

PROBLEMAS TEÓRICOS Y PRÁCTICOS EN LA REDACCIÓN DEL *NTLE**

Manuel Alvar Ezquerra

Universidad Complutense de Madrid

Cuando los organizadores de este volumen me propusieron que escribiera sobre los problemas en la redacción de diccionarios generales de la lengua, me alegró porque me hacía recordar pasados retos e imaginaba los que pueden presentarse a la actividad diccionarística en los próximos años. Sin embargo, no me atreví a sostener el envite porque en estos momentos mis preocupaciones se hallan orientadas en otras direcciones. Mirando hacia el quehacer actual que absorbe la mayor parte de mi tiempo fue cuando les propuse tratar de las dificultades que surgen, que nos han surgido, en la elaboración de un tesoro lexicográfico, el *Nuevo tesoro lexicográfico del español (s. XIV-1726) (NTLE)*¹, que, por otra parte, no deja de ser un diccionario de carácter general, si bien muy diferente de lo que concebimos habitualmente como un diccionario general de la lengua, de carácter descriptivo y sincrónico. Por otro lado, el *NTLE* se encuentra más cerca, porque lo es, de los repertorios de carácter acumulativo e histórico, y muchos de los problemas que se nos presentaron son comunes a los que aparecen en cualquier otro proyecto de lexicografía histórica. No son, pues, problemas nuevos, sino viejos conocidos –como siempre, unos más que otros– que reaparecen al emprender una nueva obra.

Una de las principales dificultades con las que se encuentra cualquiera que desee redactar un diccionario, del tipo que sea, es la informatización, pues, como resulta conocido de todos, hoy ya no se acomete una empresa lexicográfica sin acudir a los ordenadores, no puede ser de otro modo. No creo necesario recordar aquí que fueron los estudios lexicológicos y lexicográficos de los primeros en beneficiarse de los avances

* Este trabajo se encuadra dentro de los llevados a cabo para el proyecto *Nuevo Tesoro Lexicográfico del Español (s. XIV-1726)*, que goza de una ayuda del Ministerio de Ciencia y Tecnología (HUM2004-05344).

1. A propósito de la obra, véanse Alvar Ezquerra (2003a, 2003b, 2004 y 2006a). Además, Nieto Jiménez / Alvar Ezquerra (2004) y Alvar Ezquerra / Nieto Jiménez (2006).

tecnológicos, como ha ocurrido siempre, por lo menos desde que la imprenta irrumpió en nuestro mundo. El uso de los ordenadores en nuestros ámbitos ya no es una novedad: han transcurrido más de treinta años desde que comenzaron a hacerse los primeros trabajos de informatización aplicada al léxico y a los diccionarios del español. Por supuesto, los medios, técnicas, procedimientos, rapidez, volumen de datos tratados, complejidad de las consultas, etc., etc., han cambiado de forma sustancial, y lo que hoy podemos hacer en nuestra casa en unos minutos con un ordenador personal y un programa comercial, entonces requería de un gran esfuerzo personal y material, y de unas cuantas semanas, si no meses, de trabajo.

Por lo que respecta al *NTLE*, los inconvenientes en su informatización han surgido por cualquier parte cuando menos los esperábamos, y hemos tenido que acudir a múltiples frentes para ir solucionando los problemas que se planteaban cada vez que deseábamos avanzar un paso en el trabajo.

En primer lugar, para tratar informáticamente los datos contenidos en los diccionarios que deseábamos recopilar resultaba necesario darles un formato electrónico, como es obvio. Para ello había que disponer de ordenadores, pero ¿de qué tipo? En el inicio de las tareas, una parte del equipo se encontraba en Madrid, y otra en Málaga, así que debíamos buscar la máxima compatibilidad posible, para evitar problemas en el envío de datos de un lado a otro. Nos pareció lo más apropiado emplear ordenadores personales en lugar de grandes equipos, decisión que venía avalada por mi experiencia anterior durante la redacción del *Diccionario general ilustrado de la lengua española (DGILE 1987)*, en la que la utilización de grandes equipos informáticos nos proporcionó más de un sobresalto, y siempre debido a motivos ajenos a la propia redacción lexicográfica. Es más, cuando la editorial Biblograf s. a. decidió crear en Málaga el CELEX (Centro de Lexicografía Vox), se optó por los ordenadores personales, pues eran la mejor solución para facilitar el trabajo en equipo con sedes diferentes, y para el envío de materiales entre Málaga y Barcelona.

La segunda parte del problema informático inicial consistía en la recuperación de la información contenida en las fuentes que iban a servir para la elaboración del *NTLE*. El trabajo que pretendíamos llevar a cabo se nos presentaba inmenso: agrupar decenas y más decenas de diccionarios y otras obras con contenido léxico. La primera idea que acude a cualquiera que se vea en la misma tesitura es la de escanear los repertorios. Sin embargo, apenas la consideramos, entreviendo la multitud de inconvenientes que nos iban a surgir, derivados de los tipos de materiales tan heterogéneos que íbamos a manejar.

De entrada, quedaba descartado el escaneo de los manuscritos, pues los medios técnicos a nuestro alcance no pueden recuperar su información, máxime habida cuenta del pésimo estado en que se encuentra la mayor parte de los que hemos utilizado, y que no son pocos, pues del centenar y medio de fuentes manejadas 23 son manuscritos que hemos tenido que leer nosotros mismos, y dos más de los que se dispone de transcripciones fiables, junto con los tres glosarios medievales editados por Américo Castro. Por otra parte, tampoco era posible el escaneo de los repertorios impresos, pues resulta evidente que, de ningún modo, se nos facilitarían los originales para tal tarea, por lo que hubiéramos debido partir de las fotocopias de que hemos dispuesto, en las que los problemas se acentúan, debido a la falta de limpieza que tienen para poder realizar un escaneo con unas mínimas garantías. Es cierto que disponemos de ediciones facsimilares de algunas obras, de los dos diccionarios grandes de Nebrija (1492, 1495?) y de sus *Introductiones latinæ* (Nebrija 1481), del de Cristóbal de las Casas (1570), del de John Minsheu (1599), del de Hadrianus Junius (1567), de los dos anónimos ingleses de 1554 (Anónimo 1554a y Anónimo 1554b), del *Vocabulario del humanista* de Lorenzo Palmireno (1569), del *Vitruvio* de Miguel de Urrea (1582), de la *Instrucción náutica* de Diego García de Palacio (1587), del tratado de Bernardo de Aldrete (1606), así como de otras obras de gran envergadura de las que queríamos extraer datos, como el *Dioscórides* de Andrés Laguna (1555) o el tratado de las aves de Francisco Marcuello (1617). Poca cosa ante el panorama que se nos presentaba, máxime cuando las ediciones facsimilares, en algunos casos, tampoco presentan la limpieza que se necesita en estas ocasiones, y parten de reproducciones de baja calidad. En cualquier caso, fuesen originales, facsímiles o fotocopias, el escaneado no parecía la mejor solución debido a múltiples causas que se comprenden con facilidad:

- la diversidad tipográfica, que hubiera obligado a instruir al programa en cada caso particular, si es que llegaba a reconocer el tipo de letra. Piénsese en las diferencias entre los impresos de finales del siglo XV y los del siglo XVIII, o en el empleo de tipografía gótica, mantenida durante varios siglos en los textos en alemán y flamenco junto a otros tipos para las otras lenguas de los mismos repertorios.
- las manchas de tinta debidas a una impresión no excesivamente cuidada (acrecentadas en las fotocopias), que hubieran dificultado la interpretación de las letras.
- la escasa calidad de los papeles empleados para la impresión, que aportan otras manchas que dificultan la lectura, en el caso en que

el contraste hubiese sido suficiente entre el papel y la tinta, debido al deterioro de aquél.

- la variedad de lenguas utilizadas en los repertorios, y alfabetos distintos: latino, griego, hebreo, árabe.
- el empleo de frecuentes abreviaturas, no sistemáticas, y, esporádicamente, de algunos signos no alfabéticos.
- las numerosas erratas y errores debidos a que una buena cantidad de los repertorios fueron impresos fuera de nuestras fronteras, por impresores que no conocían bien nuestra lengua.

Téngase en cuenta, además, que los diccionarios eran obras de carácter didáctico, de consulta, por lo que los papeles empleados en ellos no eran los mejores, y su impresión no resultaba todo lo cuidada que hubiéramos deseado, y, por supuesto, de una calidad inferior a la de otros libros, por no pensar en el desgaste por el uso en manos de estudiantes, comerciantes y otras personas, lo que no necesariamente afecta a todos los ejemplares conservados.

Si nos hubiésemos decidido por el escaneo, las tareas de lectura y corrección hubiesen sido muy superiores a lo tolerable para que la solución fuese razonable en términos de tiempo, de trabajo y de gasto.

Aunque poco deseable, y menos apetecible, la única solución que nos quedaba era teclear, uno a uno, todos los diccionarios que debían constituir el cuerpo del *NTLE*, labor a la que nos hemos dedicado varias personas durante unos cuantos años. Pero no se trataba de comenzar el tecleado a ciegas, sin haber imaginado el resultado final. ¿Para qué iban a servir los datos codificados? Sin duda, como en tantos otros diccionarios, tendríamos que montar una base de datos para los tratamientos y consultas que fuese necesario realizar antes de alcanzar la meta propuesta. Mi decisión estaba tomada de antemano: trabajaríamos con *Access* de Microsoft, pese a que se barajó la posibilidad de emplear otra base de datos para ordenadores personales. Ello era así porque conocía el funcionamiento y prestaciones de esa base de datos, gracias a la experiencia con ella en otro tesoro lexicográfico, bien es cierto que de un carácter diferente del que estábamos acometiendo, terminado ya cuando habíamos comenzado el *NTLE*, el *Tesoro léxico de las hablas andaluzas* (Alvar Ezquerra 2000), en cuya realización me surgieron no pocos problemas, de los que he dado cuenta en otros lugares², y que alumbraron nuestros primeros pasos. Sin embargo, debido a la heterogeneidad del equipo, de la diversidad de per-

2. Véase el propio prólogo de la obra, así como Alvar Ezquerra (1996).

sonas que iban a trabajar y de la disponibilidad de equipos informáticos de todas ellas, así como por la falta de conocimientos en herramientas informáticas, y otros inconvenientes menores, no resultaba aconsejable pasar directamente los datos a la base de datos. Fue éste el motivo por el que se optó por teclear los diccionarios en un programa de tratamiento de textos, disponible en todos los equipos, en un formato RTF, con el fin de que no surgieran los problemas de compatibilidad que se nos presentaron en varias ocasiones con el *Tesoro léxico de las hablas andaluzas*.

Esa resolución conllevaba, como es lógico, que todos los colaboradores trabajasen del mismo modo, por lo que fue necesario redactar unas normas de teclado en las que se especificaban las instrucciones que debían seguirse, los signos que se debían emplear, formatos, tamaños de archivos, etc., etc. Como es fácil de imaginar, estas normas fueron modificándose durante los primeros tiempos a medida que avanzaba el trabajo, pues no fue posible prever todas las contingencias desde los inicios. Algunas de esas modificaciones nos obligaron a volver atrás y rehacer parte de lo que ya teníamos hecho, lo cual no siempre fue bien entendido por los colaboradores, aunque lo perseguido era dar consistencia al conjunto de los materiales. Bien es cierto que ahora vemos lo acertado de las decisiones, pues la base de datos nos permite avanzar con una rapidez que, de otro modo, no hubiese sido posible. El ulterior paso desde las informaciones codificadas en el tratamiento de textos a la base de datos, entre otras cosas, obligaba a marcar la separación entre algunos campos para que la carga en ella no plantease otras dificultades sobrevenidas. Pero también ello obligaba a tener en cuenta otros aspectos. Por ejemplo, la base de datos no permite los cambios de tipo de letra, por lo que no era posible el empleo de cursivas y negritas, para lo cual seguimos el mismo procedimiento que en el *Tesoro léxico de las hablas andaluzas*: poner unas marcas de apertura y de cierre, de modo que los cambios fuesen fácilmente detectables en cualquier momento y para cualquier necesidad que pudiese surgir en el futuro, como así sucedió. Tampoco eran posibles los cambios de tamaños de letra, o el empleo del alfabeto griego o el hebreo, pues todos esos signos desaparecían al cargar la base; igualmente, no era posible el empleo de algunos signos especiales, como cruces y otros símbolos que encontramos en algunos de los originales, bien es cierto que de forma muy esporádica. Para todo se fueron adoptando soluciones, que obligarían a cambios más adelante en el trabajo.

Al tiempo que se realizaba el teclado, se marcaban las supresiones de texto innecesario, los blancos, las ausencias, los rotos en el papel, los cortes que impedían la lectura, etc., etc. Pero esos no eran los únicos as-

pectos formales que se iban solucionando. En los textos manejados han aparecido numerosísimos errores y erratas debidos al poco escrúpulo con que se componían los diccionarios, a la escasa formación de algunos tipógrafos, o sus pocos conocimientos de español cuando las impresiones se realizaban fuera de nuestras fronteras, o a que no interpretaban adecuadamente los manuscritos que les llegaban. No resulta extraño encontrar confusiones de *u* por *n* o a la inversa, de *c* por *ç*, numerosos casos de seseo y algunos de ceceo, etc. Se efectuaron las correcciones de errores y erratas poniendo entre corchetes nuestra interpretación, añadiendo unas interrogaciones a aquellas de las que no nos hallábamos muy seguros, hasta una comprobación posterior. Por el contrario, las lecturas que resultan sorprendentes, o para las que no tenemos una solución satisfactoria, van seguidas de unas interrogaciones entre corchetes, o de un *sic*, con el fin de que el usuario no tenga dudas sobre lo que lee. Hemos optado por ofrecer a continuación de la palabra su reconstrucción íntegra entre corchetes, y no poner sólo la letra o signo cambiado en su interior, de modo que si en algún momento se realiza una consulta automática de una cadena de caracteres no se nos escamoteen las soluciones porque no se haya tenido en cuenta un corchete que no puede estar previsto en la cadena buscada. Esta solución es tanto más necesaria cuando lo modificado es la palabra que figura como entrada del diccionario: hemos respetado el original, pero en el *NTLE* aparece como lema la solución correcta.

Como se ve, los textos no se han tecleado tal cual figuran en los originales. Se han modificado la puntuación y la acentuación, para adecuarlas a las reglas y usos actuales en español, no así en las otras lenguas, lo cual hubiera requerido un esfuerzo para el que no estábamos capacitados, ni para el que teníamos medios personales y materiales, aunque en determinadas ocasiones, y para facilitar la lectura, se ha cambiado la puntuación o la acentuación, o se han introducido correcciones entre corchetes, pero, repito, sólo de manera muy esporádica. Por eso, el lector encontrará en lo escrito en lenguas diferentes al español unas tildes o unos signos de puntuación que resultan chocantes (por ejemplo, tildes en algunos textos latinos, empleos de diacríticos que no se corresponden con los actuales en italiano o en francés, etc., por no hablar de los signos de puntuación). Esta tarea, que parecía fácil, se nos vino a complicar con la aparición de las nuevas normas ortográficas de la Academia (RAE 2000). Tuvimos que plantearnos si volver atrás en todo lo que teníamos realizado hasta entonces, que era mucho, o seguir adelante con los mismos criterios. Optamos por esta última posibilidad para no estar tejiendo y destejiendo de modo continuo, lo que a la postre, iría en contra del desarrollo del trabajo,

y nos alejaría el puerto al que queríamos llegar, todavía lejano. Además, el mantenimiento de la ortografía anterior no suponía un atentado demasiado grande contra la lengua, de ninguna manera impediría la lectura de la obra, ya que los usuarios, al menos los actuales, están acostumbrados al sistema conservado, y sabrían comprender lo sucedido.

En este proceso también se comprobaban las remisiones que hacían los autores a otras partes de sus diccionarios, señalándose aquellos casos en que se trataba de pistas perdidas, cuando no era un simple error, habiéndose corregido como los demás.

También se desarrollaron las abreviaturas españolas para hacer fácilmente comprensible el texto, excepto las que son habituales hoy, como *S.*, *Sr.*, etc. A la vez, se fueron incluyendo remisiones cuando en el interior del texto reproducido se definía otra palabra, o había una referencia a ella digna de ser tenida en cuenta, o se utilizaban varios equivalentes españoles para una forma extranjera, o la entrada española contenía varias voces, y tantas otras posibilidades más. Se ha puesto especial cuidado para que estas remisiones no se confundan con las que se ponían los propios autores para guiar al usuario en la consulta de su obra; éstas aparecen con la misma tipografía empleada para el texto de los diccionarios, mientras que las nuestras van con un tipo de letra diferenciado.

En definitiva, hemos seguido unos criterios que resultan habituales en la edición de textos, y que no sorprenderán al lector experimentado. En este sentido, quiero dejar claro que las erratas muy evidentes se han corregido sin más, aunque en pocos casos. Además de todo ello, durante el proceso de teclado se sustituía el signo tironiano de acuerdo con la norma en la lengua correspondiente, excepto en inglés, por el uso que sigue teniendo.

El teclado resultó verdaderamente complicado con el *Vocabulista aráuigo en letra castellana* de fray Pedro de Alcalá (1505). La ventaja inicial que presentaba esta obra, al hallarse las voces árabes transliteradas en nuestro alfabeto, se vio desvanecida inmediatamente, ya que el autor empleó unos signos especiales para determinados sonidos, que no se encuentran en la tabla de signos ASCII que hemos manejado, la más reducida, para, como siempre, evitar problemas. La solución que adoptamos fue el empleo de unas combinaciones de signos que sustituyeran a los del autor, que han sido, a su vez, cambiados por los correctos en el tratamiento final del *NTLE*, antes de su envío a la imprenta. Tuvimos la suerte de contar con una hispanista marroquí para el teclado de esa obra, con lo que las dudas e inconvenientes que podría haber tenido cualquiera de nuestros colaboradores se solucionaron en el acto. Creo que éste es el mo-

mento oportuno para decir que ha sido la única ocasión en que durante el proceso de teclado hemos debido acudir, fuera de nuestro equipo, a un especialista en lenguas que no fuesen la nuestra. Nuestros colaboradores, afortunadamente, poseen unos conocimientos sobrados de otros idiomas para las tareas realizadas, y solamente el flamenco ha ofrecido alguna resistencia, no demasiada, pues, para nuestra sorpresa, una de nuestras colaboradoras se había iniciado en el conocimiento de esa lengua, y varios miembros del equipo dominaban a la perfección el alemán. Por lo demás, el latín, el griego, el francés, el inglés, el catalán, el gallego o el portugués no han presentado especiales dificultades, y sólo las pocas palabras en hebreo que hay en el *Tesoro* de Covarrubias (1611), y las escasas del Minsheu (1617) o del manuscrito de Bartolomé Valverde (1760?) nos han llevado a consultas externas.

A la vez que se realizaba el proceso de codificación, se avanzaba ya algún paso en la redacción del *NTLE*, como fue el ir efectuando una incipiente lematización, para la que se reservó un campo especial: los sustantivos y adjetivos irían en singular, estos últimos bajo la forma masculina, con la terminación femenina si fuese el caso, las formas de los paradigmas verbales se lematizaban en infinitivo, etc. Como no es difícil de sospechar, las unidades fraseológicas han presentado no pocas dificultades. Se ha procurado lematizarlas bajo la entrada que tenía el autor, salvo en aquellas ocasiones en que aparecían bajo un elemento de relación, en cuyo caso se han llevado a la palabra que soportara la mayor carga semántica de la expresión. Cuando hemos detectado que una construcción de este tipo aparece bajo entradas diferentes, porque así lo hacían los autores, hemos procurado poner unas referencias cruzadas, con el fin de que no se pierdan las informaciones. En general, hemos marcado los refranes y frases proverbiales entre unas comillas de valor, no así las frases hechas y las locuciones, máxime cuando el *DRAE* las registra como tales. Tampoco llevan una señal especial las numerosas frases y ejemplos que figuran en algunos repertorios como meros modelos de uso. En este sentido, he de decir que se han detectado algunos refranes y frases proverbiales que se presentaban truncados, y que también se han marcado, para mantener una cierta coherencia con los que no están truncados en los mismos artículos del *Tesoro*.

De modo simultáneo al teclado, se iban incorporando unas referencias internas que nos ayudasen más adelante en la agrupación de variantes, o para que sirviesen de llamadas de atención. Esas referencias se han revisado en la corrección final, para mantener únicamente como remisiones internas aquellas que fuesen necesarias y eliminar las que

sólo sirvieron para el desarrollo interno de nuestro trabajo, que no debían trascender el ámbito de la redacción. Una vez finalizado todo el proceso de corrección, en la Universidad Politécnica de Madrid nos han preparado un nuevo programa que efectúa una comprobación de todas las remisiones que hemos dejado, para que remitan al artículo correspondiente, y no a una variante que se haya lematizado bajo una entrada diferente o tenga un error de cualquier tipo, realizando los cambios correspondientes. Además, incorpora remisiones desde todas las variantes documentadas en cada autor, de modo que el usuario se encuentre siempre asistido en sus búsquedas. Por otra parte, genera un informe indicando aquellos artículos en cuyo interior hay más de tres remisiones a un mismo lugar, que hemos empleado en la revisión final para prescindir de aquellas que figuraban en cinco o más autores, sustituyéndolas por una genérica al final del artículo. Esto mismo se ha hecho en aquellos artículos con menos remisiones pero que contaban con un número reducido de fuentes, o que enviaban a lugares con pocas autoridades. Sin embargo, cuando se remitía desde un artículo con muchas fuentes, o a uno de esas mismas características, se han mantenido en cada autor, para no obligar al usuario a una tediosa búsqueda.

Los diccionarios, poco a poco, iban siendo tecleados, pero eso no era sino el comienzo de un proceso, que ya en sus inicios era pesado. Una vez finalizada la codificación electrónica de cada obra, los responsables del proyecto procedíamos a su lectura y corrección sobre papel, que devolvíamos a los colaboradores para que efectuasen las modificaciones señaladas. Cuando se concluía este proceso, se archivaba una copia, y se procedía a una alfabetización del contenido, que sufría un nuevo tratamiento, para unificar bajo una sola entrada toda la información del diccionario, y para eliminar las repeticiones innecesarias, más frecuentes en los repertorios de carácter temático que en los estrictamente alfabéticos. Después, se ponían las marcas de identificación del texto y se guardaba una copia de esta versión, que se convertía a un formato TXT, también guardado por separado, que era el empleado para ir cargando la base de datos.

Ya avanzada esta fase del proceso tuvimos un gran tropiezo imprevisto, que supusimos debido a nuestra propia inexperiencia: comprobamos que las informaciones de la base de datos no eran las mismas que había en la versión del tratamiento de textos, pues el contenido de algunos campos estaba truncado. No sabíamos a qué se debían las diferencias, y lo achacamos a algún error cometido inadvertidamente, por lo que reiniciamos una nueva carga de datos. Hicimos los controles pertinentes, y vimos que los errores detectados no se habían producido. Pero el alivio

duró poco y sonó otra alarma. En el nuevo proceso de carga volvía a suceder lo mismo, desaparecían los errores, pero se presentaban otros que no había antes. Ya no quedaba duda, el programa realizaba unos cortes aleatorios, probablemente debidos a la cantidad de información trasvasada. Recurrimos entonces al Grupo de Estructura de Datos de la Universidad de Las Palmas de Gran Canaria, que ya me había ayudado en las tareas informáticas del *Tesoro léxico de las hablas andaluzas*, y nuestras sospechas fueron confirmadas. A partir de ese instante la colaboración ha sido intensa, y las herramientas de que hemos ido disponiendo son resultado de su generosa asistencia, aunque en los últimos meses hemos recurrido a informáticos de la UPM, como he dicho, por la comodidad de la proximidad, y la necesidad del contacto directo y continuo. Fueron aquellos quienes nos propusieron la solución adoptada, y la importación de datos, desde entonces, no se ha hecho directamente empleando los recursos de *Access*, sino mediante un programa diseñado *ad hoc*, que es el que lleva la información desde el tratamiento de textos a la base de datos, con el fin de que no se pierda nada, y, para mayor seguridad, a continuación realiza una comparación para comprobar que lo que ha llegado a la base de datos sea igual a lo que se deseaba trasvasar. El proceso es extraordinariamente rápido y, sobre todo, nos proporciona una seguridad en nuestros medios que habíamos comenzado a perder. Pasado el tiempo he constatado que los errores no eran culpa nuestra, ni se debían a códigos mal empleados, pues otras personas que han llevado a cabo procesos similares en otras empresas lexicográficas se han tropezado con idénticos inconvenientes. Tal vez se trate de un defecto en el diseño del programa, cuestión sobre la que no he querido indagar más, una vez que los problemas sobrevenidos estaban ya resueltos.

En el momento en que consideramos que disponíamos de una cantidad significativa de obras en la base de datos, procedimos a generar un informe para ir viendo los resultados parciales, y animarnos en nuestro quehacer. Así, a finales de 1999 (tras varios años de intensos trabajos) elaboramos el primer borrador del *NTLE*, pasando todo el contenido de la base de datos a un tratamiento de textos, *Word* de Microsoft. Los cientos y cientos de páginas producidos nos causó asombro, pero también nos mostraba que el camino que quedaba por recorrer aún era largo.

Cuatro años más de esfuerzos nos llevaron a realizar un nuevo ensayo del resultado final, y vimos, nuevamente, cómo se producían errores en la carga de la base de datos, si bien esta vez ya no eran imputables a fallos en los programas informáticos, sino a nosotros mismos: se habían producido durante el proceso de tecleado. Esta alerta se produjo cuando

las marcas para separar campos no coincidían con los de la base de datos. Se trataba de un mero problema formal fácilmente subsanable. Pero, a la vez, nos percatamos también de la presencia de otras faltas formales, como equivocaciones en las secuencias de signos y otros errores de consistencia. Afortunadamente, se podían sistematizar, y, por tanto, corregir de una forma automatizada. El grupo informático de Las Palmas se dispuso a prepararnos un nuevo programa que comprobara si los separadores de campos estaban bien puestos, y otro que comprobara si las marcas sistemáticas que habíamos ido utilizando habían sido tecleadas correctamente. A partir de ese momento, comenzamos a buscar y detectar otros errores formales que pudiesen ser corregidos de una manera automática, o con la asistencia de sistemas informáticos, lo que nos llevó a solicitar nuevos programas. En realidad, es un solo programa que va realizando múltiples tareas encadenadas, y en un orden estricto para que unas instrucciones no puedan afectar a los cambios realizados por otras. De este modo, se incrementó el programa para detectar que siempre que se abra un paréntesis, un corchete, un signo de interrogación o de exclamación, unas comillas, etc., estén cerrados en el mismo campo, que los diferentes campos terminen de la misma manera, en especial que al finalizar el texto de cada artículo aparezca un punto, que no surjan espacios en blanco delante de los signos de puntuación, y que haya uno tras ellos (menos para el punto final), que no aparezcan dos espacios en blanco consecutivos, que los puntos suspensivos sean tres, etc., etc. Son pequeños detalles cuya corrección, en el conjunto del *NTLE*, hubiera requerido una gran dedicación de haberla realizado manualmente, y, lo que es peor, sin la garantía del resultado final. De este modo, hemos podido disponer de un corrector formal que se ha ido perfeccionando continuamente, hasta en doce versiones diferentes, algunas de las cuales suponían una marcha atrás, pues, a veces, los requerimientos para una corrección concreta afectaban a otros lugares, produciéndose cambios en exceso, o se introducían elementos que no debían aparecer. En definitiva, se trataba de ir afinando la herramienta y puliendo los errores en su diseño o en la apreciación de las tareas que se requerían.

La utilidad de este sistema es innegable, y desde los primeros instantes nos produjo una gran fascinación por las ventajas que presentaba, no sólo porque nos exoneraba de una gran cantidad de trabajo, a la vez que nos proporcionaba una considerable seguridad sobre el contenido formal de los originales, lo que nos permitió en la posterior fase de corrección manual centrar nuestra atención en otros aspectos. Por otra parte, la rapidez del mecanismo resultaba asombrosa, pues la corrección de cual-

quier diccionario no le ocupaba más de un minuto, además de generar un informe con las correcciones realizadas para que nos cerciorásemos de lo que había hecho, al tiempo que señalaba lugares en que no se habían realizado modificaciones aunque eran posibles, y marcaba otros que se debían comprobar y corregir manualmente. Esta otra tarea, en cambio, requería un tiempo considerable, infinito en comparación con la rapidez de la comprobación automática.

La alegría que nos proporcionaba el funcionamiento de esa herramienta iba siempre acompañada de un poco de amargura, pues pese a los esfuerzos realizados, al cuidado con que se efectuaba el trabajo, en cada texto que se hacía pasar por el programa se descubrían numerosos errores e inconsistencias, que, cuando el repertorio era extenso, resultaban ser varios cientos. Aquello que creíamos bien hecho, sometido a varias correcciones por distintas personas, nunca alcanzaba la perfección deseada, y nuestra euforia por ver cerca el final del tratamiento de cada diccionario se veía atemperada siempre que nos sometíamos al juicio informático, inexorable. De todos modos, no podíamos renunciar a tan valiosa ayuda para que el contenido de las obras que iban pasando a la base de datos fuese el deseado, y cumpliera, al menos, unos estrictos requisitos formales. Era la única manera de ir avanzando con un pie firme sin necesidad de mirar continuamente hacia atrás.

Los problemas formales, de organización y de informatización, como los descritos hasta ahora, con ser importantes y afectar de manera directa a la marcha del trabajo, evidentemente, no han sido los únicos con los que nos hemos ido encontrando durante la elaboración del *NTLE*, pues los ha habido de otra índole que también han incidido en la organización del trabajo, el general y el cotidiano.

Una de las decisiones que se nos antojaba fácil era la de qué repertorios habrían de formar parte del *Tesoro*. Por el carácter del proyecto, y lo ambicioso de nuestro objetivo, que ha continuado siéndolo hasta el final, la respuesta era inmediata: todos los anteriores a la publicación del *Diccionario de Autoridades* (1726-1739), así de sencillo. Sin embargo, como siempre sucede, esa simplicidad era solamente aparente. Resultaba necesario precisar, y elaborar una lista de obras para trabajar sobre ella. El elenco de lo manejado por Gili Gaya (1960) para su *Tesoro* podía constituir un buen punto de arranque, si bien pronto apreciamos los defectos que encerraba aquella nómina que nos parecía extensa. Por un lado, algunas obras estaban confundidas, otras no eran las primeras ediciones, de unas pocas había manejado varias ediciones, otras estaban en la relación pero no las encontrábamos citadas en el interior de los artículos... Había más.

No se recogían *todos* los diccionarios del periodo abarcado, y faltaban algunos anteriores, como pusieron de manifiesto quienes se encargaron de hacer reseñas del *Tesoro* cuando fue publicado (por ejemplo, Alonso 1951, Colón 1956 y Pottier 1962). Además, aquella era una lista de diccionarios en la que también figuraban obras que no eran de carácter lexicográfico, ni mucho menos, por más que su interés para la historia de nuestro léxico resultaba incuestionable, por lo que su incorporación, fuera de otras consideraciones, estaba sobradamente justificada desde el punto de vista lexicológico.

Ante todo ello, se imponía elaborar nuestra propia lista de obras, tarea a la que nos dedicamos de inmediato, y que no hemos abandonado sino en el último instante, al concluir el proceso de redacción, tal ha sido nuestro empeño por ser consecuentes con nuestras decisiones, y por proporcionar al usuario la mayor cantidad posible de materiales, siempre que no se entorpeciera el manejo de la obra y que no se acumulasen informaciones de manera farragosa hasta ocultar las informaciones verdaderamente interesantes. Éste es el motivo por el que no damos cuenta de algunas nomenclaturas que, en la parte española, no hacen sino repetir cabalmente lo que ya figura en los repertorios que toman como fuente, con escasísimos o nulos cambios, además, en un léxico, que, por la naturaleza de las obras, es el más cotidiano, por lo que aparece abundantemente documentado en la lexicografía y en la historia de la lengua.

En otros lugares he descrito cómo se ha ido configurando el conjunto de obras que ha terminado por ser el cuerpo del NTLE³. Si don Samuel había dado cabida a textos no lexicográficos, y, como he dicho, su presencia en su *Tesoro* quedaba justificada, nosotros no sólo mantuvimos aquella decisión, sino que la extendimos, de modo que retomamos los mismos textos y añadimos alguno más. De este modo, si había razones para recoger el léxico del *Dioscórides* de Andrés Laguna (1555)⁴, ya que incluye unas tablas en varias lenguas de los nombres registrados en su interior (entre ellas, por supuesto, de los españoles), no tenía sentido prescindir de los términos españoles que puso Antonio de Nebrija (1518) en la edición del *Dioscórides* impresa en 1518 por Guillén Arnao de Brocar, partiendo de la edición de Ruelio, publicada poco antes en París. Y ello nos llevó a incorporar también las voces españolas que Amato Lusitano (1553) había incluido en su versión de la obra, aparecida sólo dos años antes que la de Laguna. Los motivos que amparaban la inclusión de las palabras españolas con-

3. Véanse los trabajos que cito en la primera nota.

4. Gili Gaya empleó la edición de 1570, y no la primera, de 1555, como hemos hecho nosotros.

tenidas en esas obras son los mismos que autorizan la de las que hay en la obra de Conrad Gesner (1551-1587) (cfr. Alvar Ezquerro, 2002-2004), o en la *Historia de las yeruas y plantas* de Juan Jarava (1557)⁵, o en el anterior *Modus faciendi cum ordine medicandi* de fray Bernardino de Laredo (1527) –la primera farmacopea escrita en español–, que también tienen índices de las palabras españolas registradas en su interior. Y desde ellas llegamos al *Tesoro de medicinas para todas enfermedades* de Gregorio López (1672)⁶. De esta manera, en el interior de nuestro *NTLE* se halla un tesoro léxico de la medicina de la época, que, sin duda, presenta un notable interés y puede ser de gran utilidad a quienes se interesan por esa parcela del léxico. Como nos ha sucedido con otros grupos de obras, han sido más las consultadas, pero no han pasado a engrosar las páginas del *NTLE* por no ser de carácter lexicográfico y no presentar un índice de voces que avalase su incorporación, pese a que el vocabulario contenido en ellas sea de interés, como por ejemplo, el *Tractado de las drogas y medicinas de las Indias Orientales* de Cristóbal Acosta (1578) (cfr. Alvar Ezquerro 2006b).

Ciertamente, ese grupo de obras constituye un conjunto homogéneo de fuentes para un ámbito de léxico específico, pero no es el único que hay en nuestro *NTLE*, aunque sí el más rico. No carece de importancia el dedicado a la terminología marinera, tanto es así que Lidio Nieto (2002) dio un avance de esa parte en su *Tesoro lexicográfico del español marinerero anterior a 1726*, aunque después se ha añadido alguna obra más. En otros ámbitos, como el del derecho, las bellas artes o las ciencias exactas, la riqueza no es tan grande, y no por falta de empeño por nuestra parte. Los repertorios de léxico técnico se desarrollan después del periodo abarcado en el *NTLE*, y son muy pocos los elencos de que hemos podido disponer. Es lo mismo que sucede con las listas de palabras de uso regional, pues antes de 1726 solamente aparece la de Gerónimo de Blancas (1641)⁷, mientras que sí hay algún elenco de voces anticuadas, así como de carácter etimológico –según se entendía la etimología en la época–, aunque una buena porción de ellos sólo se han conservado manuscritos, pese a su trascendencia, como sucedió con el repertorio de voces de origen ára-

5. Disponemos de una edición moderna de M^a Jesús Mancho Duque (2005), con varios estudios preliminares donde se desentrañan los problemas relativos a la obra y su contenido.

6. De 1672, aunque redactado a finales del siglo anterior. El manuscrito hológrafo ha sido editado por Francisco Guerra (1982).

7. Son las voces aragonesas que se encuentran en el “Índice donde se declaran algunos vocablos aragoneses antiguos [...]”, que puso en las *Coronaciones de los serenísimos reyes de Aragón*, publicadas en 1641, si bien fueron escritas en 1583.

be del P. Guadix (1593)⁸ o las etimologías de Bartolomé Valverde (1760?)⁹ o Francisco del Rosal (1601)¹⁰, por no señalar nada más que un par de ellos.

Por supuesto, en el *NTLE* están recogidos los diccionarios monolingües anteriores al *Diccionario de Autoridades* (1726-1739), no solamente los específicos, sino los más extensos, así como los bilingües y plurilingües que contienen el español, sean alfabéticos, sean temáticos. La cantidad de obras recogidas es un centenar y medio, aunque son más las consultadas. De ello se da cuenta en las páginas preliminares del *NTLE*.

La lista de obras que iban a configurar el contenido del *NTLE* quedó esbozada en un primer momento, pero no cerrada. Las consultas de fondos bibliográficos, de repertorios de toda índole, de catálogos de la imprenta en España y en otros países, de variados trabajos científicos, no solamente de lexicología y lexicografía, sino también de historiografía lingüística, de historia de la ciencia, etc., etc., así como las investigaciones que desarrollan nuestros colaboradores para sus tesis doctorales y otros trabajos de investigación, nos han hecho descubrir de forma continua obras cuyo examen era necesario para decidir si su contenido debía verse en el *Tesoro* o no, con lo que han quedado por el camino muchos esfuerzos, y algún que otro libro. De unos ni siquiera se mantiene un rastro en el *NTLE*, de otros se da una somera explicación de los motivos por los que no se han tenido en consideración. De todos modos, alguna que otra obra de las que no han pasado al interior del *NTLE* ha sido objeto de estudio por parte de los miembros del equipo, por lo que, con los trabajos que ya se van derivando de nuestra empresa, no se ha perdido del todo el léxico de esas otras obras no incorporadas, y nuestros esfuerzos han dado frutos, aunque no en el interior del *NTLE*; es lo ocurrido, por ejemplo, con la breve lista de palabras que aparece en *The Key of the Spanish Tongve* de Lewis Owen (1605)¹¹, que es solamente inglés-español, con la *Janua linguarum* (1611) de Salamanca (cfr. Alvar Ezquerro 1992), latín-español, o con el vocabulario francés-español que aparece en la gramática de Monsieur de Maunory (1701). Este continuo desvelo por las fuentes ha hecho que la nómina de obras no se haya cerrado hasta el último instante, cuando la corrección final de los datos se hallaba muy avanzada, con la incorporación

8. *Recopilación de algunos nombres arábigos* [...], ms. 59-I-24 de la Biblioteca Colombina (Sevilla), manejado por los lexicógrafos posteriores.

9. De la obra se conservan, al menos, dos copias más manuscritas.

10. *Origen y etimología de todos los vocablos originales de la lengua castellana*, ms. 6929-T.127 de la Biblioteca Nacional, Madrid. También hay más copias de él. Hay una edición moderna con estudio de Enrique Gómez Aguado (1992) y reproducción del manuscrito citado.

11. Sobre ella, véase Nieto Jiménez / Alvar Ezquerro (2002).

de un pequeño repertorio multilingüe de la primera mitad del siglo XVI, el *Quinque linguarum* (1534), cuya copia tardó en llegar, por lo que habíamos introducido una edición posterior, en seis lenguas, de 1541, con escasas diferencias en la parte española (*Vocabulario de seis linguaies* 1541)¹².

Con la lista de obras esbozada comenzó una nueva complicación, la de conseguir reproducciones sobre las que trabajar. En unos casos, no muchos, había reproducciones facsimilares sobre las que podríamos trabajar, o ediciones modernas, por más que su grado de fiabilidad no fuese muy elevado, como nos ha sucedido con la citada de Francisco Guerra (1982) del *Tesoro de medicinas para todas enfermedades* de Gregorio López (1672). Pero no siempre hemos empleado las ediciones modernas, pues unas veces hemos acudido a los originales, como en el *Tesoro* de Covarrubias (1611), ya que teníamos a mano un ejemplar de la primera edición, si bien hemos consultado la edición de Martín de Riquer en caso de duda; diferente ha sido el caso de los *Diez privilegios para mugeres preñadas* de Juan Alonso y de los Ruyzes de Fontecha (1606), pues, una vez realizado nuestro teclado, ha aparecido una edición moderna (Zabía Lasala 1999), que nos ha servido de guía en el proceso de corrección posterior.

Por otra parte, teníamos fotocopias de otras obras, tampoco muchas, que conservábamos de trabajos nuestros anteriores, o que habíamos ido acumulando por diversos motivos. Con éstas, y las ediciones de que acabo de hablar, era suficiente para iniciar el teclado de los contenidos, aunque fuese de una manera desordenada, pues siempre que nos ha sido factible, hemos avanzado de manera cronológica, entre otras razones, para que aquellas obras que fueron utilizadas como fuente por otras posteriores nos sirviesen de guía o facilitasen el trabajo.

Del resto del elenco comenzamos a buscar ejemplares en las bibliotecas públicas. Primero acudíamos a las madrileñas, por ser las más cercanas y ricas en fondos. Si no encontrábamos en ellas la obra en cuestión, rastreábamos su presencia en otras bibliotecas españolas, y, en caso negativo, nuestras pesquisas se dirigían al extranjero. Incluso en alguna ocasión hemos entrado en bibliotecas particulares, españolas y extranjeras, tal ha sido la intensidad de la búsqueda, para la que nos han ayudado grandemente los repertorios bibliográficos que han ido apareciendo durante los últimos años, y los ficheros on-line de las principales bibliotecas, aunque no siempre ha sido ésa la manera de llegar hasta el ejemplar deseado.

12. Es preciso advertir que tanto la versión en cinco lenguas como la de seis coinciden en su primera parte con el *Quinque linguarum* (1526), aunque no en la segunda, que presenta diferencias, por lo que la hemos trasladado al *NTLE*, donde el lector podrá comprobar los cambios.

Con la localización de ejemplares tampoco se terminaba el proceso de búsqueda, salvo, evidentemente, de los manuscritos. Era preciso conocer el estado de los libros para que no faltasen páginas y para que el texto no estuviese mutilado. Por ello, no han sido pocas las veces en que, tras comprobar la reproducción que nos llegaba, y cerciorarnos del estado del original, hemos tenido que reiniciar el proceso de búsqueda. En otras ocasiones, las reproducciones estaban mal hechas, con los márgenes cortados, hojas sin reproducir, texto desenfocado o con tantas manchas que resultaba ilegible, y con cuantas complicaciones que ni nosotros mismos éramos capaces de imaginar. Entonces teníamos que volver a la biblioteca para reclamar, con el malestar de todos, como es fácil de suponer. Aunque estos problemas causaban trastornos, especialmente porque debíamos adaptar el ritmo de las tareas a las circunstancias, era tal la cantidad de obras que teníamos entre manos, y la cantidad de correcciones y filtros que se efectuaban, que nunca se constituyeron en un gran inconveniente para el progreso de los quehaceres.

Pese a que los servicios de información y de reproducción de las bibliotecas funcionan razonablemente bien, la obtención de fotocopias ha llegado a constituir un verdadero conjunto de adversidades. En este sentido, nunca se nos olvidarán las acaecidas con el *Diccionario español-francés* de fr. Pierre Seguin, que a punto estuvieron de hacernos abandonar y de provocar que su contenido no figurase en el *NTLE*, pero nuestro tesón estaba hecho a toda prueba, por difíciles que se planteasen las contrariedades. Se trata de un manuscrito de la Biblioteca Nacional de Francia, empleado por Gili Gaya en su *Tesoro*. Solicitamos su reproducción tanto directamente como con la ayuda de nuestra Biblioteca Nacional, y fueron nuestros colaboradores a París en varias ocasiones, sin que lo consiguiéramos; por fin, y tras una reclamación a la dirección de la biblioteca al cabo de los años, pudimos saber que había cambiado la signatura de la obra, con lo que volvimos a iniciar el proceso, pero la reproducción que nos enviaron no tenía nada que ver con nuestro manuscrito. Vuelta a las quejas, y admitieron haber cometido un error en la signatura que nos comunicaron. Para poder seguir adelante tuvimos que devolver las fotocopias recibidas, pero las nuevas no llegaban, pese a las promesas, así que aproveché un viaje a París para averiguar lo sucedido; me dijeron que el envío se había efectuado unos días antes, pero en Madrid no recibía nada. Nuevas quejas, y cada vez más serias, hasta conseguir las dichas fotocopias. Habían transcurrido unos cuantos años desde la primera solicitud. En honor a la verdad, y en defensa de las bibliotecas y de sus trabajadores, he de decir que ésta no ha sido la tónica general, aunque nunca nos han llegado las copias con

la rapidez deseada, con algunas excepciones, como la de la Real Academia Española. Si el funcionamiento de las bibliotecas no hubiese sido el adecuado, nuestro trabajo no hubiera podido llevarse a cabo, o lo hubiéramos hecho de manera más lenta y penosa, y no hubiéramos podido contar con reproducciones de ejemplares conservados en bibliotecas tanto españolas como extranjeras. Las bibliotecas españolas han sido fundamentalmente madrileñas (principalmente la Biblioteca Nacional, la de la Real Academia Española y la Biblioteca Histórica Marqués de Valdecilla de la Universidad Complutense de Madrid, pero también la Biblioteca de Palacio, la de la Real Academia de la Historia, Biblioteca General de Humanidades del CSIC, Biblioteca del Instituto de Filología del CSIC, Biblioteca Histórica Municipal, la del Museo Naval, la de la Fundación Universitaria Española, la del Monasterio de El Escorial), aunque también de otros lugares de nuestra geografía (Biblioteca Universitaria de Salamanca, Biblioteca Universitaria de Granada, Biblioteca de la Universidad de Barcelona, Biblioteca Pública de Málaga, Biblioteca Colombina de Sevilla, Biblioteca Valenciana de Valencia). Una buena parte de las bibliotecas extranjeras han sido de Francia (Bibliothèque Nationale de France, Bibliothèque de l'Arsenal, Bibliothèque Mazarine, Bibliothèque Sainte Geneviève, en París, así como la Bibliothèque Municipale de Rouen y la Bibliothèque Municipale de Montauban), pero también hemos acudido a las de Alemania (Bayerische Staatsbibliothek de Munich, Herzog August Bibliothek de Wolfenbüttel, Stadtbibliothek de Nürnberg, Universitäts Bibliothek de Augsburgo), de Bélgica (Universidad de Amberes, Rijks Universiteit de Gante), de Italia (Biblioteca Nazionale Centrale de Roma, Biblioteca Nazionale Marciana de Venecia), de Portugal (Biblioteca Nacional de Lisboa), de Gran Bretaña (British Library), de Austria (Österreichische Nationalbibliothek de Viena), de Argentina (Biblioteca Nacional de Buenos Aires), y de Estados Unidos (The Hispanic Society of America), así como alguna más particular en España y en Francia. Además, a muchas de las bibliotecas reseñadas, tanto españolas como extranjeras (en Europa y en Estados Unidos) hemos acudido personalmente para comprobar los ejemplares, o para cotejar nuestras lecturas.

En todos los casos, hemos utilizado reproducciones de las primeras ediciones de las obras, dejando de lado otras ediciones, cuando las hubo, ya que en una buena parte de ellas los cambios introducidos no eran muchos, y, de todas maneras, no se trataba de realizar ediciones críticas para saber la evolución de la obra, por más que el *NTLE* sea la base fundamental para realizarlas. Únicamente se ha hecho una excepción con el diccionario nebrisense, por su importancia para la lexicografía posterior, en que, además de la primera edición, hemos incorporado las variantes

(sólo las variantes) de la edición de 1520, última hecha en vida del autor y corregida por él, que coincide, sustancialmente, con la de 1516, también corregida por Nebrija¹³.

Aunque se ha introducido todo el contenido de los repertorios seleccionados, como se dijo antes, en ocasiones se ha suprimido texto (especialmente en las obras no lexicográficas), cuando resultaba irrelevante para el conocimiento de la palabra. Como el *NTLE* es un repertorio léxico, no han pasado a él los numerosos nombres propios que aparecen en las fuentes, sean geográficos, de personas, de seres mitológicos, de animales fantásticos, de constelaciones y cuerpos celestes, etc., salvo, evidentemente, cuando están lexicalizados, o cuando constituyen parte de una expresión pluriverbal fijada.

Por lo que se refiere a los diccionarios bilingües y plurilingües, únicamente hemos tenido en consideración la parte español-otra lengua, no la parte en la que nuestra lengua es a la que se vierten las palabras. No se nos oculta que en esa otra parte de los diccionarios hay léxico de enorme interés para la historia de nuestra lengua, pero no es menos cierto que la información, en bastantes ocasiones, hubiese sido muy repetitiva, alargando en exceso el *NTLE*, y, además, nuestros recursos tanto humanos como económicos resultaban limitados para acometer esa tarea, que queda para fases posteriores del proyecto, si es que retomamos energías para continuarlo. De todos modos, son muchas las excepciones que se han hecho a este principio, ya que hemos llevado al *NTLE* las voces españolas de todos los repertorios anteriores a 1500, independientemente de su posición en la obra (como entrada o como lengua meta). Igualmente se ha hecho excepción de algunos repertorios posteriores, entre ellos todos los de Nebrija, que no sólo son sus dos diccionarios grandes –los más conocidos, por otra parte–, y los diccionarios que iniciaban la lexicografía bilingüe del español con otra lengua. El lugar que ocupa el español no se ha planteado para las nomenclaturas, ya que en ellas lo importante no es la posición de la lengua, sino el modo de ordenar las informaciones, que, al pasar al *NTLE*, se muestran según el orden alfabético.

En el momento en que se tuvieron seleccionadas todas las obras, y se hubieron tecleado, se comenzó el proceso de lematización definitiva, esbozado cuando se pasaban los textos a formato electrónico. La diferencia es que entonces se realizaba aisladamente para cada una de las obras manejadas, mientras que ahora se procedía de una manera conjunta, y

13. De esta última hay una edición debida a Gerald J. Macdonald (1981). El contenido léxico de esta salida ha sido analizado por María Lourdes García Macho (1993, 1995 y 2005).

definitiva. El problema con que nos hemos encontrado en esta fase ha sido el de las numerosas variantes gráficas que constan en los repertorios. Para facilitar la consulta del *NTLE*, las hemos agrupado todas bajo una sola entrada, remitiendo a ella desde el lugar alfabético que les correspondía a cada una de ellas. Es cierto que son numerosos los reenvíos, pero no lo es menos que de otra manera hubiéramos tenido una inmanejable red de referencias cruzadas que habrían dificultado la consulta de la obra, y habrían enfadado no poco al usuario. Como hizo Gili Gaya, podríamos haber dejado cada forma en el lugar que le hubiese correspondido por el orden del abecedario, pero así la información hubiese aparecido desperdigada en mil lugares de la obra, inconexos entre sí, tanto más cuanto más distantes eran las formas. Pero no nos pareció una forma adecuada de ofrecer los materiales.

La solución adoptada permite ver la historia lexicográfica de la palabra, además de presentar agrupadas las variantes tras la forma de la entrada. Para esta tarea se ha desarrollado otro programa informático en la Universidad Politécnica de Madrid que, además, resalta la forma coincidente con el lema –lo que permite, con un simple vistazo, saber si figura en las fuentes utilizadas–, o pone un asterisco tras él cuando no coincide con ninguna de las variantes registradas, lo cual puede ser de interés para los estudiosos que acudan a consultar el *NTLE*.

Gracias al sistema informático empleado, la información contenida en los diversos textos ha aparecido ordenada de manera cronológica, para lo cual cada obra va acompañada de una clave con la que se ordena. Si dos o más obras pertenecen a la misma fecha, esa clave, que no figura en la versión impresa, va acompañada de un dígito diferenciador. De esta manera, se ofrecen los datos en una secuencia histórica que vale para conocer no sólo la evolución formal de la palabra en cuestión, o su tratamiento lexicográfico, sino también para ver la filiación y deudas de unas obras con otras.

Cuando se hubo procedido a la lematización del conjunto de las informaciones, que en total han sobrepasado largamente los 600000 registros en la base de datos, se procedió a una impresión en papel, con un formato similar al que tiene la versión impresa, sobre la que se ha efectuado un doble proceso de corrección, para asegurarnos de la calidad del contenido, lo cual, por otro lado, es obvio. Esta tarea ha llevado más tiempo del previsto, por los errores detectados, lo que nos ha obligado a una tercera lectura, cuando creíamos que ya estaba el trabajo culminado. Sin dudas, los yerros e inconsistencias serán menores de este modo, aunque quedarán todavía, pues ninguna obra humana es perfecta.

Con todos los datos unificados, nos ha sido posible corregir errores que habían pasado desapercibidos al tratar las obras aisladamente, o que no habíamos podido solucionar a falta de mayor información. Ha sido también éste el momento en que hemos podido solventar muchas de las dudas que habían quedado durante el proceso de tecleado. Pero también nos ha obligado a unificar la puntuación en aquellos lugares en que lo contenido es similar, y que, al ser tratado independientemente, se había codificado según el criterio de cada colaborador.

Sabíamos de la dependencia de unas obras de otras, que, ahora, queda bien patente, pero también, al reunir todos los datos, hemos podido comprobar nuevas filiaciones y deudas. Queda, a partir de estos momentos, una enorme labor de estudio, pues antes no se habían entrevisto muchas de las concomitancias que se muestran con nitidez al presentar todos los diccionarios agrupados y ordenados, como se puede apreciar en la larga lexicografía con el inglés (cfr. Nieto Jiménez / Alvar Ezquerro 2006). En otras ocasiones se aprecia cómo se han repetido afirmaciones que no tenían consistencia, o que deben ser precisadas, como la deuda de Stevens respecto a Minsheu (cfr. Alvar Ezquerro / Nieto Jiménez 2003), o la de Pedro de Alcalá con respecto a Nebrija (cfr. Alvar Ezquerro, 2006c). “Disponer de un modo conjunto de todas las informaciones nos proporciona un excelente panorama de la historia de nuestro léxico y de la representación de cada una de las palabras a través de los diccionarios, pero además se convierte en un instrumento de primera magnitud para ver la filiación y la interdependencia de los diccionarios, y eso que únicamente hemos recogido las primeras ediciones de las obras, excepción hecha, como he comentado, de Nebrija, aunque en una sola de sus obras. Es más, las posibilidades que nos brinda una base de datos nos permite comparaciones entre obras, entre series de obras, entre repertorios de una misma época, o la búsqueda de soluciones para una voz de otra lengua en diversos elencos, e, incluso, rastrear la actividad de un lexicógrafo cuando es autor de varios repertorios, etc.” (Alvar Ezquerro 2006a: 55).

Ha sido ciertamente revelador comprobar cómo los errores y erratas se transmitían de un elenco a otro. En este sentido, el *NTLE* se convierte en un excelente instrumento para averiguar las fuentes empleadas por determinados autores, o para seguir la presencia de un lexicógrafo en los diccionarios posteriores, o para conocer qué edición de una obra pudieron manejar sus seguidores. En este sentido, cabe hablar de la fidelidad de los diccionarios ingleses con respecto a los anteriores, pese a las innovaciones que hay en todos ellos. En ocasiones, nuestro trabajo nos ha llevado a consultar no sólo varias ediciones de un mismo repertorio, sino

hasta diversos ejemplares de una misma edición, lo que nos ha permitido descubrir emisiones distintas, como las que detecté para la *Bibliotheca hispanica* de Richard Percivale (cfr. Alvar Ezquerro 2002: 175-176).

El trabajo está finalizado. Han sido muchos años de dedicación y de renunciaciones. Me gusta recordar las palabras del humanista Juan Lorenzo Palmireno en la nota “Al lector” que hay después del octavo abecedario del *Vocabulario del humanista*:

Mi propósito fue poner nueve abecedarios, tratando en el último las antiguallas que al principio prometo: pero quien tiene familia que sustentan en años tan caros, y sin Mecenate, que ayude al papel, por fuerza ha de boluer a tras; contemplando que pone sus ojos, salud, fama, y dinero a votos del vulgo: el qual con su acostumbrada ingratitud dize, que vendo muy caro este librero: en el qual, allende de los pescadores y caçadores que he estrenado, y combidado, para ver cómo quadraba lo que yo facua de los libros con lo que ellos experimentan. En sólo Hippólito Saluiano, y los otros autores que de peces escriuen gasté veynte y cinco ducados: y por no dexar agraviado al lector de lo que aquí esperaua, haña que yo tenga dinero, para facer a parte lo que le he prometido, lea entre tanto lo que se figue (Lorenzo Palmireno 1569: II, 114).

Con otras apariencias, bajo otras formas, en otros tiempos, las dificultades y los problemas continúan siendo los mismos. Dicho con la desesperación y el estoicismo del autor del *Eclesiastés* (1, 9), y como advertencia, y alivio, para quienes en el futuro intenten abrirse paso en este mundo: *nihil novum sub Sole*.

Esperamos que los esfuerzos no sean baldíos y sirvan para el mejor conocimiento de nuestro léxico, de nuestros diccionarios y de nuestra lengua. *Indocti discant, et ament meminisse periti* (“que los que no saben aprendan, y los instruidos gocen al recordar”), que cuentan decían los edictos romanos.