



Julio Borrego y José Antonio Bartol en el claustro de la Universidad de Salamanca. ICAL

## Sabios de la lengua buscan el ADN de cada palabra

Equipos de lingüistas de las universidades de Salamanca y León recopilan 400 millones de vocablos en el Corpes XXI para su estudio en el futuro

Y. R. T. (ICAL) SALAMANCA

Las palabras expresan emociones, sentimientos, órdenes, pero también inquietan, movilizan, hacen sentir amor y odio, en definitiva son poderosas. De ahí la importancia de estudiarlas, investigarlas como si de una «fotografía fija» se tratara para, a partir de ellas, conocer cómo era la lengua en una época determinada.

Esa es la labor que llevan a cabo los profesionales que trabajan en el

desarrollo del Corpus del Español del siglo XXI, también denominado Corpes XXI, una investigación sobre el léxico y la gramática del idioma, que permite saber qué palabras se usan más frecuentemente en un determinado país, cuándo se introdujeron y si después fueron olvidadas.

Por ejemplo, si hay una palabra que está «de moda» hoy en día en España ésa es la de «populismo», y el Corpus del Español da la posibi-

lidad a quien quiera satisfacer su curiosidad de saber dónde se utiliza con más asiduidad, que en este caso es en países latinoamericanos como Bolivia y Venezuela, y si es realmente novedosa en España.

En el futuro, su uso en los medios de comunicación, publicaciones y programas de televisión podrá ser analizado, y algún investigador del mañana podrá comprobar que en 2015 comenzó a usarse de forma más frecuente en España.

En esta labor de recopilación, impulsada por la Real Academia de España, colabora un equipo de filólogos liderado por los profesores Julio Borrego y José Antonio Bartol de la Universidad de Salamanca, con expertos de la Universidad de León. Participan junto a docentes de otras cuatro instituciones académicas españolas.

Borrego detalla que el objetivo es reunir en 2018 un conjunto textual formado por 400 millones de vocablos y formas de la lengua de 500 millones de hispanohablantes. Aunque será una herramienta indispensable para los investigadores, afirma que es de acceso abierto y gratuito para toda la sociedad y, una vez que se conozcan sus posibilidades, cualquier ciudadano puede resolver sus dudas o curiosidad acerca de las palabras que de-see.

¿Que usted quiere saber qué se usa más, si pseudo o seudo? Pues al Corpes. O ¿cuándo comenzó a usarse el término «teléfono móvil»? Tendrá que mirar este programa, porque no se podrá estudiar la lengua ni la gramática sin él. Es «imprescindible».

Del mismo modo opina Bartol, quien además asegura que es el proyecto de la Real Academia de la Lengua «más importante» de cara al futuro y es una herramienta «potentísima e imprescindible». De hecho, explica que es el «más potente» que existe hoy en día de todos los que hay, como el modelo anglosajón, o el dedicado al francés, por el número de entradas que recoge. Sin embargo, se trata de «un trabajo sin fin», porque el idioma está en continua evolución.

Cada grupo tiene adjudicado un apartado muy concreto sobre un tema, por ejemplo «periódicos de

Cuba sobre salud», y a partir de ahí se pone a trabajar para incluir todas las palabras. Se calcula que se introducen unos 25 millones de términos por año y cada equipo se compromete a un millón al año, aunque hay quienes tienen mayores posibilidades, como la Universidad de Barcelona, y puede llegar a cinco millones de palabras.

Por tanto, detallan que este Corpes se elabora a partir de textos orales y escritos, por lo que también tienen que trabajar en formatos como programas de televisión u obras de teatro.

### EL BANCO SANTANDER, MECENAS DE LA INICIATIVA

Conscientes de la importancia que tiene este recurso, el Banco Santander decidió apoyar esta iniciativa en 2007 y recientemente, el director de la Real Academia Española (RAE), Darío Villanueva y el director general de Banco Santander y director de Santander Universidades, José Antonio Villasante, renovaron su vigencia hasta 2018. El objeto fundamental de este nuevo convenio, cuya aportación económica por parte de Banco Santander asciende a 700.000 euros, es continuar con las actividades de investigación vinculadas al desarrollo del Corpus del Español del Siglo XXI.

El banco, a través de Santander Universidades, y la RAE establecieron un nuevo marco de actuación para ambas instituciones, con el propósito de ampliar y reforzar su colaboración en el ámbito de la promoción y mejora de la actividad investigadora y académica sobre el léxico y la gramática del español, establecido por el anterior convenio. Este nuevo convenio, se enmarca dentro de la celebración del III Centenario de la RAE, a través de iniciativas en los ámbitos tecnológico e institucional de mutuo interés para un mejor servicio a los integrantes de la comunidad académica y a toda la sociedad.

El Corpes XXI, que supone la continuación del trabajo realizado con los corpus CREA y CORDE, es una iniciativa de la RAE y de la Asociación de Academias de la Lengua Española (Asale) en la que trabajan, además del equipo central, ocho externos a las Academias: seis de universidades españolas (Alcalá de Henares, Autónoma de Barcelona, León, Salamanca, Santiago de Compostela y Valencia), la Academia Argentina de Letras y la Fundación Comillas.