

VARIA

FILOLOGÍA E INFORMÁTICA
EPIGRAFÍA UGARÍTICA

JESÚS-LUIS CUNCHILLOS
JOSÉ MANUEL GALÁN
CSIC. Madrid

I. CIENCIAS HUMANAS Y TECNOLOGÍA INFORMÁTICA

Entre los dos extremos del dicho latino *scripta manent, verba volant irreparabile*, hoy puede decirse que existe un término intermedio y aglutinador: «el soporte informático», que permite que la palabra perdure como en un escrito y, al mismo tiempo, se mantiene dinámico y abierto como el discurso.

Es necesario que el humanista haga un esfuerzo por conocer el ordenador y las posibilidades que éste ofrece a su investigación. Puede parecer, cuando se compara el uso de la informática en las ciencias humanas con su uso en otras ciencias, que la máquina y sus aplicaciones avanzan sólo marginalmente en el sentido deseado por el humanista. Ello se debe, entre otras causas, a que el humanista no está interviniendo en la evolución de este nuevo sistema de trabajo. Hay que cambiar esta actitud: no se puede esperar que investigadores de otras disciplinas solucionen los problemas que son propios de las humanidades.

Las «bases de datos relacionales» ofrecen muchas oportunidades para el desarrollo de las humanidades. En el mercado pueden hallarse a la venta diversidad de ellas: se ha de elegir aquella que se adecue mejor a las necesidades científicas de cada uno. Pero, ¡atención!, no habrá ninguna que le responda completamente si antes alguien no la ha puesto a punto.

Trabajando sobre soporte informático, como sobre cualquier otro, hay que recorrer ineludiblemente numerosas etapas para llegar a ver el producto deseado en su estado final. El especialista debe ser

consciente de los problemas que le plantea la mecanización. La informática, por ejemplo, requiere que los datos sean seleccionados y homogeneizados de forma rigurosa. El humanista, capaz de apreciar rasgos diferenciadores en muchos de los datos con los que trabaja, puede pensar que pierde información a la hora de homogeneizarlos. La solución no está en invalidar el sistema de trabajo, sino en mejorar los criterios de selección, clasificación y definición de los fenómenos estudiados.

Introducir la totalidad de los datos sobre los que se trabaja supone una inversión de tiempo considerable. Pero el tiempo empleado se amortiza cuando más adelante se hace uso del banco de datos. Además, evitará a sus colegas el trabajo de reescribir los datos una y otra vez. Cada paso que se dé en la automatización de una de las fases para el análisis y relación de los datos almacenados es un paso que se evita repetir a generaciones futuras.

Pero no basta con introducir datos: hay que relacionarlos entre sí. La utilización de una base de datos relacional conduce al humanista a diseccionar sus procesos mentales en pequeñas o, para él, minúsculas etapas. En esta tarea de reconocimiento regresivo, el investigador toma conciencia de los diferentes componentes de su saber y de las relaciones que éstos establecen entre sí. Al hacer explícitas operaciones que antes realizaba de forma mecánica y que, por tanto, formaban parte de su subconsciente de especialista, el investigador puede corroborar o desechar premisas e hipótesis, e incluso descubrir nuevas relaciones entre los datos con los que trabaja.

El investigador advierte enseguida que la base de datos no le proporciona todo el juego de relaciones al que él está acostumbrado trabajando en su campo. Es indispensable, para salvar este escollo, crear dentro de la base de datos aplicaciones adaptadas a cada especialidad. Hay que programar la base de datos, es decir, dotarla de las aplicaciones propias a las relaciones que cada disciplina establece con sus datos. Este trabajo de programación no puede llevarse a cabo sólo por un informático, sin la ayuda del futuro usuario. El especialista en una materia es quien mejor conoce los datos que utiliza y la relación que entre ellos hay que establecer para hacer de su saber, de sus conocimientos, ciencia. Así, él es el más indicado para introducir en el ordenador el material sobre el que trabaja, para relacionarlo según las posibilidades que le ofrece

una base de datos relacional y para construir aplicaciones más adaptadas a su propia disciplina.

La base de datos relacional así desarrollada para cada especialidad puede dar nacimiento a un «banco de datos». El banco de datos tiene la ventaja de contener todos los datos y herramientas que interesan a una disciplina o a varias. Puede mejorarse y ampliarse continuamente: es un banco abierto.

En realidad, todo el trabajo que se acaba de describir se reduce a conocer la base de datos relacional que el especialista posee en su mente, que ha ido creando con su esfuerzo y sus conocimientos en el decurso de su vida profesional, y reproducirla externamente, formalizándola lo más minuciosamente posible.

Hoy en día, el humanista puede tener dificultades para homogeneizar los datos con los que trabaja y para formalizar las relaciones que constituyen el entramado de su saber, pero tendrá que seguir trabajando en ello. Lo que no pueda homogeneizarse y formalizarse puede ser «saber», pero no es «ciencia». El humanista tiene que aceptar este reto si quiere hacer ciencia. La tecnología informática, que en principio se percibe como una mera herramienta de trabajo, acaba siendo para las humanidades un vehículo hacia un método de investigación científica ¹.

¹ Diversos grupos de investigación dentro del vasto campo de los estudios del Próximo Oriente llevan ya años trabajando en esta línea. La mayoría de los proyectos tienen como objetivo la elaboración de concordancias, índices y/o vocabularios. El *CATSS*, *Computer Assisted Tools for Septuagint Studies*, comenzó a finales de los años setenta con la finalidad de producir un diccionario de la traducción griega de la Biblia conocida como de los LXX; cf. R. A. KRAFT - E. TOV, «Computer Assisted Tools for Septuagint Studies», *Bulletin of the International Organisation for Septuagint and Cognate Studies* 14 (1981) 22-40; ÍDEM, *LXX. Computer assisted Tools for Septuagint Studies (CATSS) Volume 1: Ruth*, Atlanta 1986. M. CIVIL, en el *Oriental Institute* de la universidad de Chicago, está desarrollando un banco de datos de textos sumerios. G. PETTINATO, *Studi per il vocabolario sumero*, I, Roma 1985, trabajó sobre un grupo de textos de Lagaš de finales del III milenio a. C. Un tratamiento exhaustivo recibieron un conjunto de antropónimos por I. J. GELB, *Computer-Aided Analysis of Amorite*, Chicago 1980. La elaboración de estadísticas de ocurrencias de grafemas ha sido también el objetivo de proyectos de investigación; p.ej., G. BUCCELLATI, «Comparative Graphematic Analysis of Old Babylonian and Western Akkadian», *UF* 11 (1979) 89-100. En egiptología, han sido principalmente los corpus de textos funerarios los que han sido informatizados; cf. los artículos de C. CROZIER-BRELOT, M. HAINSWORTH y de W. SCHENKEL en *L'Égyptologie en 1979*, II, París 1982, 13-17; 19-23; 45-53; y, más recientemente, el resumen de diferentes proyectos coordinados elaborado por D. VAN DER PLAS, «Computer-Aided Research on Funerary Texts and Iconography», en *Informatique & Égyptologie Bordeaux 94*, en prensa, accesible por Internet; W. SCHENKEL, «Das Tübinger Konkordanz-Programm», *ZAS* 121 (1994) 142-154; ÍDEM, «Wörterbuch vs. Textkorpus, oder: Wie und ob man überhaupt ein Wörterbuch

II. FILOLOGÍA Y BASES DE DATOS RELACIONALES

Ninguna base de datos está preparada, en su estadio original, para recibir inmediatamente datos filológicos. Con frecuencia el filólogo, por ejemplo, necesita hacer identificables cada uno de los textos y partes de un texto de donde extrae sus datos, necesita hacer uso de caracteres o tipos especiales para connotar fenómenos lingüísticos, etc.

En el caso de Ugarit, como en el de cualquier otro yacimiento arqueológico donde se haya encontrado años atrás documentación escrita, las *siglas* que identifican cada texto no fueron concebidas pensando en su utilización informática ². Para la introducción de textos ugaríticos en el *Banco de datos filológicos semíticos noroccidentales* ³ se evitó crear una nueva sigla, pues ya coexistían varios sistemas de identificación documental, lo que viene generando cierta confusión en el desarrollo de la investigación ⁴. Así, se adoptó como

machen kann», *ZÄS* 121 (1994) 154-159. La informatización de un texto literario ha sido realizada por J. WINAND, *Le voyage d'Ounamon. Index verborum. Concordance. Relevés grammaticaux*, Lieja 1987. La importancia de los bancos de datos es recalcada por C. CROZIER-BRELOT, «Égyptologie et informatique», *CdE* 56 (1981) 49-54.

² Las siglas que han recibido los textos desde su descubrimiento son varias; cf. J.-L. CUNCHILLOS, *Manual de estudios ugaríticos*, Madrid 1992, págs. 19-22. Tampoco la nueva numeración RS propuesta en *TEO*, I, fue creada pensando en la informática; cf. P. BORDREUIL - D. PARDEE, *La trouvaille épigraphique de l'Ougarit*, París 1989. Teniendo en cuenta, entre otras, las exigencias informáticas, se crearon las siglas epigráficas del Tell de Doña Blanca (Puerto de Santa María, Cádiz), que se componen del año del descubrimiento seguido de tres dígitos, suponiendo lo que es más que razonable, que no se encontrarán más de 999 inscripciones en un solo año. La numeración va precedida por las siglas que identifican el yacimiento, en este caso TDB. Ver J.-L. CUNCHILLOS, «Las inscripciones fenicias del Tell de Doña Blanca (I). Primera aproximación», en A. GONZÁLEZ BLANCO, *et alii* (eds.), *El mundo púnico. Historia, sociedad y cultura*, Murcia 1994, 205-216.

³ J.-L. CUNCHILLOS - J.-P. VITA, *Banco de datos filológicos semíticos noroccidentales, I: Datos ugaríticos, I.- Textos ugaríticos*, Madrid 1993 (= *TU*); versión en soporte papel de los textos que contienen los datos sobre los que opera la aplicación informática del *Banco de datos*, que los va corrigiendo y ampliando regularmente.

⁴ M. DIETRICH - O. LORETZ publicaron una concordancia de las diversas numeraciones de los textos en *Konkordanz der ugaritischen Textzählungen*, Neukirchen-Vluyn 1972. Cuatro años más tarde ellos mismos crearon una nueva numeración al publicar, junto con J. SANMARTÍN, un corpus de inscripciones ugaríticas, *Die keilalphabetischen Texte aus Ugarit*, Neukirchen-Vluyn 1976 (= *KTU*). J.-L. CUNCHILLOS incluyó unas tablas de equivalencia entre los sistemas de numeración de C. H. GORDON, *Ugaritic Textbook*, Roma 1965, y *KTU* y viceversa en, «Ras-Shamra», *Supplément au Dictionnaire de la Bible*, IX, París 1979, cols. 1442-1466; ÍDEM, *Manual de estudios ugaríticos*, págs. 241-301.

base la sigla más aceptada en los trabajos de esta especialidad, la de *KTU*, que comprendía el mayor número de textos.

La sucesivas lecturas o colaciones de un texto, en tanto que presenten novedades filológicas, son recogidas en el *Banco de datos* y cada una de ellas recibe su propia sigla. A la sigla de *KTU* se le añadieron las mínimas modificaciones posibles, de manera que siempre pudieran reconocerse: se le antepusieron dos dígitos seguidos de un guioncito, con el propósito de que registrasen las diversas lecturas publicadas de un mismo texto. A través de la sigla, el ordenador distingue entre cada una de las colaciones de las que ha sido objeto un texto.

El prefijo de la sigla *KTU* es 00-. Para los nuevos textos y colaciones publicados después de 1976 se adoptó el criterio de numerarlos siguiendo el orden cronológico de su edición, comenzando por el prefijo 10-. La numeración entre 00- y 10- se ha reservado por si en algún momento se cree conveniente incorporar al *Banco de datos* las colaciones anteriores a *KTU*. Recientemente se ha elegido el prefijo 50- para registrar la lectura resultante de la corrección filológica (no epigráfica) de la 00-, es decir, de *KTU*. El texto de la sigla que comienza por 50- incorpora las siguientes mejoras: a) corrección de errores tipográficos advertidos hasta la fecha; b) separación de cadenas grafemáticas que son interpretadas como compuestas de dos palabras, a pesar de no llevar signo separador⁵; c) supresión de los signos de interrogación, puesto que la advertencia del epigrafista de que la lectura de un signo puede no ser correcta queda ya convenientemente reflejada mediante la superposición del *circellus*; d) indicación de las diversas lecturas posibles de las palabras que presentan uno o más signos dudosos, separadas por tres barras oblicuas contiguas y agrupadas entre llaves, con el fin de facilitar al usuario la elección de la palabra más plausible de entre las combinaciones apuntadas por el epigrafista.

Además, el *Banco de datos* ha asignado siglas a elementos integrantes del texto, de significación para la ulterior crítica textual, que no habían sido antes numerados; como por ejemplo, las líneas

⁵ La identificación y separación de estas palabras ligadas se fue produciendo, al igual que la corrección de errores tipográficos, en el proceso de elaboración de la concordancia de palabras; cf. J.-L. CUNCHILLOS - J.-P. VITA, *Concordancia de Palabras Ugaríticas en morfología desplegada*, Madrid - Zaragoza 1995.

incisas que seccionan el texto en la tablilla, o las líneas hoy no legibles al comienzo y final del texto ⁶.

Para la informatización de información lingüística y filológica, o de documentación recogida en algunas lenguas, como las pertenecientes a la familia de las lenguas semíticas, se necesita disponer de caracteres o *tipos especiales* capaces de reproducir los fonemas de forma adecuada. Es importante que cada fonema o grafema corresponda a un sólo número ASCII. Así, para la correcta informatización de los textos ugaríticos hubo que diseñar un conjunto de tipos especiales. En un primer momento, se creó un tipo en *mapa de bits* que permitía imprimir en impresoras matriciales, es decir, de agujas (p.ej. *Imagewriter* de *Apple*); después, se creó un tipo *postscript* para una mejor impresión láser; finalmente, se ha mejorado el diseño tipográfico con la creación del tipo *Ilu*. Los signos cuneiformes de la escritura ugarítica se reproducen mediante el tipo *Cuneifugarit*. Todos estos tipos de creación propia conservan los mismos números ASCII para cada grafema, por lo que el usuario puede pasar el texto de un tipo a otro a su conveniencia, sin que su contenido sufra alteración alguna.

El filólogo debe tener en cuenta en su estudio de un texto los resultados de la lectura previa del epigrafista, materializados en la transcripción. El epigrafista utiliza en su transcripción un código convencional de signos para connotar los distintos accidentes que presenta la lectura de un texto. Los epigrafistas no hacen todos la misma lectura de una tablilla ⁷ y, para los textos de Ugarit, ni tan siquiera utilizan unos y otros el mismo código de signos ⁸. Esta falta

⁶ Para más detalles sobre el sistema de siglas y sobre el criterio seguido para la introducción de colaciones adoptado para los textos ugaríticos del *Banco de datos*, ver introducción de J.-L. CUNCHILLOS - J.-P. VITA, *TU*, págs. XV-XVI.

⁷ Ver, por ejemplo, las discrepancias (166 en total) entre la lectura de Virolleaud y la transcripción de Herdner tan sólo en la detección o no del grafema separador de palabras en ocho tablillas del Louvre, W. J. HORWITZ, «Discrepancies in an Important Publication of Ugaritic», *UF* 4 (1972) 47-52.

⁸ Nótese la variación de los códigos para connotar accidentes epigráficos incluso en publicaciones de los mismos autores; e.g., D. PARDEE, *Les textes hippatriques*, París 1985, pág. 11, ÍDEM, *Les textes para-mythologiques*, París 1988, pág. 290; y entre la primera y segunda edición de *KTU*. Para connotar, por ejemplo, que la lectura de un signo no es del todo cierta, por estar éste en parte dañado, se ha utilizado una

de consistencia y de consenso añade incertidumbres al trabajo ulterior del filólogo.

Tanto el método científico como el ordenador hacen indispensable la selección y la definición de los fenómenos observados y la homogeneización de las connotaciones para los mismos fenómenos. En lo concerniente a los textos ugaríticos, el *Banco de datos* ha realizado un esfuerzo considerable y ha tomado una serie de opciones, que pueden leerse en la introducción a la publicación de la primera parte, I, *Textos ugaríticos*⁹.

Los accidentes o *fenómenos epigráficos* detectados en los documentos de Ugarit han sido señalados en el *Banco de datos* de la siguiente forma¹⁰:

— La tablilla, en muchos casos, no se encuentra en buen estado de conservación y, aunque probablemente el texto fue inscrito comenzando desde su borde superior, los grafemas pueden en su estado actual percibirse sólo desde unas líneas más abajo. El signo ... al inicio de la transcripción sirve para indicar esta posible laguna inicial en la lectura del texto. Tal vez deberían indicarse los milímetros del espacio indeterminado, pues ofrece pistas sobre el posible número de líneas perdidas.

Una posible laguna de tamaño indeterminado puede encontrarse en cualquier otro lugar de la tablilla, en cuyo caso se indicará mediante el mismo signo ... en el lugar correspondiente de la transcripción. De nuevo, sería aconsejable señalar los milímetros que ocupa ese espacio hoy anepígrafo.

— El texto legible puede no comenzar en la margen izquierda de la tablilla porque ésta ha sufrido daños. Si el espacio epigráfico hoy perdido es indeterminado, se indica mediante el signo [...] al

interrogación entre paréntesis tras el signo en cuestión, la letra normal (frente a la cursiva), el asterisco *, el *circellus* °, los semi-corchetes " " y el estilo de letra cursiva (frente al estilo normal).

⁹ J.-L. CUNCHILLOS - J.-P. VITA, *TU*. A los signos convencionales de la pág. X, hay que añadir tres más: /// entre palabras, indica las diferentes palabras que pueden formarse a partir de una cadena grafemática con uno o más signos con lecturas alternativas; { } agrupa las diferentes lecturas posibles de una palabra a partir de las alternativas ofrecidas por el epigrafista; [bn] corchetes con signos dentro, significa que, existiendo un espacio epigráfico determinado, los grafemas son restituidos por el epigrafista. A pesar de que no los lee en la tablilla, el tamaño de la laguna y el contexto le conducen a deducir que éstos son los grafemas que fueron inscritos.

¹⁰ Nótense las doce consideraciones sobre epigrafía mencionadas en J.-L. CUNCHILLOS, *Manual de estudios ugaríticos*, págs. 137-143.

comienzo de cada línea. Si el epigrafista cree que puede calcular el número de caracteres desaparecidos de una línea, lo señala con el número de guioncitos entre corchetes equivalente al número de caracteres; p.ej. dos caracteres = [--]. En el caso de que dude sobre el número de caracteres que contenía el espacio, por ejemplo si se trataba de dos o tres caracteres, bastaría que indicase por defecto el número menor posible: dos.

Una línea o conjunto de líneas del texto pueden encontrarse dañadas en la margen derecha o en su parte central, o incluso en varios lugares a la vez. La laguna o lagunas se indican de la forma señalada en el párrafo anterior [...], en el lugar correspondiente de la transcripción del texto.

Cuando no hay indicios para discernir si un espacio en una línea de texto contenía grafemas o fue originariamente dejado anepígrafo, se utiliza el signo [...] para hacer constar ese espacio y la posibilidad de que hubiera caracteres inscritos.

— Un grafema puede no estar conservado en su integridad, siendo parte de él hoy imperceptible. Si el trazo que falta no es significativo para la identificación del signo, la transcripción no tiene por qué reflejar este accidente.

En el caso de que la identificación del grafema, por estar dañado, no sea segura, el epigrafista connotará este accidente superponiendo a la transcripción del signo que parece más plausible por su contexto un *circellus* °. Los semi-corchetes " " utilizados en la transcripción de escritura silábica son, en principio, innecesarios cuando se trata de un alfabeto de treinta signos, puesto que los posibles grafemas restituibles se reducen a cuatro o cinco, que pueden ser convenientemente indicados por el epigrafista ¹¹.

— Si el epigrafista cree que existen varias restituciones posibles para un signo parcialmente dañado, explicitará en su transcripción cuáles son todos los signos susceptibles de ser restituidos, separando cada uno de ellos con una barra oblicua /. Por ejemplo, *k/r/a/n*. Cuando no se sabe si el espacio contenía un grafema significativo o el grafema separador de palabras (transcrito por un punto), se explicitan las dos posibilidades, separadas por una barra oblicua y ambas con un *circellus* superpuesto, ° / °.

¹¹ D. Pardee reemplaza el *circellus*, presente en algunas de sus publicaciones, por los semi-corchetes siguiendo motivos lingüísticos y/o tipográficos; cf. D. PARDEE, «The Ugaritic umma izbu Text», *AJO* 33 (1986) 117-147, 117 n. 5.

— Cuando los signos dudosos son tantos que podrían leerse varias palabras diferentes, se señalan éstas separadas entre sí por tres barras oblicuas contiguas ///. Se ha preferido este signo y no // para evitar así que el signo se interprete como «paralelo de».

— Las diferentes lecturas posibles de una cadena grafemática se agrupan entre llaves {///} con el fin de que las variantes sean fácilmente identificables como tales. Ejemplo: [...] {r̄hd///k̄hd} . bd . yr̄t.

— La lectura de una palabra puede suscitar dificultades semánticas, lo que conduce al epigrafista a conjeturar que el escriba ha omitido, añadido o tergiversado algún signo de la palabra. Entonces, después de la transcripción de su lectura literal, indicará entre paréntesis () la palabra propuesta como correcta.

La informática, que aún hoy es percibida por la mayoría de los humanistas como un mero apoyo de escritorio a sus labores investigadoras, se convierte en un método científico para las humanidades. La selección, homogeneización y formalización de los datos objeto de estudio, otorgan a la argumentación y a las conclusiones alcanzadas por el humanista el mismo nivel de rigurosidad que la máquina y la programación previamente le exigieron a él.

El sistema de siglas, la creación de tipos especiales teniendo en cuenta la numeración ASCII, así como la formalización de los accidentes epigráficos, son ejemplos inmediatos en los que la informática influye en la tarea investigadora del humanista para racionalizarla; desde luego, hay muchos más.

Con estas líneas esperamos abrir un foro de discusión tanto sobre estas tres cuestiones concretas, de relevancia para los estudios sobre lengua ugarítica, como sobre otras formas de interacción entre la informática y los diversos campos de las ciencias humanas ¹².

¹² Correo electrónico de contacto con los autores para la discusión de estos temas: cunchillos@fresno.csic.es y jmgalan@shapshu.filol.csic.es

RESUMEN

Las bases de datos relacionales ofrecen muchas oportunidades para el desarrollo de las ciencias humanas. El investigador ha de buscar criterios de selección, clasificación y definición de los fenómenos estudiados. Una vez introducidos los datos, las relaciones entre ellos se establecen reproduciendo el conocimiento del especialista y las diversas etapas que recorre su razonamiento científico. Es necesario desarrollar aplicaciones adaptadas a las relaciones que cada especialidad establece entre sus datos. Como ejemplo en el campo de la Filología, se presenta la elaboración de un sistema de siglas, la introducción de colaciones y de connotaciones epigráficas de los textos ugaríticos en el *Banco de datos filológicos semíticos noroccidentales*.

SUMMARY

Relational data bases offer opportunities for the development of Human Sciences. The researcher has to look for criteria to select, classify and define the phenomena under study. Once the data have been introduced, their relations reproduce the specialist's knowledge and the various stages which his scientific reasoning runs over. It is necessary to develop computer applications in accordance with the relationships that each field establishes with its data. This article presents, as an example within the field of Philology, the creation of a sigla system and the introduction of collations and epigraphic connotations of Ugaritic texts in the *Banco de datos filológicos semíticos noroccidentales*.