



CONCORDANCIA CALDERONIANA: ESTADO DE LA CUESTIÓN

Guido Mensching Institut für Romanische Philologie Freie Universität Berlin Habelschwerdter Allee 45 D-14195 Berlin mensch@lingrom.fu-berlin.de

Jürgen Rolshoven
Institut für Linguistik / Informationsverarbeitung
Universität zu Köln
Albertus-Magnus-Platz
D-50923 Köln
rols@spinfo.uni-koeln.de

Manfred Tietz Romanisches Seminar Ruhr-Universität Bochum Universitätsstr. 150 D-44801 Bochum Manfred.Tietz@rub.de

1. HISTORIAL Y PRELIMINARES

El artículo que se lee a continuación tiene como motivo la próxima publicación de la tercera parte (*Dramas*) de la *Concordancia calderoniana*. Este proyecto se debe, como bien se sabe, a la iniciativa de Hans Flasche, quien —en colaboración con Gerd Hofmann— publicó hace casi tres decenios la primera concordancia del teatro caldero-

niano¹. Se trata de una concordancia parcial, ya que su objeto fueron tan solo los Autos sacramentales, es decir aquella parte del teatro calderoniano que por aquel entonces más interesaba a Hans Flasche y al círculo de sus discípulos en la Universidad de Hamburgo. Aunque esta concordancia abarca tan solo una tercera parte de la producción teatral completa de Calderón, fue el resultado de largos y muy penosos años de trabajo. La iniciativa técnica de la concordancia se debió a las actividades de la Deutsche Forschungsgemeinschaft (DFG), el homólogo alemán del Consejo Superior de Investigaciones Científicas, que por aquellos años fomentó la naciente tecnología de los ordenadores en la ciencias y quiso que esta tecnología se aplicase también a las humanidades (Geisteswissenschaften). Se creó un Centro Alemán de Procesamiento de Datos (Deutsches Rechenzentrum) en Darmstadt, donde se elaboró también, desde el punto de vista electrónico, la Concordancia calderoniana y donde siguen existiendo los datos en versiones electrónicas actualizadas.

La novedad del tratamiento electrónico de una gran cantidad de textos y su transformación en una concordancia se manifiesta en el subtítulo español, algo pedante y orgulloso a la vez: «Concordancia aplicada a las obras de Calderón con auxilio de una computadora electrónica». En lo que se refiere a la base textual de la concordancia, Hans Flasche se decidió por la edición de los autos sacramentales publicada por Ángel Valbuena Prat², siguiendo en ello las

exigencias que T. H. Howard-Hill³ considera como fundamentales para la base de una concordancias literaria: 1° que se trate de una edición lo más completa posible [y] 2° que sea una edición accesible⁴.

En aquel entonces, por falta de los medios técnicos adecuados, no era todavía posible escanear los textos calderonianos. Tenían que transferirse, palabra por palabra, en unas tarjetas perforadas que formaron la base del procesamiento electrónico posterior. El resultado de este procesamiento no fue una mera lista de las palabras por orden alfabé-

¹ Flasche y Hofmann, 1980-1983.

² Madrid, Aguilar, 1967.

³ Howard-Hill, 1979.

⁴ Mensching, 1998, p. 300.

tico. Se realizó una presentación de las palabras según el sistema de «key word in context» (KWIC), que, en este caso, cita cada palabra en su contexto métrico (verso completo). Hemos aquí un ejemplo de cómo se imprimieron, exclusivamente con mayúsculas y con signos diacríticos auxiliares (E´ en vez de É o N´ en vez de Ñ, los textos de la concordancia con las impresoras y copiadoras todavía deficientes de la época, de modo que en no pocos casos resulta bastante difícil descifrar los signos impresos:

```
## 485 2 17

| STATE |
```

Ilustración 1: Extracto de la concordancia de los Autos sacramentales⁵.

Además, en el lado izquierdo de la página, hay una «parte de referencia», donde se indica la fuente de la palabra con una sigla del título del auto sacramental correspondiente (por ejemplo, para la primera línea: VG = La vacante general), la página de la edición de Ángel Valbuena Prat (485), la columna (2) y el número del verso contado desde el principio de cada una de las dos columnas de cada página (17)⁶. No se tomó en consideración un elemento muy importante y específico de las obra teatrales, la indicación del personaje que habla. A pesar de estas deficiencias técnicas (evidentes pero no insuperables) de la Concordancia de los Autos sacramentales, publicada en cinco volúmenes bastante gruesos que abarcan más de 7000 páginas, esta obra constituye una valiosa herramienta para cualquier estudioso del teatro áureo⁷.

Según resulta del título de la obra, Hans Flasche tuvo la intención de «aplicar» la concordancia al conjunto de las obras de Calderón, lo que no logró hacer. La preparación de la *Concordancia de los Autos sa-cramentales*, que tardó mucho más de lo previsto en realizarse, coincidió con la jubilación de Hans Flasche (1977) en la Universidad de

⁵ Flasche y Hofmann, 1980-1983.

⁶ Los números de los versos faltan en al edición de Valbuena Prat y resultan del cálculo de los editores de la concordancia, suprimiendo las acotaciones.

Aunque, bien es verdad, la obra se encuentra tan solo en un número reducido de bibliotecas y por ello queda fuera del alcance de muchos de sus posibles usuarios.

Hamburgo, lo que significó que a partir de aquel momento ya no disponía de los recursos humanos y materiales para llevar adelante el proyecto de concordancia para los dos otros tomos de las *Obras completas*, los dramas (I) y las comedias (II). Además, otro proyecto le tenía muy ocupado por aquellos años, una historia de la literatura española que se publicó entre 1977 y 1989 en tres tomos impresionantes.

A la hora de la muerte de Hans Flasche en 1994, la Concordancia calderoniana quedó en estado de fragmento durante casi un decenio. Lo que sí siguió funcionando fueron los Coloquios anglogermanos, cuya organización y «dirección» (completamente informal) el fundador de estos encuentros, Hans Flasche (el co-fundador Alexander A. Parker había muerto en 1989), confió a Manfred Tietz, quien también volvió a poner en marcha el proyecto abandonado de la Concordancia calderoniana. Sin embargo, ya que —a diferencia de Hans Flasche— no disponía de unas competencias tanto de lingüista como de literato, pudo hacerlo tan solo en colaboración con lingüistas especializados en el tratamiento electrónico de datos lingüísticos. Juntamente con Jürgen Rolshoven y Guido Mensching se reformuló el proyecto que, después de una fase de trabajos previos, se puso en marcha con el apoyo financiero de la Deutsche Forschungsgemeinschaft (DFG) y se integró, además, en un proyecto más amplio todavía, el de la serie Concordancias de Obras Literarias Españolas (COLE)8. Los resultados más bien teóricos de la fase inicial y los principios básicos de estos trabajos se presentaron en el Coloquio anglogermano sobre Calderón que se celebró en St. Andrews en 19969.

El primer resultado concreto de los trabajos conjuntos fue la concordancia del *Teatro cómico breve* (bailes, entremeses, mojigangas) de Calderón, publicado en la editorial Olms en 2003¹⁰. La concordancia está basada en la edición crítica de María Luisa Lobato¹¹, ya que los textos del «teatro cómico breve» —a primera vista poco «calderonianos»— no se incluyeron en los tres tomos de la edición de las *Obras*

⁸ A parte de la *Concordancia calderoniana*, se publicó, como primer tomo de dicha serie, la «Nueva concordancia completa» del *Poema de Mio Cid* (König, Mensching y Rolshoven, 2003).

⁹ Véanse las respectivas publicaciones en las actas de dicho coloquio: Tietz (1998), Bedoya Touriño (1998), Mensching (1998), Rolshoven (1998).

¹⁰ Mensching, Rolshoven y Tietz, 2003.

¹¹ 1989.

completas de Ángel Valbuena Prat y de Ángel Valbuena Briones, o no se habían todavía identificado como obras suyas. Hemos aquí un ejemplo de esta concordancia que, cómo se ve muy bien, sigue también el sistema KWIC:

ETO 161 CAB 129 BZ 175 MENG 94 EDR 200 VEJ 39 ETO 164 CAB 197 MPV 300 AMA 277 EPP 45 NIC 192 ERG 135 PED 160 MPV 297 DMA 218 ETO 157 CAB 58 ERS 416 CARD 49 EF 342 ESCR 24 ESM 116 BOR 146 EPS 53 SALA 68 MPR 263 ZAP 306

por heredarme aprisa. / CAB Pesado sois, por Dios, sin ser de plomo. / RANA Digo que torearé, mas / Cantando... "fuese Bras de la cabaña. / sabe Dios si volverá." / BRAS No volverá aunque se ¡Gracias a Dios que ya llegó mi tanda! / VEJ A Dios soldado, en buena casa queda. / TER No muy / Sale... / MUJ2 ¡Qué furioso! / CAB Dios te libre el rigor de sus castigos. / RANA Juanico traiga. / AMA Anda, niño, anda, / que Dios te lo manda. / ... / MOR Mama, coco, coco. sin dar dinero. / NIC ¿Dices que no me quieres? / Dios te lo pague, / que el cuidado me quitas / de no pasa? / DMA ¡De espaldas estoy, por Dios! / DLU ¡Vive Dios que estoy de espaldas! / / CAB Demos cuatro paseos, / y decidme, por Dios, vuestros deseos, / que de veros tan flaco me de consolaros en tan grande pena. / Quedad con Dios y ahorcaos norabuena / que harta razón tenéis. que me tenéis molido? / ESCR ¡Ah, síl ¡Válgame Dios! Ya sé que ha sido... / Señor alcalde... / ALC grande insolencia? / Esto es rogar. / SALA En Dios y en mi conciencia, / que me ha picado el verme / mas no de viejo. / ZAP Pues vayan / con **Dios** y, si a esto no vienen. / adelante, que

hay quien lo entienda. / PED No se haga, por Dios, todo contienda, / sino recemos mucho aquesta noche / y plegue a Dios que me harten. / BOR A Dios, y aventura empiezo: / Ojos... / MLO Ojos que

Ilustración 2: Extracto de la concordancia del *Teatro cómico breve*¹².

Los avances en cuanto al formato con respecto a la primera concordancia de Flasche son evidentes: el «keyword» está en negrilla y está ubicado en un contexto mucho más complejo, no limitado a un solo verso, sino se sitúa dentro de su estructura sintáctica y textual compleja, lo que posibilita una «lectura» auténtica del «keyword» en su contorno lingüístico y permite comprender todo el sintagma o la frase, sin recurrir, con mucha pérdida de tiempo, al texto mismo de la edición¹³. La «parte de referencia» a la izquierda de la página contiene (tal y como se hizo en la concordancia de los Autos sacramentales) la sigla de la obra¹⁴, la página de la edición crítica de María Luisa Lobato y, al final, el verso en que se halla la palabra. Entre la indicación de la página y la del verso se encuentra —como novedad de esta

¹² Mensching, Rolshoven y Tietz, 2003.

¹³ Cada concordancia impresa presupone un cierto compromiso: por un lado, el tamaño de la concordancia necesita ser lo más limitado posible, y por otro se quiere garantizar la legibilidad de los keywords y del contexto y, además, asegurar que los keywords se localicen fácilmente. Por ello, la Concordancia Calderoniana usa un sistema de formato homogéneo que garantiza la identificación inmediata del keyword.

¹⁴ Por ejemplo: ETO = «Entremés del Toreador», BZ = «Baile de los Zagales», EDR = «Entremés del Dragoncillo». La lista entera (para las cuarenta obras) está en la página 10 de la concordancia.

concordancia— la sigla del personaje que está hablando¹⁵. Se sabe que es esto es sin duda un elemento indispensable para cualquier análisis literario o lingüístico de un texto teatral, y de esta forma la identificación del personaje que habla (o los personajes si intervienen varios personajes en el texto citado) es parte de la concordancia, así que no hay que recurrir al texto de la edición.

2. Las concordancias de los dramas y de las comedias

2.1. Presentación

Como continuación del proyecto de la *Concordancia calderoniana*, publicaremos próximamente las concordancias de los *Dramas* (ya en 2012) y de las *Comedias*. Estas concordancias se publicarán también en el formato KWIC, siguiendo el modelo de Flasche con los avances ya expuestos arriba en relación a la concordancia del *Teatro cómico breve*. La decisión de seguir este modelo fue tomada conjuntamente con la editorial Olms-Weidemann, y no sólo es debida a razones de copyright y al deseo de querer completar la obra de Flasche en un formato más o menos uniforme, sino también a la tradición y la eficacia probada de este tipo de concordancias¹⁶. La base textual será otra vez la edición de las *Obras completas* de Aguilar, en este caso preparada por Ángel Valbuena Briones, ya que las razones indicadas por Hans Flasche y reiteradas por Mensching¹⁷ son aún válidas hoy en día.

Hemos aquí a modo de ejemplo algunas entradas de la concordancia de los *Dramas*, la cual ya está completa y se está preparando en

¹⁵ Por ejemplo: CAB = caballero, MENG = Menga, VEJ = vejete, NIC = Nicolás. La lista entera se encuentra en las páginas 10-11 de la concordancia.

¹⁶ Esto no implica que los textos que hemos preparado como base de la concordancia (véanse los detalles descritos en el parágrafo 2.2) no puedan ser usados en el futuro para publicar un sistema de búsqueda electrónica como aquel esbozado en Rolshoven (1998, siempre y cuando se puedan solucionar las cuestiones de copyright). De hecho, el etiquetaje básico que hemos usado para preparar los textos y la información que nuestro programa extrae y codifica automáticamente en un primer paso antes de generar la concordancia va a ser usado en el futuro para crear textos en XML, como se explicará más adelante en el 2.2.

¹⁷ 1999, p. 300.

este momento para la publicaión (véanse también las reproducciones de la primera y de la última página en los apéndices I y II):

```
ya rendido al paso / que he sustentado hoy, por Dios, / y quisiera descansar, / si es que algún descanso espera me despedi tiernamente, / puse mi espíritu en Dios, / y repitiendo mil veces / / las misteriosas palabras / de que en Una y mil veces le ofrezco; / que le debo mucho a Dios, / y seré feliz si pierdo / por El la vida. / POL (Dent.) En lo ver estas distancias dos, / pues no nació en cuanto Dios / y sí murió en cuanto hombre. / DARÍ Pues ¿no es más
CV 54 a FED 16
PP 207 b LUD 50
DAC 1105 a DARÍ 27
DAC 1099 b CRIS 34
                                 aleve y crüel, / falso, engañador, fingido, / sin fe, sin Dios y sin ley, / que como inocente pierdo / mi honor, venganza me
MH 327 b DaLE 37
                                   aleve y cruel, 1 halso, enganador, inigido, 1 sin le, sin blos y sin ley, 1 que como incorte piedo in incore de dueño, 1 y, en fin, deseo en un hombre, 1 que, sin blos y sin respeto, 1/o abominable y lo horrible / estago alo por Dios, se salga con ello. 1 DLOP No se saldrá tal, por Dios; 1 y si por ventura vos, 1 si sale o no, queréis vello, 1 decid que un templo labrar intenta / adonde viva su Dios, 1 y su fábrica desea / ilustrar con dones tuyos. 1 Mi embajada
PP 183 a LUD 48
AZ 568 a DLOP 31
SO 1165 b HIR 26
                                        de la rama inmortal círculo breve / al defensor de Dios y su ley pía, / horror de la gentil idolatría. / ABS Himnos la fama
CAB 829 a ADON 7
ECR 1003 a ANAS 20
                                     con pérdida la victoria / de tu rey, de tu nación, / tu Dios y tu religión, / quiero creer que la gloria / de ella te alcance por
CAM 1795 a CLA 31
                                               de pobre, / las vanidades de ilustre. / Quiso Dios y tu ventura / que en este estado te acude / la herencia de un
                                         cielos, / nueva luz a mis sentidos; / que entre un Dios y una hermosura / anda delirando el juicio abrazaos. Toca al arma. / TODO Ea, caballeros, Dios / y vuestra razón os valgan. ... / CO A los brazos han venido
CD 658 b LIC 26
PDE 1309 b TODO 23
```

Ilustración 3: Extracto de la concordancia de los Dramas¹⁸.

La concordancia de los Dramas consistirá en unas 6200 páginas y contiene un total de más de medio millón de entradas. Formará la Parte III de la Concordancia calderoniana, donde ocupará los volúmenes 7-14 (A-B / C-D / F-G / H-L / M-O / P-Q / R-S / T-Z), mientras que la concordancia de las Comedias será publicada como parte IV (que consistirá probablemente de los tomos 15-22). Debido a los avances tecnológicos serán volúmenes muy superiores desde el punto de vista técnico-formal, como lo ilustra la ilustración 3 en comparación con la 1: los contextos izquierdo y derecho incluyen aún más caracteres y el formato también fue optimizado a causa de la utilización del formato PDF, como explicaremos más adelante. A pesar del tamaño considerable (aunque repartido en tomos de menor número de páginas y por lo tanto más manejables que los tomos de Flasche), se trata de una concordancia reducida. La reducción se debe a una decisión de principio que tomaron los editores y la editorial Olms-Weidmann: no se incluirán en los volúmenes impresos de las partes III y IV de la Concordancia calderoniana las palabras de muy alta frecuencia, como son los artículos, algunas preposiciones y conjunciones muy frecuentes (por ej. a, de, o, y, que) y pronombres personales y posesivos, cuyas referencias llenarían volúmenes completos¹⁹. No obstante, los usuarios no ten-

¹⁸ Mensching, Rolshoven y Tietz, en prensa.

¹⁹ Véase el cálculo en Mensching, 1998, pp. 303-304.

drán que renunciar a las informaciones relacionadas con estas palabras, por cierto de gran interés para los especialistas (quizás más bien de lingüística que de literatura). Las entradas de las palabras de alta frecuencia se publicarán como apéndices en CD-ROM y/o en el web, en el mismo formato que el resto y probablemente con opciones avanzadas de búsqueda. La editorial tiene la intención de publicar lo más pronto posible el primer tomo de la concordancia de los dramas calderonianos (principios de 2012)²⁰ y hacerle seguir los demás tomos a un ritmo bi- o trimensual.

2.2 Algunos aspectos técnicos

Los aspectos técnicos implicados en la elaboración electrónica de una concordancia y de la nuestra muy en concreto se han discutido y explicado con muchos detalles en los artículos ya mencionados de las *Actas del coloquio anglogermano sobre Calderón* de St. Andrews²¹. No parece necesario que todo aquello tenga que repetirse punto por punto para unos lectores interesados más bien en el funcionamiento práctico y la utilidad inmediata de la concordancia y no tanto en sus fundamentos computacionales. Sin embargo conviene tener en cuenta que una concordancia es mucho más que una lista alfabética de palabras y que tan solo puede elaborarse a base de una serie de estándares científicos indispensables. Por ello parece necesario llamar la atención, aunque sea muy brevemente sobre algunos des estos aspectos. También nos parece oportuno informar los lectores sobre algunas innovaciones recientes del proyecto.

Los textos de los dos tomos de Aguilar (*Dramas* y *Comedias*), de alrededor de 900.000 palabras cada una, fueron escaneados²² y corregi-

²⁰ Tenemos, pues, que corregir nuestros pronósticos demasiado optimistas formulados en St. Andrews (Tietz, 1996, p. 282), si bien es verdad que la concordancia del *Teatro cómico breve* logró publicarse en 2003. El retraso de la publicación de las partes III y IV fue debido sobre todo a problemas de personal y de financiación, pero también a una serie de complicaciones que se mencionarán hacia el final de este parágrafo.

²¹ En particular en Rolshoven (1998) y Mensching (1998). Véase también König, Mensching y Rolshoven (2003, pp. 595-598).

²² Los dos tomos editados por Ángel Valbuena Briones fueron escaneados ya en los años 90 con una «máquina de lectura» Kurzweil.

dos, en gran parte a mano, como se ha descrito en el primer esbozo del proyecto²³. Un trabajo bastante árduo fue el etiquetaje básico²⁴, que no incluye sólamente marcas para cada página y cada columna, sino también etiquetas para distinguir el texto dramático de las acotaciones escénicas²⁵. Este trabajo se pudo hacer de forma semiautomática, pero la mayor parte se tuvo que hacer completamente a mano, porque las acotaciones no son homogéneas, ni linguística ni tipográficamente. En total, aparte de varios pasos automáticos, los dos tomos completos tuvieron que pasar por varias lecturas «humanas». En paralelo a la preparación de los textos, el programa que genera la concordancia²⁶ y que ya fue usado para la concordancia del Teatro cómico breve (cuya edición tiene un formato muy homogéneo) fue adaptado para el tratamiento de las ediciones de Aguilar, que son bastante mucho más difíciles de procesar. En particular, el programa tuvo que ser modificado para reconocer automáticamente las partes habladas por cada personaje, la extensión de los versos y otros detalles más que sirven para la concordancia o para la construcción de una base de datos en un futuro²⁷. Para la concordancia impresa, en la cual la «parte de referencia» necesita tener una extensión limitada, fue necesario abreviar los nombres de los personajes o combinaciones de tales nombres a un total de unas 900 expresiones alfanuméricas unívocas de un máximo de cuatro letras. Como se puede observar en los apéndices III y IV, hemos decidido usar estas abreviaturas también en los contextos izquierdo y derecho de cada palabra, para darle más espacio al texto dramático.

²³ Mensching, 1998.

²⁴ Fue desarrollado para este etiquetage un código especial, ver la ilustración 1 en el apéndice III. Esta forma de etiquetas era usual en los años 90, pero resultó ser bastante laboriosa, sujeta a errores y difícil de manejar. Ver más detalles al final de este párrafo.

²⁵ Sobre el problema de las acotaciones escénicas, véase Bedoya Touriño (1998). En la concordancia, estas no se toman en cuenta para la determinación de los «keywords», pero sí aparecen en el contexto derecho o izquierdo, normalmente de forma abreviada, con el sistema descrito en Mensching, Tietz y Rolshoven (2003, p. 12).

²⁶ Se trata del programa *Powerkonk* desarrollado por Jürgen Rolshoven. Véase la decripción de una versión anterior de este programa en Rolshoven (1998) y en König, Mensching y Rolshoven (2003, pp. 595–598).

²⁷ Véase la ilustración 2 en el apéndice 3.

Para la concordancia de los *Dramas* hemos cambiado también el método de imprenta. Para preparar la concordancia del *Teatro cómico breve*²⁸ para la imprenta, se usó aún el Rich Text Format (RTF)²⁹, que presentaba algunas desventajas (por ejemplo a la hora de centrar los keywords³⁰) y necesitaba un postprocesamiento complicado y algunas intervenciones manuales. Por eso, nuestro programa fue modificado de forma que pueda producir un output en PDF. Este nuevo sistema está ya funcionando y se usará a partir de la parte III de la *Concordancia calderoniana* (*Dramas*). Los extractos presentados aquí en la ilustración 3 y en los apéndices I y II fueron ya producidos con este sistema en formato PDF³¹.

Este tipo de avances pertenecen a toda una serie de innovaciones que son el resultado de nuestra experiencia con el procesamiento de los textos calderonianos. La correción y el etiquetaje manual de tal cantidad de textos de un formato particular (un texto dramático en una edición no del todo uniforme) resultó ser problemático, entre otras cosas porque las respectivas tareas fueron realizadas por varias personas en diversos lugares. Además, se tenía que trabajar con textos en dos medios diferentes (texto fuente: papel/texto escaneado y corregido: electrónico/pantalla). Así, los correctores y etiquetadores se veían enfrentados con un permanente cambio de medio. En su totalidad, estos problemas conducían también a inconsistencias frecuentes. Estos problemas se consiguieron resolver, pero sólo parcialmente y de forma provisional, con un sistema de WIKI, el cual permitió que varias personas pudiesen trabajar —a través de la red— en un mismo texto electrónico centralizado (pero aún usando los dos medios, papel y pan $talla)^{32}$.

A base de estas experiencias con la *Concordancia calderoniana*, Jürgen Rolshoven y su equipo de lingüística computacional de la Universidad de Colonia desarrollaron, en el marco de otro proyecto³³, un sistema

²⁸ Y también para la Nueva concordancia completa del Cantar de Mio Cid.

²⁹ Ver König, Mensching y Rolshoven (2003, p. 597).

 $^{^{30}}$ De hecho, si se mira con lupa la ilustración 2 en el parágrafo 1, se aprecia que los «keywords» no están centrados perfectamente.

³¹ Este resultado profesional (desde el punto de vista de la imprenta) fue facilitado por el software *iText*[®] (ver Lowagie, 2010).

³² Fue también fundamental la decisión de encargar una gran parte de los trabajos a una sola persona, por lo menos durante varios años. Queremos agradecer a Isabel

colaborativo basado en web, que presenta simultáneamente varias versiones de un texto por digitalizar. Así, el texto escaneado puede aparecer en un lado de la pantalla en su versión de imagen y en otro como texto electrónico tras el reconocimiento automático de caracteres (ROC) (véase el ejemplo en el apéndice IV), lo que evita la heterogéneidad medial. Como el sistema es accesible a través de Internet por medio de los navegadores estándard, los trabajos de corrección y de etiquetaje pueden ser efectuados de forma repartida. La corrección y otras modificaciones del texto son organizados según los principios elaborados y practicados con éxito por Wikipedia³⁴.

Un último cambio que no se tuvo aún en cuenta para los *Dramas* y que queremos mencionar brevemente toca al sistema mismo de etiquetaje. Cómo se muestra en las ilustración 2 del apéndice III, el texto electrónico que forma la entrada inmediata de la parte del programa que genera las concordancias está codificado con un sistema poco transparente y complicado de procesar³⁵. Hemos decidido generar, con el mismo sistema descrito más arriba en el texto, anotaciones basadas en XML, que seguirán, además las propuestas de la *Text Encoding Initiative*³⁶ (véase el ejemplo en la ilustración 3 del apéndice 3). Esto nos va a facilitar el trabajo a la hora de producir la parte IV (*Comedias*), pero también pensamos convertir los textos electrónicos que tenemos del *Teatro Cómico Breve* y de los *Dramas* al mismo formato para que queden listos en un estandard moderno para aplicaciones futuras.

Bedoya Touriño, a cuya competencia y a cuyo empeño la *Concordancia Calderoniana* debe muchísimo. Agradecemos también al programa *ProSpanien* (Programa de Cooperación Cultural entre el Ministerio de Cultura de España y centros de enseñanza superior alemanes), que subvencionó algunos de estos trabajos.

- ³³ Se trata del proyecto de la digitalización avanzada de la *Crestomatía retorroma*na (publicada por Caspar Decurtins entre 1896 y1919), financiado desde el 2009 por la DFG.
- ³⁴ Es la comunidad de los que participan en el proyecto la que decide si una modificación o una correción debe ser aceptada o no. Todas las modificaciones se almacenan, así que las versiones anteriores son recuperables. Véanse www.spinfo.phil-fak.uni-koeln.de/forschung-drc.html y <a href="https://www.spin
- ³⁵ Este sistema fue introducido en nuestro proyecto en una época en la cual XML ya no estaba establecido. El precursor de XML, SGML, no era ni muy práctico ni muy divulgado y dejaba demasiada libertad al usuario, una cosa contraproductiva para el desarrollo de estándardes.
 - ³⁶ Véase http://www.tei-c.org/index.xml.

3. El porqué y el para qué de las concordancias (no tan solo calderonianas)

En los años 60 y 70 del siglo pasado se generó —por lo menos en los círculos de los interesados— un cierto entusiasmo por las concordancias de textos literarios, en parte también debido a la idea de hacer participar las «humanidades» (Geisteswissenschaften) de los avances tecnológicos de las ciencias naturales³⁷, donde, en la práctica diaria de la investigación y ante la necesidad de manejar enormes cantidades de datos, el procesamiento electrónico de datos se imponía cada vez más. Entre los resultados destacados de esta «transferencia tecnológica» se sitúan también, no cabe duda, la concordancia completa de las obras de Shakespeare de Marvin Spevack (1968-1980) y la Concordancia calderoniana de Hans Flasche³⁸. Lo que importa subrayar en este contexto es el hecho siguiente: no eran necesidades irremediables de la investigación lingüístico-literaria las que llevaron a la elaboración de las concordancias, sino, al revés, la existencia de las concordancias reivindicaba (o posibilitaba) nuevos enfoques y nuevos métodos de la investigación lingüístico-literaria. Bien es verdad que la concordancia de los Autos sacramentales fue y sigue siendo un valioso instrumento de trabajo para la realización de las ediciones críticas de los Autos sacramentales completos de Calderón, colección dirigida por Ignacio Arellano (1992-...). También es verdad que permite tomar en consideración, por parte del estudioso individual, no tan solo el «canon de siempre» sino la totalidad de los autos calderonianos para analizar el vocabulario, los rasgos estilísticos o las imágenes y metáforas de dichas obras. No obstante hay que constatar que estas posibilidades todavía no se han podido poner en práctica con la debida insistencia por parte de los calderonistas. Parece que, por un lado, faltan todavía estudios (y coloquios) especializados y que, por otro lado, los estudios se centran todavía mayoritariamente en el análisis de una sola obra o de un número reducido de obras, es decir, se ocupan de datos que son todavía compatibles con la capacidad del lector individual y para los cuales no es indispensable el recurso a concordancias o sistemas electrónicos.

³⁷ La Association for Literary and Linguistic Computing se fundó en 1973.

³⁸ Si Spevack logró terminar su obra y Flasche no lo logró, este hecho no se debe tan solo a las contingencias biográficas arriba aludidas sino también a la gran diferencia que existe entre el número de obras de Shakespeare y el de Calderón.

Este recurso se convertirá en una necesidad imperante cuando los estudiosos intentarán analizar, a base de «enfoques» o «cuestiones» que quedan por formular, grandes complejos de obras sea de un solo autor (en este caso Calderón) sea de un género literario determinado (en este caso las obras teatrales como lo son la comedia o el auto sacramental). El teatro áureo, con su inmensa producción de textos dramáticos (recuérdese que se estima que el número de las obras representadas fue entre diez y treinta mil) equivale a un tamaño de producción que no se dio en ningún otro ámbito cultural europeo, v parece ser un campo muy apto para tales estudios futuros. El cultural turn, tan presente en los estudios filológicos de los dos últimos decenios, favorece precisamente el estudio de grandes cantidades de obras, para lo cual se exige que el análisis sitúe las obras en sus contextos históricos y discursivos, creando y aprovechando la existencia de enormes «archivos» o bancos de datos que evidentemente exceden las capacidades de cualquier investigador individual³⁹. Todo aquello conlleva a un gran auge de las así llamadas «digital humanities» o «e-humanities» y del «humanities computing», corrientes para las cuales un texto determinado no es nada más que la «punta visible de un iceberg textual» que solo puede entenderse y explicarse al tener en cuenta también su enorme «corpus invisible»⁴⁰.

Evidentemente, hay dos posibilidades para crear estos «archivos»: por un lado la digitalización de textos individuales con sistemas de búsqueda, tal y como se está realizando, por ejemplo, en los bancos de datos CREA y CORDE de la Real Academia, en la *Biblioteca virtual* del Instituto Cervantes o, más recientemente, en la base de datos ARTELOPE del equipo de investigación dirigido por Joan Oleza de la

³⁹ Véanse las revistas especializadas como el bastante reciente *Jahrbuch für Computerphilologie* [Anuario de filología computacional] (desde 2000), congresos internacionales importantes como Language Resources and Evaluation (LREC) o el estudio sistemático y rico en cuanto a las perspectivas de investigaciones posibles de Moritz Baßler (2005), especialmente el capítulo VI («Volltextdatenbanken in der kulturwissenschaftlichen Praxis», pp. 321–332). Agradezco a Felix Kurt Ernst Schmelzer el haber llamado mi atención sobre esta obra (Manfred Tietz).

⁴⁰ Esto significaría en concreto que el *auto sacramental* como género no podría estudiarse sin conectarlo estrechamente, por ejemplo, con la numerosa predicación de la fiesta del Corpus, ni la comedia sin relacionarla con la inmensa poesía petrarquista y los numerosos tratados del amor (tanto profano como divino) del Renacimiento y del Barroco.

Universidad de Valencia: «Patrimonio Teatral Clásico Español: Textos e Instrumentos de Investigación»⁴¹). Por otro lado está la elaboración de concordancias de la obra de autores de la categoría de Calderón de la Barca o de Shakespeare que —en forma impresa en papel o en forma electrónica— pueden servir de «puntos fijos» o de áncoras en la mar de datos que inundará al filólogo del futuro. Daniel Eisenberg tenía toda la razón cuando elogiaba en una ponencia presentada en 1997 la existencia de una versión digital de la obra de Cervantes⁴². Aunque evidentemente tenía razón al calificarla como «algo mucho mejor» que una concordancia, se refería naturalmente a las concordancias en uso de aquel entonces, es decir una mera presentación de palabras por orden alfabético. Una concordancia como la Concordancia calderoniana presenta muchas más informaciones para el usuario, y le presenta ya hecha aquella «concordancia» de palabras, modismos y demás fenómenos que el usuario de la versión electrónica de un texto tiene que ir buscando para así constituir al banco de datos específico que necesita para la temática del trabajo que está realizando. Estamos convencidos de que una visión de conjunto de todo el vocabulario de nuestro autor, donde se pueda «navegar entre las palabras» sin dejar dejar de leer a Calderón a través de sus contextos será un primer paso importante y un instrumento útil para abrir la vista hacia nuevas preguntas de investigación sobre el teatro áureo español y su época.

⁴¹ Ver http://www.tc12.es.

⁴² Eisenberg, 1999.

BIBLIOGRAFÍA

- Babler, M., Die kulturpoetische Funktion und das Archiv. Eine literaturwissenschaftliche Text-Kontext-Theorie [La función poético-cultural y el archivo. Una teoría de la ciencia de la literatura sobre textos y contextos], Tübingen, Francke, 2005.
- BEDOYA TOURIÑO, I., «Texto y contexto: problemas para una concordancia de la obra de Calderón», en *Texto e imagen en Calderón. Actas del Undécimo Coloquio Anglogermano sobre Calderón, St. Andrews, Escocia, 17-20 de julio de 1996*, ed. M. Tietz, Stuttgart, Steiner, 1998, pp. 284-298.
- EISENBERG, D., «Los textos digitales de las obras de Cervantes», en *Cervantes* 1547-1997. Jornadas de investigación cervantina, ed. A. González, México, El Colegio de México, 1999, pp. 53-61.
- FLASCHE, H., Geschichte der spanischen Literatur, Bern, Francke, 1977-1989, 3 vols.
- FLASCHE, H. y HOFMANN, G., Konkordanz zu Calderón. Concordancia aplicada a las obras de Calderón con auxilio de an computadora electrónica. Computerized Concordance to Calderón. Parte Primera, tomos 1-4, Concordancia de los Autos sacramentales, Hildesheim / Zürich / New York, Olms, 1980-1983.
- HOWARD-HILL, T. H., Literary Concordances. A Guide to the Preparation of Manual and Computer Concordances, Oxford / New York, Pergamon Press, 1979.
- König, H., Mensching, G. y Rolshoven, J., Poema de mio Cid: Nueva concordancia completa, Hildesheim / Zürich / New York, Olms, 2003.
- LOBATO, M. (ed.), Calderón de la Barca, P., *Teatro cómico breve*, Kassel, Reichenberger, 1989.
- Lowagie, B., *iText in Action*, Stamford (CT), Manning Publications Company, 2010, 2.^a ed.
- Mensching, G., «Planteamiento general de la Concordancia calderoniana», en Texto e imagen en Calderón. Actas del Undécimo Coloquio Anglogermano sobre Calderón, St. Andrews, Escocia, 17-20 de julio de 1996, ed. M. Tietz, Stuttgart, Steiner, 1998, pp. 299-311.
- Mensching, G., Rolshoven, J. y Tietz, M., Concordancia calderoniana: Parte segunda, tomo VI, Teatro cómico breve, Hildesheim / Zürich / New York, Olms, 2003.
- Concordancia calderoniana: Parte tercera, tomos 7-14, Dramas, Hildesheim / Zürich / New York, Olms, en prensa.
- ROLSHOVEN, J., «Realización técnica y perspectivas de la Concordancia calderoniana: hacia un sistema hipermedial», en Texto e imagen en Calderón. Actas del Undécimo Coloquio Anglogermano sobre Calderón, St. Andrews, Escocia, 17-20 de julio de 1996, ed. M. Tietz, Stuttgart, Steiner, 1998, pp. 312-322.
- Spevack, M., A Complete and Systematic Concordance to the Works of Shakespeare, Hildesheim / Zürich / New York, Olms, 1968-1980, vols. 1-9.

Tietz, M., «La concordancia completa de las obras teatrales de Calderón de la Barca», en *Texto e imagen en Calderón. Actas del Undécimo Coloquio Anglogermano sobre Calderón, St. Andrews, Escocia, 17-20 de julio de 1996*, ed. M. Tietz, Stuttgart, Steiner, 1998, pp. 281-283.

Apéndice i

DAC 1073 1/207-5 FAP 75/2 a CRI CLI 25 Micros National Control (1997) and selection (1997) about a control (1997) and selection (1997) per sersa que en medio de eicos / un aguita adatenta / su mais retionitatos vietor / jan antiguio es en en rutino de la funtion inspirent. A ellos. / Per Esto se pero, que Abdalá, / alentado en mi despeño, / creyendo que mantro cago desesperad. A ellos. / Per Esto se pero, que Abdalá, / alentado en mi despeño, / creyendo que mantro cago desesperad / a contrares, a tempo que i de Abdalá, / carleda en ser suya la victoria, al paliga es desmanda dicha mipera? / pues como yo el veriorimento / de Abdalá, contrar la en ser suya la victoria, al paliga es desmanda dicha mipera? / pues como yo el veriorimento / de Abdalá, contra la la, cuesta el el el mio aqui, / Disk. / que blem mas que no com abbará! e y ela mostrado es Abdalá. / Carleda belleto le envia el el Liona, como elesta crafas

Concordancia calderoniana, parte III (*Dramas*), tomo 7 (A-B), p. 1 (tamaño reducido).

APÉNDICE II

6224

MRT Concordancia Calderoniana

TMP 1582 a FLOR 4 PC 252 b MUL 2 GPF 1365 b PRÍ 18 ADM 376 a DÁLV 33 MMM 488 b TET 19 GPF 1368 a BUEN 2 AC 1855 b APOL 16 EJM 918 b DE 34 HAP 730 b MEN 55 FAP 1662 b CELI 16 GS 1735 a ESC 4 FAP 1650 a FINE 35 HF 1254 b PET 28 AZ 548 a ISA 11 TMP 1574 b AR 6 ACO 1322 b IDOL 46 AHE 970 b VET 15 ADM 360 b DÁLV 3 AA 1691 a MAL 3

AC 1850 b SÁT 30 pazguato, un majadero, / que sin dignidad de loco, / zorrero bajel de hueso, / se deja venir a fondo / en busca de aquel las olas, / perdido el piloto suyo, / a todas partes zozobra. / UNO Los verdinegros cristales / teñidos en la espumosa sola / una nave, y que en el mar / mal defendida zozobra: / porque, según después supe, / de una tormenta que todas ALA 1962 b SIQU 19

alevosa mi lengua. / Entre Cariódis y Solia / Ian zozobrada padezcas, / que desees por bonanzas / las circes y las

LAP 1743 a DAFN 49

GS 1724 b ESC 3

vista... / C ¡Qué mal a mi oído suena... / ESC ...el zozobrado huracán... / C ...la desesperada queja... / ESC ...de aquel el sueño / avanzadas centinelas / en zozobrado sosiego / descansan muchos dormidos / en fe de pocos AA 1688 b FLORA 11 parte surca. / FLORA Embates de mar y tierra / le zozobran y le asustan. / AUR Y tanto, que desbocado / choca con las HF 1258 b IDAS 1 con la corriente, / contrastando a los embates, / zozobrando a los vaivenes, / rozándose en una peña, / al tope la quilla con la corriente, / contrastando a los embates, / zozobrando a los vaivenes, / rozándose en una peña, / al tope la quilla a quien no en balde / dieron este nombre, ya / zozobrando sobre mares / de púrpura que la anegan, / de llamas que la y hay fuego / donde hay agravio sumo. / para zozobrar suspiros, / y para hacer llorar humo. / ¿Qué sesto? ¡Ay de la lid, / que arde interior en su pecho) / zozobraré tus aplausos / y turbaré tus trofeos, / sacando de sus azares dichas divulguen... / APOL Por ti venciendo zozobras ... / CLI Por ti gozando quietudes? / TODO Que vivan y que lo sepas, no / has de tener, por lo menos / sin zozobras y pesares, / persecuciones y riesgos, / fatigas, ansias y y bien hecha, / una valla de cristal, / sin que zozobrase en ella / la perfección, siendo así / que la nariz más perfecta las turbadas alas mueve. / haciéndola que zozobre / al espolón de su frente. / al tiempo que amotinado / de ellas estriba, haciendo / que el impulso de caer / le zozobre a los embates / de un vaivén y otro vaivén. / Y a esta villana no hay bajel / de cuantos al puerto llegan / que no zozobre a su vista; / porque su estatura inmensa, / si se mueve, es tome / las vivas aguas del Nilo, / en sus corrientes zozobre, / pues no podrá contrastarlas / fusta de tan poco porte. / A la si cortés / ponéis en obligación / nuestras vidas, no zozobre / tan presto a la intercesión. / Que dejeis este soldado / os sobre / sus espumas bajel sea / que a poco tiempo zozobre, / yéndose a pique contigo; / y desde la quilla al tope / hecho aquel bajel con las nóticias vaya, / le embata, le zozobre y le persiga, / por más que ahora, viento en popa, diga / en ni en las mismas flores hay / que no den, rojas o **zules**, / tósigo a la araña amargo / y miela la abeja dulce. / Y pues caballero / sabré ser cuando conviene; / que soy **Zúñiga** en Castilla / antes que Justicia fuese, / así, arrimando esta Vamos por esta otra mano, / por si es más quieta la zurda. / ¡Ah señor! / DANT ¡Válgame el cielo, / y qué crueldad tan DAL 1466 a MOR 38 debo / a derechos ni a derechas; / que también soy **zurdo**. / SOL1 Vaya / el mandria... / SOL2 La mosca muerta

> Concordancia calderoniana, parte III (Dramas), tomo 14 (T-Z), p. 6224 (tamaño reducido).

APÉNDICE III

..JM

. . 5

..(

JUDAS MACABEO

PERSONAS QUE HABLAN EN ELLA

LISÍAS. JUDAS MACABEO. GORGIAS.

SIMEÓN. TOLOMEO. CHATO, villano. JONATÁS. ZARÉS, dama. JOSEF, soldado.

MATATÍAS, viejo. Un CAPITÁN. CLORIQUEA, dama.

Soldados, etc.

[La escena es en Jerusalén y otras partes]

JORNADA PRIMERA

..) ..a

..([Campo]..)

..(Salen por una parte, habiendo tocado cajas y trompetas, JONATÁS, SIMEÓN, JUDAS, y por otras, MATATÍAS, viejo, ZARÉS y Músicos. ..)

MÚSS.-..((Cantan.)..)
Cuando alegre viene
Judas vencedor,
su frente coronan
los rayos del sol.
MAT.-Valerosos Macabeos,
legítima sucesión
de palestinos hebreos,
cuya gloriosa opinión

Ilustración 1: Inicio de la obra *Judás Macabeo* con etiquetaje básico.

```
$[Campo]...&/
$...&/
JM 5 a MÚSS 1#$MÚSS%MÚSS.-$(Cantan.) &/
JM 5 a MÚSS 2#
                 Cuando alegre viene/
JM 5 a MÚSS 3#
                  Judas vencedor,/
JM 5 a MÚSS 4#
                 su frente coronan/
JM 5 a MÚSS 5#
                  los rayos del sol./
JM 5 a MAT 6#$MAT%MAT.-Valerosos Macabeos,/
JM 5 a MAT 7#
                legítima sucesión/
JM 5 a MAT 8#
                de palestinos hebreos,/
              cuya gloriosa opinión/
JM 5 a MAT 9#
```

Ilustración 2: Inicio de la obra *Judás Macabeo* con etiquetaje avanzado (automático).

```
<body >
 <div type =" act ">
 <head >JORNADA PRIMERA </ head >
 <stage type =" setting ">[ Campo ]
 <stage type =" entrance ">
  ( Salen por una parte , habiendo tocado cajas
y trompetas , JONATÁS , SIMEÓN , JUDAS , y por
otras , MATATÍAS , viejo , ZARÉS y Músicos .)
 </ stage >
 <stage type =" sound ">MÚSS cantan .
 <sp >
 <speaker >MÚSS .
   <1>Cuando alegre viene </1>
   <1>Judas vencedor ,</1>
   <1>su frente coronan </1>
   <1>los rayos del sol .</1>
   </sp>
 </ div >
</ body >
```

Ilustración 3: Versión en XML.

APÉNDICE IV

Sistema colaborativo (basado en web) para la digitalización de textos escaneados: cuando el usuario clica en una palabra del texto ya interpretado por el software ROC (izquierda), aparece un marco en la parte correspondiente en la imagen (derecha). Aquí se ha clicado en la ventana de edición sobre el nombre BONIFACI, que luego aparece en caracteres verdes. La ventana de la derecha muestra la imagen original del texto. Los datos del ROC permiten que se calcule la posición de la palabra en cuestión en la imagen. Esta aparece luego en un margen con fondo verde en la parte derecha. Este procedimiento permite al usuario (sin que este recurra al libro «físico» original) juzgar si el texto reconocido por el software ROC corresponde al original. Si no es el caso, el texto puede ser corregido en seguida.

