homepage.ruhr-uni-bochum.de/bernt.karasch/cet/.

En realidad, lo interesante no es tanto la recensión de los programas disponibles cuanto el propio hecho de que se ofrece a los profesionales de la Filología Latina la oportunidad de intervenir en el desarrollo de la tecnología necesaria para facilitar su propio trabajo. Tanto CTE como CET podrán, o no, seguir siendo herramientas útiles, pero su valor es, sobre todo, el de la demostración de que el desconocimiento de nuestras capacidades no nos exime de su desarrollo, si se permite la paráfrasis jurídica.

La *collatio* es el proceso más evidentemente automatizable. Hay una amplia descripción de los procesos que se deben seguir para codificar con éxito el texto en una desproporcionada monografía de F. Marcos Marín²⁸, pero pueden resumirse en: captura del texto, codificación, colación y preparación para impresión. Para estas tareas se puede recurrir a herramientas como TUSTEP²⁹ o URINKA, o a otras aplicaciones más amigables, como *Collate*, desarrollado por un equipo del Center for Technology and the Arts de la Universidad De Monfort, en el Reino Unido, y que da la impresión de ser bastante útil si se ve necesario abordar la colación basándose en el método estadístico-distribucional³⁰.

3. FILOLOGÍA LATINA Y SOFTWARE PARA LA EDUCACIÓN

Casi a nadie se le escapa que la labor filológica tiene un correlato docente. En este aspecto, las posibilidades que ofrece el desarrollo informático también son muy amplias, desde los cada vez más numerosos programas destinados a la práctica de morfología, léxico³¹, sintaxis o métrica³², hasta la utilización de las nuevas tecnologías en función de apoyo a la docencia³³. Muy seguramente, la pizarra seguirá (y deberá seguir) siendo usada, pero esto no impide que, por

²⁸ F. MARCOS MARÍN, *op. cit.*, p. 366.

²⁹ F. MARCOS MARÍN, *op. cit.*, p. 372: «Lo propio de TUSTEP es la edición y el análisis de fuentes históricas, la *collatio* de versiones distintas de un texto, así como la preparación y publicación de índices, concordancias, bibliografías y ediciones críticas.»

³⁰ www.cta.dmu.ac.uk/projects/collate/: «The Project edition of Collate includes a number of advanced tools. These include the ability to explore the relationships between witnesses (using statistical tools)...».

³¹ Vid. P.-E. BARRERA EDO, «Vocabulari llatí: ensenyament per ordinador», *Annals de l'Institut d'Estudis Gironins* 31 (1990-1991), pp. 289-295; A. FERREIRA - D. CAMPOS - E. RUGGERI, «Verbum: una aplicació multimedial para la ensenjanza del latín», *EClas.* 40 (1998), pp. 121-134.

³² Una recopilación de sitios a los que acudir en busca de este software puede verse en www.ual.es/Universidad/Depar/Filesla/latin/software.htm. De ellos, el repertorio más completo, actualizado y asequible es el mantenido por el IES. «Melchor de Macanaz», de Hellín, en alerce.pntic.mec.es/~rmarti41/software.htm. El predominio de los programas anglosajones es abrumador, pero también conviene reseñar el esfuerzo que se está haciendo en Francia, con Yves Ouvrard y François Collard en cabeza, y la tarea de traducción de esos programas franceses que está acometiendo Rogelio Martínez del Oro en ese Instituto. Como tantas veces ocurre, los docentes de Secundaria constituyen la auténtica avanzadilla de la renovación pedagógica.

³³ Vid. E. E. BENEDICTO, «Eines informatiques a l'ensenyament universitari», *Anuari de Filologia. Seccion D. Studia Graeca et Latina* 13 (1990), pp. 33-53.

ejemplo, determinados organizadores avanzados, árboles de descripción, esquemas o incluso lecciones magistrales completas se les puedan proporcionar a los estudiantes bajo la forma de presentaciones, sea en el momento mismo de la clase (un ordenador y un proyector son suficientes), sea listas para ser descargadas en sus propios ordenadores. En este caso, el medio de transmisión de la información puede proporcionar la ventaja de que, al estar más relacionado con la esfera de intereses y aptitudes de nuestros estudiantes, facilita la no siempre sencilla tarea de interesarlos en la materia.

4. DEL TRABAJO INDIVIDUAL A LA INTERCONEXIÓN LABORAL

No es nuevo señalar que los nuevos marcos de trabajo han cambiado sustancialmente en pocos años. Primero fue la posibilidad del tratamiento automatizado de la información el que permitió acceder a la versatilidad de los ordenadores y sus capacidades. Ya hace años que contamos con repertorios de textos latinos como el clásico PHI (Packard Humanities Institute) o la Bibliotheca Teubneriana en CD-ROM, que proporcionan servicios de valor añadido, cuales la posibilidad de búsquedas extendidas, o la generación de concordancias³⁴.

No obstante, pronto se pasó a la convicción de que la resultante del trabajo colaborativo de varios ordenadores conectados entre sí es superior a la mera suma de sus respectivas capacidades de procesamiento de la información. El filólogo, en su despacho, tiene una serie de posibilidades que amplía muy notablemente si se coordina en sus actuaciones con otros colegas.

Hoy en día, no se conciben los ordenadores si no es como un conjunto de puestos de trabajo interconectados a través de redes locales (LAN, *Local Area Networks*) o de redes ampliadas (WAN, *Wide Area Networks*), un conjunto de redes locales conectadas entre sí. El ejemplo extremo de fusión de múltiples redes es lo que llamamos Internet. Puede a primera vista hacerse extraño compartir ideas a través del correo electrónico, pero no es lo importante el medio de transmisión, sino la finalidad perseguida. Esto agiliza los procesos de adquisición del conocimiento y es, sin duda, una gran ventaja.

Un ejemplo que podemos tomar de los métodos de trabajo de los programadores es el llamado *Concurrent Version System* (CVS, Sistema de Versiones Concurrentes), con un funcionamiento difícil de desarrollar y fácil de entender³⁵. Se trata, en esencia, de coordinar los trabajos de distintos programadores

³⁴ Por hacer una prueba, un ordenador bastante modesto para las modas actuales (AMD K6-II a 300 Mhz. con 196 Mb. de memoria RAM y un lector de CD-ROM 4X), tardó escasamente cinco minutos en generar una concordancia de todo el texto de Quintiliano recurriendo a la opción «Make a concordance» del PHI#5.

³⁵ www.hispalinux.es: «CVS es un sistema de mantenimiento de código fuente extraordinariamente útil para grupos de desarrolladores que trabajan cooperativamente usando alguna clase de red. Para ser más concreto, CVS permite a un grupo de desarrolladores trabajar y modificar concurrentemente ficheros organizados en proyectos. Esto significa que dos o más personas pueden modificar un mismo fichero

que van trabajando en partes de un mismo programa; de ellos, unos irán desarrollando una parte, otros irán probándola y otros irán decidiendo qué incorporar a las sucesivas nuevas versiones del programa. Evidentemente, se corre el peligro de que se solapen líneas de código, desaparezcan de una versión a otra, se incorporen elementos no suficientemente comprobados... Lo que se hace es mantener un sistema con varias líneas de desarrollo: habrá una rama «estable» (todo lo que hay, funciona bien), otra «inestable» (no todo lo que hay funciona bien) y otra de «desarrollo», en la que se van almacenando las nuevas piezas de código.

No hace falta un gran esfuerzo de imaginación para ver que esta técnica puede perfectamente aplicarse a un entorno laboral de esfuerzo distribuido en la Filología Latina. Se podría hacer una edición crítica, traducción y comentario de Quintiliano al modo tradicional, verbigracia, encargándose una o pocas personas, cada una de un bloque completo; pero también se podría distribuir la tarea de otro modo, esto es, con mayor número de gente que aporta fragmentos más pequeños de su trabajo coordinados por un grupo menor que toma las decisiones finales. Los frutos del trabajo de cada cual se pueden consultar en cualquier momento; se pueden sugerir enmiendas a esas piezas pequeñas del gran mecano de la edición; se puede observar cómo el trabajo va avanzando gracias a las contribuciones de distintas personas. Y, no menos importante, se puede disponer de la versión no corregida (algo así como los borradores), de las correcciones y de la versión final.

5. USO DEL TRABAJO EN RED: INTERNET Y OTROS

Otro elemento útil y que debe agradecerse a la tecnología informática en general, y a las redes en particular, es la amplia gama de posibilidades que brinda un uso acertado de la Internet. Una somera búsqueda revela que prácticamente cualquier información está accesible sin desplazamientos innecesarios, lo que supone, de nuevo, la posibilidad de eliminar tiempos muertos en el trabajo. Prácticamente todos los textos latinos clásicos están disponibles para consultar. Más interesante aún, también pueden ser almacenados en los discos duros (o en cualquier otro medio semejante, incluyendo la posibilidad de su conversión al formato de los PDA, *Personal Digital Assistants*) y trabajar con ellos. Desde luego, no sustituye esto el discreto encanto de las ediciones impresas, ni el placer del bibliófilo que tan borgeanamente describe D'Ors³⁶, ni siquiera ahorra

sin que se pierdan los trabajos de ninguna. Además, sus funciones más usadas son muy sencillas de usar. Además CVS guarda las versiones antiguas de los ficheros. Esto permite recuperar en cualquier momento versiones anteriores a la actual.»

³⁶ Cf. A. D'ORS, «Del volumen al codex», *EClas* 121 (2002), pp. 107-108. Sólo hay un aspecto en el que disiento de sus opiniones. Afirma en las palabras finales de su escrito que no se imagina «a un joven estudioso del mañana leyendo a Homero en Internet». Personalmente, prefiero que Homero siga siendo leído. No importa tanto el medio como el mensaje.

el principal problema de estas ediciones electrónicas: disponer de los textos, pero no de sus aparatos críticos (al contrario que el texto, están sometidos a las leyes de propiedad intelectual).

El gran inconveniente del modelo actual de trabajo con Internet es, si se quiere, doble. De un lado, la dependencia de los intereses personales de los latinistas de las instituciones del ámbito anglosajón; de otro lado, el fuerte componente de azar con el que hay que enfrentarse cuando se aborda cualquier búsqueda de información. En la Red, la información no es paritaria; lo paritario es la capacidad de acceso. Las Universidades estadounidenses llevan una gran ventaja (si es que ventaja puede llamarse) en esto de poner cada cual sus trabajos y materiales a disposición del resto de la comunidad. Y no es extraño que así sea, toda vez que es el país en el que se inventó y perfeccionó el trabajo en red, y donde el sistema de compartición del conocimiento tiene mayor arraigo. El problema surge cuando resulta que en buena medida dependemos de esa disponibilidad: en Internet se encontrará lo que alguien voluntariamente haya depositado, no lo que se pudiera considerar objetivamente necesario. Su crecimiento, por lo anárquico y no planificado, es imprevisible; su uso, por lo ingente del volumen de información accesible, es cada vez más dificultoso.

En el interesante trabajo de Macías Villalobos aparecido en el número inaugural de esta REVISTA DE ESTUDIOS LATINOS³⁷ se explicaba bien cómo y dónde buscar, y también se citaban y valoraban un buen número de direcciones útiles. Lo que no se desarrollaba (acaso dándolo por sabido) es el problema de la esencial inoperancia de los actuales métodos de búsqueda en Internet.

RESULTADOS DE UNA BÚSQUEDA EN ARGOS

Marcus Tullius Cicero

The Cicero Homepage Cicero Texts Cicero Chronology Cicero Bibliography More Cicero Images Cicero's Biography (according to Plutarch [trans. Dryden]) Cicero Texts Texts provided by Chris...<http://www.utexas.edu/depts/classics/documents/Cic.html> (35.4k) [«CICER*»=213, «OFFICIIS»=1]

Greske og latinske navneformer inntil r 1500

NB! Dette er WWW-versjonen av autoritetslisten. Den er nesten identisk med den trykte utgaven. Bortsett fra at layouten er noe mer «rufsete» er det et...<http://www.ub.uio.no/uhs/sok/fag/teologi/baser/GreLatWWWv.1.html> (221.8k) [«CICER*»=8, «OFFICIIS»=6]

³⁷ MACÍAS VILLALOBOS, CRISTÓBAL, «Internet y la didáctica del latín», *RELAT* 1 (2001), pp. 203-235.

Sem. für Klass. Philologie der Univ. Marbug

Seminar für Klassische Philologie der Philipps-Universität Marburg Wilhelm-Röpke-Straße 6, Block D 35039 Marburg Tel.: +49/(0)6421 / 28-47 53 Fax: +49/(0)6421 / 28-48 50 e-mail: klass.phil@Mailer.Uni-Marburg.DE Stand: ...<http://www.uni-marburg.de/fb07/klassphil/KVV-SS-99.html> (50.4k) [«CICER*»=12, «OFFICIIS»=2]

ULg - Laboratoire d'Analyse Statistique des Langues Anciennes

Le Laboratoire d'Analyse Statistique des Langues Anciennes Interrogation en ligne de textes du LASLA Description Textes disponibles Modalités d'inscription pour obtenir un accès Description Le...<http://www.ulg.ac.be/cipl/bdlatla/> (22.3k) [«CICER*»=13, «OFFICIIS»=1]

Warped Intertextualities

Warped Intertextualities: Naevius and Sallust at Tacitus Histories 2.12.2 Rhianon Ash (St. Hilda's College, Oxford) ABSTRACT This paper argues that Hist. 2.12.2 alludes to Naevius...<http://www.dur.ac.uk/Classics/histos/1997/ash.html> (30.5k) [«CICER*»=12, «OFFICIIS»=1]

Vox Patrum - vol. 15

[1] [2] [3] [4] [5] [6-7] [8-9] [10] [11] [12] [14] [16] [17] [18] [19] [20-23] [24-29] [30-31] [32-33] [34-35] vol. 15 (8), 1988...http://pomoerium.com/contents/vox_patrum/voxp_15.htm (12.8k) [«CICER*»=1, «OFFICIIS»=4]

Tertullian: Sources for this information

Sources for these pages These entries show what I have read (and liked), and some comments on them to help amateurs like myself. The intention...<http://www.tertullian.org/sources.htm> (131.1k) [«CICER*»=3, «OFFICIIS»=1]

De officiis - Index - IntraText CT

Index - Help | Words: Alphabetical - Frequency - Inverse - Length - Statistics | IntraText Library | ulogos Marcus Tullius Cicero De officiis I...<http://www.intra-text.com/X/LAT0102.htm> (11.1k) [«CICER*»=1, «OFFICIIS»=2]

Diotima

Sulpicia's Complaint: On the State of the Nation and the Age of Domitian Translation and notes copyright 2000 James Lawrence Peter Butrica; all rights reserved....<http://www.stoa.org/diotima/anthology/complaint.shtml> (36.0k) [«CICER*»=1, «OFFICIIS»=1]

Eos - vol. 31

[1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25]...http://pomoerium.com/contents/eos/eos_31.htm (15.7k) [«CICER*»=1, «OFFICIIS»=1]

Lektüreliste

[Campus | HHU | ULB | RZ | Verwaltung | AStA | Düsseldorf | Navigator] Seminar für Klassische Philologie [Home] [Mitarbeiter]...<http://www.phil-fak.uni-duesseldorf.de/klassphil/lektuere.htm> (12.4k) [«CICER*»=1, «OFFICIIS»=1]

So - vol. 9

[1] [2] [3] [4] [5] [6] [7] [8] [10] [11] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]...http://pomoerium.com/contents/symbolae_osloenes/so_09.htm (5.7k) [«CICER*»=1, «OFFICIIS»=1]

Wiener Studien

Wiener Studien Our commission is co-editing Wiener Studien, a journal for classical philology, patristics and Latin tradition. The publisher is Verlag der Österreichischen Akademie der...<http://www.oeaw.ac.at/~kvk/kv04.htm> (8.6k) [«CICER*»=1, «OFFICIIS»=1]

Como ejemplo de esa inoperancia, podemos observar brevemente en el cuadro de arriba los resultados de una búsqueda en *Argos*, sin duda el buscador más difundido para estas cuestiones nuestras. Su funcionamiento es el tradicional: una entrada como *Cicerón* (aun formulada como *cicer**) devuelve por igual textos, biografías, traducciones, ensayos, títulos de novelas... Si se quiere, por ejemplo, el texto latino del *De officiis*, habrá que introducir algo así como: *CICER* AND DE OFFICIIS*, y no está garantizada la fiabilidad de los resultados, ya que entre ellos bien podría aparecer una página en la que se haga una bibliografía sobre este tratado ciceroniano sin proporcionar el texto original³⁸.

6. A MODO DE CONCLUSIÓN

A lo largo de estas líneas, se ha intentado trazar un bosquejo apresurado de algunos campos en los que la Filología Latina puede establecer una fructífera re-

³⁸ Un intento de contribuir a la solución de este problema es el que se desarrolla en la Universidad de Almería a través del Proyecto CVMas, mencionado en los agradecimientos preliminares de este artículo, y que busca, en esencia, generar una base de datos jerarquizada y clasificada de materiales accesibles por Internet.

lación con la Informática. Las conexiones con el ámbito de la Lingüística computacional podrán generar, no ya resultados, sino incluso procesos de transferencia de métodos y resultados que, a corto o medio plazo, serán beneficiosos para ambos campos de estudio.

El desarrollo de los métodos de trabajo colaborativo es una de las claves de la propia supervivencia de la Filología Latina, que podrá adaptarse a este cambiante ritmo de nuestros acelerados tiempos proponiendo resultados asequibles y calendarios de realización más ajustados. La correcta y sensata combinación de la tradición didáctica con los procedimientos de las nuevas tecnologías de la información nos permitirá intentar que los estudiantes adopten una actitud al menos considerada hacia la Filología Latina, no ya un reducto de una civilización en la que incluso les cuesta trabajo reconocerse, sino una serie de contenidos, actitudes y procedimientos que ayudarán a insertarlos en el contexto cultural en el que se mueven ya.

Para todo esto, la Filología Latina debe abandonar decididamente el autoimpuesto *limes* de la tecnología y avanzar con resolución por los todavía no hollados territorios de la Informática. No necesariamente pasa esto por una reconversión generalizada de los filólogos en informáticos (eso sería, amén de poco posible, poco práctico), sino por la producción de sinergias entre ambas profesiones.

El filólogo debe indicarle al programador qué necesita y cómo se opera; el programador deberá desarrollar la herramienta específica. El filólogo, para entenderse con el programador, deberá al menos tener una cierta idea de cómo opera la informática; el programador, por el contrario, al entrar en contacto con el filólogo descubrirá un nuevo horizonte de trabajo que le hará comprender que, en el fondo, si él diseña herramientas de tratamiento de la información, el filólogo trabaja con esa información.

La zona común de actuación de unos y otros será la que pueda servir de catalizador de esa siempre interesante reacción químico-académica que se denomina interdisciplinaridad. Pero esa zona común difícilmente puede aparecer en el ámbito de las conversaciones entre compañeros. Necesita de la creación de estructuras institucionales³⁹ (Centros de Computación Humanística, al estilo inglés, o Institutos Universitarios de Computación en Humanidades, al estilo español) que específicamente trabajen en la creación de las herramientas necesarias. Un buen primer paso se puede dar con cierta facilidad por la sencilla vía de los Proyectos de Investigación y Desarrollo; los demás dependerán de nuestras capacidades y voluntades.

manlopez@ual.es

³⁹ A este respecto, hay una interesante reflexión en «Institutional Models for Humanities Computing», www.kcl.ac.uk/humanities/cch/allc/archive/hcim/.