

LEXICOGRAFÍA E INFORMÁTICA EN EL NDHE

OCTAVIO PINILLOS LAFFON
NDHE de la Real Academia Española

La lexicografía actual plantea retos que difícilmente pueden ser solucionados por métodos exclusivamente manuales. Por esa razón la Academia emprendió hace más de quince años la modernización de sus métodos de trabajo con el auxilio de la informática.

Para entender la complejidad de los problemas que plantea la revisión de los diccionarios es necesario acercarse a la lexicografía desde el punto de vista del especialista. Si para el profano el diccionario es una recopilación de definiciones ordenadas alfabéticamente, para el lexicógrafo se trata de una obra compuesta por miles de textos interdependientes que mantienen un delicado equilibrio. La informática facilita el trabajo complejo en la explotación de estos textos.

En efecto, la informática, y más concretamente la tecnología de bases de datos, proporciona soluciones a muchos de los retos que plantea la lexicografía actual. Si en el diccionario tradicional cada artículo era tratado como un bloque de información textual marcado tipográficamente, en el diccionario informatizado los artículos pueden lograr organizarse en complejas estructuras de datos que facilitan la revisión y garantizan la integridad. La relación entre lexicografía e informática ha llegado a ser tan fuerte que se podría calificar de simbiótica. Tanto es así, que en ocasio-

nes me asalta la duda de si entre nosotros es la informática lexicográfica la que se beneficia de las teorías del Dr. Pascual, o por el contrario son sus postulados los que no se conciben sin el auxilio de la informática. El hecho es que esta ha tenido efectos tan beneficiosos como los siguientes:

Ha venido a cambiar de raíz los que parecían males endémicos de la Lexicografía, como es controlar –lo hacemos ahora mediante procedimientos informáticos– la integridad del sistema de referencias y el léxico empleado en las definiciones, lo cual evita, por ejemplo, la circularidad, las remisiones a artículos suprimidos, etc.

Permite, por otro lado, fomentar la coherencia en estas obras tan complejas como son los diccionarios, estableciendo patrones uniformes en el aparato de abreviaturas y marcas, que se reducen así a un universo controlado.

Resulta incluso beneficiosa para la organización del trabajo de redacción, ya que posibilita, por un lado, la creación de modelos de definición aplicables a grupos homogéneos de voces o acepciones, a la vez que hace posible la especialización del trabajo lexicográfico, una necesidad impuesta por el carácter multidisciplinar que se va imponiendo en los equipos de redacción.

También los tiempos de revisión del diccionario se reducen de manera apreciable mediante el empleo de técnicas informáticas que auxilian al lexicógrafo en la corrección de pruebas.

Y finalmente, posibilita la disminución del esfuerzo para la publicación, pues las ediciones se confeccionan automáticamente en formatos estandarizados compatibles con los modernos sistemas de edición electrónica y fotocomposición por ordenador. He de añadir que, fuera ya del marco de papel, nos encontramos con posibilidades que sorprenden a los propios lexicógrafos, al crear un tipo de texto que permite relacionar entre sí los mate-

riales de una forma que hubiera sido impensable en el formato en papel.

Las posibilidades que brinda la informática se extienden a un ámbito que podríamos considerar paralexigráfico, por cuanto sirve para apoyar lo propiamente lexicográfico. Se trata de la construcción de fuentes documentales en formato electrónico, a través de los corpus documentales, imprescindibles para la comprobación de la “realidad lingüística”, pues si un diccionario pretende reflejar la realidad de una lengua no puede obviar el uso de materiales de consulta de muy amplio espectro. Prácticamente todos los grandes diccionarios actuales recurren a ingentes cantidades de materiales de muy variada procedencia para la comprobación de los hechos lingüísticos. La Academia no podía obviar esa realidad, y ha construido una serie de corpus. Entre los que aquí interesan está el *Corpus del Diccionario Histórico (CDH)*, específicamente diseñado para las necesidades del *Nuevo diccionario histórico del español (NDHE)*, o el *Mapa de Dictionaries*, un mapa evolutivo del léxico moderno a partir de los diccionarios de la Real Academia Española.

A medio camino entre estos dos tipos de acciones informáticas estaría la manipulación de la información del corpus, que proporciona al lexicógrafo ventajas sustanciales para realizar su trabajo, por ejemplo la ordenación de los ejemplos bajo determinados criterios, o la posibilidad de mantener la integridad de las referencias establecidas entre textos y ejemplos.

¡Qué te voy a decir, querido Juan, que no sepas ya a estas alturas de tu vida en que hasta te has permitido meterte a lexicógrafo! Tú mismo has visto cómo las humanidades se han beneficiado de la informática en la medida en que los informáticos hemos sido capaces de entender las humanidades. Si bien es cierto que en ocasiones surge la impaciencia ante las limitaciones y servidumbres que impone la tecnología, una mirada retrospecti-

va basta para comprobar que si la Academia de los años noventa empezaba a servirse de la informática con cautela, la actual no tiene la menor duda de recurrir sistemáticamente a las nuevas tecnologías. Este despliegue tecnológico de la Academia ha permitido a personas de tu calidad humana y profesional conciliar disciplinas aparentemente contrapuestas.