

Análisis y primeras conclusiones sobre las formas léxicas mayoritariamente buscadas en el *DLE 23*

Analysis and first conclusions about the lexical forms mostly searched in the *DLE*

PAZ BATTANER

Universitat Pompeu Fabra-Real Academia Española
paz.battaner@upf.edu

Resumen: El uso de los diccionarios, tanto monolingües como bilingües, siempre ha preocupado a los profesores y a los lexicógrafos. Conocer qué se consulta en un diccionario es asunto tanto de la adquisición y del desarrollo de la competencia en una lengua por parte de sus hablantes, como de saber en qué aspectos podrían mejorar estos recursos lingüísticos por parte de la lexicografía. En esta línea de investigación, los macrodatos que hoy ofrecen los buscadores de los diccionarios digitalizados consultados en la red ofrecen conocimiento objetivo y cuantificado del registro que dejan sus usuarios. En este artículo se analizan datos numéricos y léxicos de las consultas concretas recogidas en el mes de abril de 2022 en el *DLE* (RAE-ASALE).

Palabras clave: uso del diccionario; lexicografía; *DEL*; problemas de didáctica de la lengua.

Abstract: The use of dictionaries, both monolingual and bilingual, has always worried teachers and lexicographers. Knowing what is consulted in a dictionary is a matter of both the acquisition and the development of competence in a language by its speakers, and knowing in what aspects these linguistic resources could be improved by lexicography. In this line of research, the big data offered today by search engines of digitized dictionaries consulted on the Internet offer objective and quantified knowledge of the record left by their users. This article analyzes numerical and lexical data from the specific queries collected in the month of April 2022 in the *DLE* (RAE-ASALE).

Key words: use of the dictionary; lexicography; *DLE*; language teaching issues.

Fecha de presentación: 22/02/2023 *Fecha de aceptación:* 30/07/2023

1. INTRODUCCIÓN

Hace ya más de treinta años se empezaba con entusiasmo en la Universitat de Barcelona un programa de Doctorado en *Didáctica de la Lengua y de la Literatura*. El Departamento universitario de la División de Educación que pudo ponerlo en marcha respondía así al llamamiento de incorporación a la investigación de esta rama de didácticas aplicadas que es objeto de la formación de docentes en Primaria y en Secundaria.

PAZ BATTANER,

«Análisis y primeras conclusiones sobre las formas léxicas mayoritariamente buscadas en el *DLE 23*»,
Revista de Lexicografía, XXIX (2023), pp. 7-24

ISSN: 1134-4539, e-ISSN: 2603-667. DOI: <https://doi.org/10.17979/rlex.2023.29.1.9764>

El programa atrajo como alumnos a muchos profesores que ya ejercían, o bien en la primera o en la segunda etapa de la entonces EGB, o bien en los Institutos de Secundaria. Eran licenciados concienciados de su trabajo, con experiencia en las aulas. Los temas que planteaban en las sesiones y en los seminarios tenían su base en la observación directa tanto de su quehacer como de la eficacia de este en sus alumnos. Los profesores del programa tenían un recorrido algo más largo en la reflexión sobre esos temas. La investigación en estas materias había tenido hasta entonces escaso desarrollo y carecía de un cuerpo teórico mantenido en el tiempo; no estaba constituido formalmente en disciplina¹.

En una de las materias sobre el léxico y su adquisición que se implementaron en aquel programa, una doctoranda planteó su investigación sobre el conocimiento real del manejo de los diccionarios en alumnos de bachillerato, BUP en aquel momento. Pretendía conocer las búsquedas concretas que un grupo de estudiantes realizaba espontáneamente, sin consigna, en un diccionario general monolingüe. Como punto de partida esta cuestión se había analizado a partir de encuestas a estudiantes (Ibrahim y Zalesky 1989, Battenburg 1991) o se había observado en el comportamiento filmado de un número reducido de alumnos y sobre un número reducido de unidades léxicas contextualizadas, lo que en aquellos momentos se consideraba investigación empírica (Battaner 1994). La respuesta a encuestas podía estar teñida de subjetividad de los encuestados y la observación directa exigía un seguimiento difícil de describir.

Más ambiciosa, la doctoranda lo planteaba en un grupo de clase (30 alumnos mínimo), sin ofrecer un material concreto sobre el que trabajar, pues lo que le interesaba eran las búsquedas en sí: qué unidades buscaban espontáneamente adolescentes de 15 o 16 años en diccionarios generales monolingües; qué tipo de lemas; si discernían entre las acepciones; qué cuestiones lexicográficas les despertaban curiosidad para abrir un diccionario y buscar en él.

En las tutorías se planteó, como no podía ser por menos, la metodología que se utilizaría para la recogida de datos y el modo de recogerlos. Las dificultades metodológicas eran varias: a) por aquellos años no había todavía diccionarios informatizados, se tenía que trabajar con diccionarios en papel, moviendo para atrás y para adelante las páginas de los gruesos volúmenes que suelen tener estos recursos; b) había varios diccionarios de este tipo en el mercado con diferente alcance descriptivo, de manera que el grado de información y la satisfacción de lo encontrado presentaría gran variación; c) la observación tenía que trasladarse a la que tomaran los mismos sujetos en sus búsquedas, es decir, sería auto-observación, pues la investigadora no podía seguir a la vez esa actividad inquisitiva, ni fuera del aula, ni de todos los estudiantes del grupo; d) además, el hecho de ser los mismos usuarios los que recogieran lo buscado les ralentizaría mucho, rebajaría su interés y, más complicado aún, pervertiría el seguimiento real

¹ Ante el anuncio de una nueva reforma de la educación secundaria, convendría seguir la evolución de las didácticas específicas desde los años ochenta del pasado siglo en que la LOU y la LOGSE dieron en España salida a los continuos cambios efectuados en la normativa sobre educación primaria y secundaria posterior (Battaner 2002), para evaluar si ha habido un progreso real.

de su comportamiento con el diccionario. Hubo que descartar el trabajo tal como le interesaba a la doctoranda y hubo que plantearse otros objetivos de investigación más llevaderos. La cuestión era sin duda muy válida si hubiera podido obtener o recibir una respuesta, pero en aquel momento no había posibilidad de recoger metódicamente los datos que interesaban.

A día de hoy, el comportamiento de los usuarios en una página informatizada, de una plataforma o de una aplicación, puede ser objeto de análisis. Las estrategias con las que operan las grandes empresas comerciales de venta en línea se basan en el análisis de los rastros que dejan sus usuarios; es decir de los datos concretos recogidos para sugerir a sus clientes otras compras posibles. De manera similar los recursos lingüísticos ofrecidos en abierto por un diccionario informatizado pueden procurar también en la actualidad datos fehacientes de cada operación realizada en él y pueden ser objeto de recogida y de análisis posterior.

Este tipo de consultas telemáticas, posibles hoy en el *Diccionario de la Lengua Española* (DLE) no obtiene datos de los usuarios, como sí tienen las empresas comerciales. Las consultas son totalmente anónimas y descontextualizadas. Sin embargo, alcanzan un número suficiente como para poder avanzar hipótesis sobre la razón/explicación de esas búsquedas. La hipótesis general de partida para iniciar la búsqueda en un diccionario es la ignorancia del consultante sobre algún aspecto de cualquier forma léxica. La segunda hipótesis es que una cantidad de miles de consultas sobre una forma lingüística determinada en un tiempo determinado permite formular hipótesis sobre los condicionamientos de la búsqueda de esa forma o de ese tipo de formas por los usuarios del diccionario, aunque estos sean anónimos.

Si los datos recogidos sobre la consulta a una misma forma son superiores a cinco mil en un mes, alguien con experiencia sobre el acceso y uso de los diccionarios puede caracterizar las preguntas típicas que han llevado a un usuario desconocido a esa búsqueda. Lo deducido será algo parecido a lo que proponía aquella doctoranda de los años ochenta del pasado siglo que tuvo que abandonar su primera cuestión de investigación por no disponer de las indispensables ayudas tecnológicas. Esta condición puede satisfacerse hoy.

En la actualidad se tiene la oportunidad de obtener datos numéricos para poder enfrentar con alguna objetividad aquel tema de tesis: saber qué se busca en un diccionario abierto a la red mundial a través de las consultas recibidas; a partir de esas búsquedas y de la información que proporciona el diccionario se podrá caracterizar a grandes rasgos, por ejemplo, cuál es el manejo del diccionario de RAE-ASALE consultable en la red.

El diccionario del que tengo datos numéricos de las consultas es el *DLE, Diccionario de la Lengua Española* (RAE-ASALE), ofrecido electrónicamente en abierto al mundo de internet en su versión 23.5. Las consultas han sido recogidas desde navegadores y aplicaciones de dispositivos celulares; el lapso temporal en que se recogieron las consultas cuantificadas es el mes de abril de 2022; el número de las consultas que se manejan para este trabajo es de 50 millones, correspondientes a las que están geolocalizadas, provenientes de más de 200 países; unos, en los que la lengua española es la oficial, otros, en que solamente es objeto de consulta o de estudio. Por la serie de la

contabilidad mensual de estas consultas que ha seguido y sigue la RAE desde hace años, se puede promediar que hay un potencial de más de 10 millones de usuarios que lo consultan informáticamente.

2. ORGANIZACIÓN DE LOS DATOS

Los datos geolocalizados, limitados para este trabajo al mes de abril de 2022, ofrecen dos magnitudes: las formas lingüísticas reconocibles por el diccionario, no propiamente lemas lexicográficos, y el número de veces que se ha buscado esa forma en el *DLE* en línea. Por ejemplo, la forma *persona* recibió 17.269 consultas, la forma *altruismo* recibió 5.903.

Estos datos distan todavía de proporcionar todos los detalles que necesita un investigador para calibrar cualitativamente el uso del recurso lexicográfico. Al no estar las formas contextualizadas pueden haber sido buscadas desde diferentes cuestionamientos. Por ejemplo, tomamos la forma *capital* como si correspondiera al lema *capital* del *DLE*². Se desconoce: a) qué se pregunta al diccionario sobre esta forma, *capital*; b) si el consultante es adulto o no; c) si el español es su primera lengua o es lengua extranjera; d) si el interés de su búsqueda parte de su uso como sustantivo o bien como adjetivo; e) qué acepción de las que ofrece el diccionario del lema *capital* ha sido la buscada, bien entre las sustantivas (en masculino, cuatro acepciones, y en femenino, tres) o bien entre las tres adjetivas recogidas en el *DLE*, pues *capital* es una palabra polisémica y los datos que recoge la informática por el momento esconden la acepción que satisfaría al usuario

Ante la falta de estos datos cualitativos, este trabajo solo ofrece datos cuantitativos, muy lejos aún de poder extraer, por el momento, cualquier conclusión lingüística y sociolingüística. Sin embargo, sí puede dibujarse a grandes trazos el panorama de los usuarios del diccionario y se pueden ofrecer características generales sobre los aspectos que despiertan interés mayoritario entre la rica información que guardan los diccionarios.

Con estos presupuestos condicionantes, revisaré en este trabajo los datos de ese limitado periodo de tiempo y aventuraré posibles causas del número tan alto de consultas que reciben ciertas formas. Los comentarios cualitativos que adelantaré han de ser entendidos como sugerencias de los motivos de búsqueda; han de interpretarse exclusivamente como primeras hipótesis sobre la causa que lleva a muchos usuarios a haber hecho su consulta. Generaré estas hipótesis a partir de mi experiencia como profesora de lengua española conocedora del su vocabulario desde el punto de vista lexicográfico, así como la información sobre ortografía, gramática, educación y cultura que recoge el *Diccionario de la Lengua Española*. El análisis presentado y las conclusiones a las que llegue serán solamente una primera aproximación basada sobre gran número de datos reales cuantificados obtenidos de la consulta a un solo diccionario, el *DLE*.

Primeramente las hipótesis se formularán sobre el porqué de la búsqueda, apoyadas en la cantidad de consultas recibidas en ciertas formas concretas; no se puede interpretar que todas las consultas realmente busquen lo hipotetizado, pues algunas podrían

² *capital* ocupa el rango 287 por haber recibido 7.279 número de consultas entre las 95.000 formas consultadas ordenadas decrecientemente en el mes de abril 2022.

responder a otras cuestiones; solo tienen valor como apreciación global y primer acercamiento a partir de datos numéricos reales.

Otro caso, si la búsqueda de las formas *catarsis* o *epifanía* se contabilizan en un tramo entre 8000 y 9000 consultas en ese mes de abril de 2022, se sugiere que la mayoría de las búsquedas sobre esas formas responde a una duda propiamente de su significado como palabra reconocible y, sin embargo, que la forma *no*, en el mismo tramo de número de consultas, responde a una duda sobre su categoría gramatical, adverbio de negación, cuestión puramente escolar en los tradicionales análisis gramaticales.

Estas dos primeras hipótesis posibles se apoyan en el comportamiento de los hablantes en su relación con la lengua escrita y con el currículo de la educación reglada, comportamiento conocido a través de la experiencia profesional.

Otra tercera hipótesis de partida manejada es que los usuarios son estudiantes. Dado que la mitad de las consultas al *DLE* se hace desde España 29,37%, México 11,03% y algún otro país de habla española, que puede rondar sobre el 8%³, los estudiantes de la enseñanza obligatoria de estos países reciben instrucción de la consulta de un diccionario monolingüe y esto permite reforzar, primero, que el número tan elevado de visitas que reciben algunas formas tenga que atribuirse a jóvenes estudiantes en formación y, segundo, que estas se realizan frecuentemente cuando responden a ejercicios didácticos de su asignatura de lengua española⁴. La otra mitad de consultas se realiza desde el resto del mundo y representa un 52%, pero en este trabajo no se han tratado por separado; por lo que esta tercera hipótesis no incluye el que los consultantes solo sean de países con lengua española mayoritaria.

Para una redacción más operativa el volumen manejado de consultas de 50 millones de datos se distribuye en veinticinco tramos, correspondientes a consultas geolocalizadas por la web y desde aplicaciones para dispositivos móviles, como se ha dicho. Los primeros tramos corresponden a un número limitado de formas que reciben un número muy alto de consultas. Vienen a continuación los tramos constituidos por formas de consultas decrecientes, cuyas búsquedas parecen relativamente fáciles de caracterizar por el tipo y por el número todavía reducido de formas que cada tramo recoge. A partir del tramo 16 se establecen los tramos de cien en cien consultas y el número de formas que constituye cada tramo crece espectacularmente, lo que rebajaría el valor de las hipótesis formuladas en un análisis de búsquedas sin contexto. Quedan pues fuera de este trabajo los últimos tramos, más variados y numerosos en las formas que recogen (Tabla 1, *infra*).

Todas las formas contabilizadas están presentes en el diccionario académico pues, si al teclear lo buscado la grafía no coincide, no se contabilizan como tales consultas, sino que pasan al conjunto de formas que no aparecen en el *DLE*, el cual suministra otro tipo de información aquí no considerada. Las 27.820 formas que comprende el

³ En las estadísticas correspondientes a abril de 2022, las consultas desde Estados Unidos llegan a 6,96%, cifra, que ronda a la de otros países americanos, lo que indica la fuerte presencia del español en ese país.

⁴ El seguimiento hecho con encuestas por las investigadoras Santamaría Pérez y Azorín Fernández (2017) concluye que el diccionario más consultado por un grupo de estudiantes limitado y conocido, es también el *DLE* en línea.

tramo 25 recibieron solo una consulta. Corresponden muchas de ellas a femeninos de adjetivos o sustantivos con flexión de género⁵, a nombres propios que aparecen en las locuciones o definiciones, a siglas y a otros casos particulares.

TRAMO	NÚMERO DE CONSULTAS	NÚMERO DE FORMAS
1	Más de 50.000	5
2	Entre 30.000 y 50.000	4
3	Entre 20.000 y 30.000	19
4	Entre 15.000 y 20.000	16
5	Entre 12.500 y 15.000	22
6	Entre 10.000 y 12.500	48
7	Entre 9 000 y 10.000	41
8	Entre 8 000 y 9 000	65
9	Entre 7 000 y 8 000	103
10	Entre 6 000 y 7 000	167
11	Entre 5 000 y 6 000	318
12	Entre 4 000 y 5 000	581
13	Entre 3 000 y 4 000	1 194
14	Entre 2 000 y 3 000	2 609
15	Entre 1 000 y 2 000	7 119
16	Entre 900 y 1 000	1 398
17	Entre 800 y 900	1 661
18	Entre 700 y 800	1 989
19	Entre 600 y 700	2 570
20	Entre 500 y 600	3 224
21	Entre 400 y 500	4 413
22	Entre 300 y 400	6 406
23	Entre 200 y 300	10.347
24	Entre 100 y 200	20.697
25	Con menos de 100	27.820

TABLA 1. *Distribución de los 50 millones de datos sobre consultas contabilizadas en el DLE (abril 2020)*

3. FORMAS RECOGIDAS EN LOS ONCE PRIMEROS TRAMOS

Se intentará caracterizar el contenido léxico que componen los primeros once tramos junto a las causas que pudieron originar su búsqueda. Los lemas a que pueden remitir las formas listadas del décimo segundo al vigésimo quinto tramo son tan numerosas que quedan en una zona oscura para avanzar hipótesis sobre los posibles motivos de sus búsquedas.

⁵ Cuando alguien consulta una forma femenina, el *DLE* ofrece la forma del lema, generalmente en masculino, bajo el que aparece todo el artículo lexicográfico. Ello hace que muchas consultas en femenino se tengan contabilizadas una sola vez.

Tramo 1, compuesto de 5 formas que contabilizan más de 50 mil consultas cada una: Tienen esta consideración exclusivamente cinco formas, *diccionario*, *consulta*, *palabra*, *de* y *que*. El lema *diccionario* se contabiliza por defecto al abrir cualquier búsqueda de *Diccionario de la Lengua Española* en línea. Le siguen *consulta* y *palabra*, que responden a ejemplos de palabras fácilmente ocurrentes cuando no se sabe qué o cómo buscar en la página. Empiezan a tener entidad, *de* y *que*, palabras puramente gramaticales, que coinciden con las de mayor frecuencia en lengua española. La preposición *de* y la conjunción, pronombre o adjetivo relativo *que*, pueden responder a consultas en que estas partículas se escribían o se copiaban junto a la palabra buscada para obtener la respuesta a una pregunta escolar sobre su categoría gramatical. Son muy poco informativas a pesar del alto número de consultas que reciben.

Tramo 2, compuesto de 4 formas que contabilizan entre 30 y 50 mil consultas cada una: *haber*, *hacer*, *follar*, *ir*, son formas que parecen responder a preguntas ortográficas o morfológicas pues tres pertenecen a los verbos más usuales e irregulares. *Follar* puede responder a la curiosidad, generalmente infantil, de buscar cómo se definen o explican las palabras que tienen relación con el sexo; ambas se mantienen en los subconjuntos o tramos siguientes.

Tramo 3, compuesto de 19 formas que contabilizan entre 20 y 30 mil consultas cada una: *a*, *amor*, *ciencia*, *cultura*, *dar*, *decir*, *definir*, *el*, *en*, *estar*, *haz*, *o*, *poder*, *pornografía*, *procrastinar*, *ser*, *significado*, *tener*, *ver*. Tres sustantivos, *amor*, *ciencia*, *cultura*, y un verbo, *procrastinar*, podrían responder a objeto de búsqueda de su definición. El resto responde a posibles dudas ortográficas o gramaticales (*a*, *el*, *en*, *o*), a la conjugación irregular de verbos, *dar*, *decir*, *estar*, *ir*, *poder*, *ser*, *tener*, *ver*, y a la forma *haz*, sea que pertenezca a un sustantivo *haz*, o a la forma de imperativo del verbo *hacer*⁶. Un aspecto de contenido sexual, *pornografía*; una novedad en la lengua por ser un cultismo moderno, *procrastinar*; y finalmente tres voces cuyo contenido podría encontrarse también en un repertorio enciclopédico, *amor*, *ciencia*, *cultura*. No hay adjetivos aun en este tramo de búsquedas.

Tramo 4, compuesto de 16 formas que contabilizan entre 15 y 20 mil consultas cada una: Si del conjunto de formas de este tramo, se explican *puto* y *sodomía* por curiosidad sexual; *hallar* como único verbo del tramo por dudas ortográficas; y *aun*, *como*, *este*, *para*, *qué*, *se*, *e*, *y*, como voces cuyo análisis morfológico puede ser una pregunta didáctica; se empiezan a encontrar cinco típicas búsquedas de significado en el diccionario: dos adjetivos esdrújulos, *cínico* y *efímero*; un neologismo reciente, *resiliencia*; una palabra expresiva, *trotamundos*; y otra, por el contrario, de significado polisémico y difícil, *persona*, cuya consulta puede responder a la ambigüedad con la que suele ser usada. Habría pues en este tramo 4.º cinco palabras objeto de búsqueda puramente lexicográfica: *cínico*, *efímero*, *persona*, *resiliencia* y *trotamundos*.

Tramo 5, compuesto de 22 formas que contabilizan entre 12.500 y 15 mil consultas cada una: En este tramo se encuentran consultas explicables por su irregulari-

⁶ El diccionario reconoce lemas y también formas flexivas, por lo que la forma *haz* del tramo 2, por ejemplo, puede haber sido buscado como sustantivo o como forma de imperativo del verbo *hacer*, que en las zonas de seseo puede presentar dudas ortográficas. Me inclino por esta segunda opción.

dad morfológica y ortográfica, *hallar*, *querer*, *saber*, y la forma del verbo *ir*, *vaya*; por el análisis de su categoría gramatical, *todo*, *más* y la preposición *por*. Consultas semánticas propias y espontáneas a un diccionario serían las de lemas de significado muy amplio, polisémicas y siempre difíciles de definir, lo que puede despertar curiosidad en un usuario: *acción*, *bien*, *bizarro*, *caso*, *estado*, *fe*, *historia*, *humildad*, *indulgente*, *respeto*, *unir* y *vehemente*. Destacamos entre ellas los adjetivos, *bizarro* (que lleva tiempo apareciendo entre las consultas al *DLE* por su relativamente extendido uso de origen galicista), *indulgente* y *vehemente*; las voces semitérminos⁷ *acción*, *bien* (altamente polisémica), *estado* y los sustantivos *humildad*, *historia*, *fe* y *respeto*; solo *historia* se podría relacionar a lo encontrado hasta ahora con *ciencia* y *cultura*; los otros tres responden a curiosidad por vivencias personales de concienciación interior. Quizá los términos representantes en este tramo de búsquedas morbosas sean *petardo* y el verbo *coger*, que por su valor en América puede haber despertado curiosidad sexual, aunque su *g* y su conjugación puedan plantear también dudas ortográficas.

Tramo 6, compuesto de 48 formas que contabilizan entre 10 mil y 12.500 mil consultas cada una: Los lemas de este tramo se caracterizan por: a) las dificultades ortográficas y morfológicas que pueden a la vez plantear *hola*, *venir*, *echar*; b) las categorías gramaticales de formas como *algo*, *cómo*, *con*, *del*, *si* como conjunción, para resolver ejercicios escolares de análisis; c) temas de interrogación escolar y sociológica, como *género* y *lengua*. *Puto* y *hostia* responden a lo que venimos caracterizando como curiosidad morbosa infantil; en este caso por una palabra, *hostia*, que responde a una blasfemia devaluada hoy; y un adjetivo con más de 16.000 consultas, *puto*⁸. Fuera de estas, se encuentran en la lista unas 30 formas interesantes semánticamente para un diccionario. Se destacan, como va a ser habitual en los siguientes tramos establecidos: adjetivos, *derecho* (que posiblemente responda mejor a una búsqueda de sustantivo, que no de adjetivo), *ético*, *hegemónico*, *inefable*, *intrínseco*, *moral*, y *vasto* (que puede cuestionarse ortográficamente con *basto*⁹); verbos, *comer*, *discernir*, *inferir*, *permitir*, *pertenecer*, *preceder*, *prever*, *redimir*, *satisfacer*, *concebir*, todos, salvo *comer*, quizá respondan a dudas sobre su significado; sobresalen entre ellos los de conocimiento, *discernir*, *prever*, *inferir* y *concebir*, junto a los que responden a análisis introspectivos: *redimir*, *permitir*, *satisfacer*. Quedan en la lista sustantivos, *arte*, *democracia*, *distopía*, *filosofía*, *justicia*, *libertad*, interesantes en cuanto a su valor conceptual, y *conciencia*, *diligencia*, *emoción*, *empatía*, *sesgo*, *valor*, en cuanto a posibles interrogaciones sobre la autoconciencia de los jóvenes. Como eco posiblemente de las continuas tomas de posición del feminismo exteriorizado en formas lingüísticas, es curiosa la coincidencia entre las formas masculina y femenina de *presidente* y *presidenta* en el mismo rango de número de consultas del diccionario en el mes de abril de 2022. La forma femenina

⁷ Tomamos la denominación de *semitérminos* del trabajo doctoral de Victoria Boschioli. Unidades léxicas no propiamente terminológicas de uso habitual y frecuente en textos académicos de muchas disciplinas, en distintos grados de especialidad, que ayudan a desempeñar funciones textuales fundamentales en la prosa expositiva (Boschioli, TFM, UNED 2014). En cada ámbito de uso su referente puede ser muy variado, mientras que su significado conceptual se mantiene.

⁸ Su femenino, *puta*, como primera búsqueda obtiene solamente 46 consultas.

⁹ El adjetivo *basto* aparece en el tramo 8 con 8211 consultas.

presidenta está en el español documentada, según el CORDE de la RAE, desde 1448 en la traducción anónima al castellano del *Llibre de les dones* del autor catalán Francesc Eiximenis (1330-1409).

Tramo 7, compuesto de 41 formas que contabilizan entre 9 mil y 10 mil consultas cada una: Solo las formas *sí* y *sino* pueden responder a la solución de un análisis morfológico. Las consultas gramaticales más finas o exigentes, como serían usos del *sí*, pronombre personal o del adverbio afirmativo, y de la conjunción adversativa *sino* (o quizá sustantivo con graves discusiones teológicas tras él) no se pueden resolver con los datos manejados; los sustantivos elementales *agua*, *gato*, *mujer* y *perro* responderían a consultas tópicas o quizá metódicas para adquirir el manejo del diccionario electrónico. En cuanto a cuestiones del *syllabus* de lengua podrían colocarse *comunicación*, *etimología* y, decididamente, *lenguaje*, que seguiría a la consulta *lengua*, que se ha visto en la lista del tramo anterior. Solo hay dos lemas, el del adjetivo *pendejo* y el del verbo *sodomizar*, que respondan a la búsqueda morbosa e infantil de palabras cargadas de tabús. Entre los adjetivos que se va viendo que tienen entidad significativa en las búsquedas, se encuentran en este tramo: *ambiguo*, *asertivo*, *baladí*, *bellaco*, *inherente*, *mezquino*, *vicario*. Los verbos *elegir*, *exacerbar*, *exhortar*, *inhibir*, pueden ofrecer dudas ortográficas, pero también semánticas; *deber*, *conocer*, *leer*, *realizar* y *remitir*, no parecen ofrecer cuestiones propiamente semánticas pero sí flexivas. Entre los sustantivos de este tramo de consultas, *carácter*, *efecto*, *examen*, *identidad*, *orden* y *tiempo*, parecen semitérminos; se encuentran también *ética* y *soberbia*, coincidentes con otros que se van encontrando de reflexión moral. Y llaman la atención en este estadio el coloquial *curro*, el exótico y negativo *sátrapa*, con más usos de adjetivo que de sustantivo, *dizque* (que habría que reconsiderar en el *DLE*) e *idiosincrasia*, característica semánticamente afín a todo lo que tiene que ver con las unidades léxicas (Hockett *dixit*).

Tramo 8, compuesto de 65 formas que contabilizan entre 8 mil y 9 mil consultas cada una: Se recogen en este tramo formas reconocibles como sustantivos que pertenecen al vocabulario elemental, *familia*, *banco*, *mano*, *pie*; junto a otros que responden a cuestiones escolares: *vocabulario*, *libro*, *ejercicio*, *nombre*; entre los semitérminos encontrados en los textos didácticos se podrían recoger *parte*, *período* y *punto*. Algunas búsquedas pueden explicarse por conocer la categoría gramatical a la que se asignan: *cuando*, *no*, *solo*, *una*. Formas que despiertan curiosidad morbosa por su definición pueden ser: *gilipollas*, *cabrón* y *verga*. Entre los verbos de este tramo de consultas pueden despertar dudas ortográficas *botar*, *caber*, *rallar*, *rebelar*, *revelar*, *vacilar*, *yacer*; dudas de morfología flexiva, *asir*; la forma *haya*, es posible que responda al verbo *haber* y no al árbol que recibe ese nombre, también en el tramo 4 se encontró la forma *halla*; con explicación parecida se puede recoger *hecho*, como derivado de *hacer*. Pueden responder a dudas semánticas junto a ortográficas, *aprehender*, *conllevar*, *corregir*, *enajenar*, *proveer*, *recoger*, *reír*, *rielar*. Los adjetivos, salvo *gilipollas* y *cabrón* que responden a otro tipo de búsquedas, pueden ser consultas propias por su significado: *banal*, *basto*, *común*, *consciente*, *inocuo* y *patriotero*, aunque alguno responde también a dudas ortográficas como la pareja *banal* y *basto*. Se encuentran bastantes sustantivos con significado cuestionable, unos por los conceptos que se manejan en el currículo escolar, *actividad*, *analogía*, *cualidad*, *relación*, *sistema*, *sociedad*; otros

por su mismo significado: *catarsis, eficacia, elocuencia, epifanía, feminismo, ideología, ironía, responsabilidad, tolerancia y utopía*; son llamativos *masoquismo y fetichismo* como consulta sobre sexualidad en un aspecto más especializado que en los anteriores tramos comentados.

Tramo 9, compuesto de 103 formas que contabilizan entre 7 mil y 8 mil consultas cada una: Se encuentran entre las formas de este tramo pocas formas que puedan asignarse a palabras gramaticales, explicables como respuestas de análisis escolares; son la contracción *al*, el indefinido *otro* y una onomatopeya, *muac*, introducida en el *DLE* en la edición en línea 23.5 (diciembre 2021), que en la versión en papel del *DLE* no figura. Voces que se vienen atribuyendo a la curiosidad sexual de los jóvenes serían *encular, coño, culo* y en otro nivel de mayor ignorancia semántica, *lascivia*. Como duda ortográfica se encuentra la forma *valla*, que no por su significado sustantivo, sino como alternancia de la forma del verbo *ir/vaya*, parece obtener este alto número de consultas. En cuanto a voces consideradas palabrotas, aparece *ostia*, sin hache; con hache se encuentra en el tramo 5.

Los lemas adjetivos de este tramo muestran bien el interés por lo que resuelve el diccionario; entre ellos, *afable, fatuo, hereje, indolente, necio, plausible, pusilánime, sobrio, vano*, seleccionan personas o acciones humanas; otros no son restrictivos en este aspecto, *recíproco, pragmático, objetivo, ignoto, político, subjetivo; etéreo* selecciona sustantivos espaciales o es utilizado poéticamente; *capital*, como *político*, puede ser buscado, como adjetivos tanto como sustantivo. Parece más lógico pensar que las búsquedas mayoritarias para esos lemas son usos adjetivos, pero los datos de las estadísticas manejadas no hilan fino en estas cuestiones.

Los verbos son abundantes y bien variados en este número de consultas: *abolir, abrumar, andar, buscar, cesar, comprender, condescender, conjuntar, consolidar, continuar, denotar, determinar, devenir, eludir, erigir, errar, establecer, evocar, eximir, generar, incidir, inmolar, llevar, oír, percibir, recoger, regir, supeditar, tergiversar, traer, transmitir, ultrajar, votar*. Entre los anteriores varios tienen significados negativos, algunos presentan problemas ortográficos y uno, claramente defectivo, *abolir*, es el ejemplo de este tipo de verbos en todas las gramáticas.

Junto al largo número de verbos, los sustantivos son también variados; desde los semitérminos *causa, forma, guion, imagen, ley, trabajo, vez*, o los elementales, *cabeza, día, masa*; significan capacidades y sentimientos humanos *actitud, análisis, aversión, conocimiento, decisión, dignidad, gracia, hipocresía, melancolía, orgullo, prejuicio, virtud*, junto al neologismo *sororidad*. *Acervo, anatema, argumento, contexto, dicotomía, eficiencia, estereotipo y vicisitud*, son conceptos analíticos que los estudiantes han de poder reconocer en un texto expositivo; junto a estos, *mito, quimera, oración, religión* tocan temas trascendentes. Extraña encontrar en una lista como la anterior el latinismo *statu quo*, el sustantivo coloquial *zocato* para ‘zurdo’, y el término de botánica *hifa*.

Tramo 10, compuesto de 167 formas que contabilizan entre 6 mil y 7 mil consultas cada una: En este tramo, ya muy numeroso en formas para seguir sugiriendo

razones de su búsqueda, se encuentran 61 verbos; 28 adjetivos; 70 sustantivos; y otras categorías gramaticales.

Por conocer sus categorías gramaticales se explican las búsquedas de *alguien, así, donde, quien, si, su* y los tratamientos, *don* y *señor*. El caso de la presencia de la forma *era* en este tramo puede explicarse mejor como respuesta a una pregunta de análisis gramatical (Pto. Imperfecto del verbo *ser*), que no por el sustantivo *era*, ‘período de tiempo’ o ‘espacio despejado’.

Entre los setenta sustantivos de este tramo se pueden recoger abstractos que tienen que ver con sentimientos, actitudes, capacidades de los seres humanos: *amistad, aptitud, astenia, circunspección, consciencia, desdén, ímpetu, inteligencia, osadía, paciencia, pasión, razón, sarcasmo, vanidad*. Sustantivos abstractos puramente conceptuales: *apología, autonomía, competencia, comunidad, consecuencia, demagogia, disposición, eficiencia, epistemología, equidad, ignominia, incidencia, monarquía, paz, política, tecnología*. Lo que venimos llamando semitérminos: *acepción, canon, concepto, consejo, disciplina, educación, estatus, hipótesis, ontología, origen, precepto, principio*; otros sustantivos relacionados claramente con la reflexión escolar sobre lengua, como *letra, léxico, sintaxis*; también voces generales polisémicas y cuyo uso es variado y de alta frecuencia, *leyenda, lugar, magnitud, número, origen, umbral*. Estos se pueden prolongar en *aguacate, carta, cónyuge, hijo, hito, llama, mono* (que puede ser un adjetivo junto a *pijo*), *naturaleza, planeta, planta, vasera, vela, vida, video; carajo* (que responde a las palabras tabúes); *salbanda* (término técnico de mineralogía).

Los adjetivos de este tramo continúan en la línea de los anteriores; señalan al diccionario y a los profesores el trabajo que hay que hacer para facilitar la comprensión y la producción de esta categoría léxica, siempre difícil. Los adjetivos que seleccionan principalmente referentes humanos son *austero, diligente, doctor, erudito, nómada, promiscuo, sabio, taciturno, tenaz*, curiosa la coincidencia de *erudito* con *sabio* en el número de consultas; los que pueden combinarse con sustantivos más restrictivos serían *apócrifo, autócrata, cabal, exento, gratis, icónico, inexorable, inmarcesible, laico, literal, prolijo, relativo, sórdido, superfluo, sutil, trivial, verosímil*; un adjetivo que es pluriempleado y polisémico, *medio*, admitiría clasificarlo como semitérmino. *Pijo* responde a ambientes juveniles españoles propiamente.

Los verbos que reciben este número de consultas presentan formas que despiertan dudas ortográficas o de su flexión irregular: *abrir, absorber, acaecer, aprehender, callar, ceñir, coadyuvar, consentir, constituir, creer, diferir, dirigir, disponer, embelesar, encontrar, excitar, expedir, grabar y gravar, halagar, huir, incluir, ir, jugar, persuadir, prescribir, producir, promover, pujar, raer, rayar, reivindicar, revocar, salir, seguir, servir, soslayar, sucumbir, surgir, transcender, velar*; puede que algunos hayan sido consultados por la diptongación en alguna de sus formas. Extraña encontrar en este alto tramo de consultas, pues son regulares, los verbos *articular, ceder, discurrir, emitir, llegar, parecer, pasar, pesar, prescindir, representar, suscitar, tomar*. Y los que han podido levantar dudas sobre su significado parecen *acuñar, aunar, derogar, emanar, mitigar, ratificar, refutar, topar*, regulares de la primera conjugación pero de sig-

nificado preciso, el último, *topar*, pendiente de figurar en el *DLE* con sus últimos usos de ‘parar’, ‘detener’.

Tramo 11, compuesto de 318 formas que contabilizan entre 5 mil y 6 mil consultas cada una: Será el último tramo analizado por su creciente número de formas, lo cual evidencia que estos análisis, basados en hipótesis que no se pueden confirmar por su descontextualización, resulten arriesgados y sin ningún valor empírico, como se ha ido explicando.

Los verbos que siguen pueden tener problemas de ortografía y de conjugación irregular, *afligir, exigir, fungir, transigir*, por ejemplo, o bien *aducir, deducir, inducir*; los terminados en *-bir, exhibir, recibir, suscribir*, frente a *vivir*; o tienen alguna *h, hacinar, hojear, hostigar, rehusar*, también el anterior *exhibir*, o tienen alguna sílaba que puede provocar dudas ortográficas, como *abocar, abordar y avocar, anegar, atisbar, avalar, azorar, bordar, enviar, existir, exonerar, instar, observar, obtener, oler*; otros de este tipo son *converger, dejar, devengar, obtener, proteger, recabar, revisar, subyacer, subyugar, trabajar, vislumbrar*. Verbos de significado muy conocido serían casos como *amar, aplicar, cantar, consultar, correr, crear, entender, estudiar, rezar, interpretar, manifestar, quedar, significar*, cuyo número de consultas en el diccionario se podría explicar por su polisemia, que exige repasar las distintas acepciones; podría ser también *aprender*, pero quizá su búsqueda responde a dudas con *aprehender*, verbo que se encontraba entre los del anterior tramo de consultas; algo parecido ocurre con la presencia de *estimar* y *amar* en este bloque. Hay dos lemas *placer* y *regular*, cuya simple presencia no deja determinar a qué búsquedas responden, por sus otras posibles categorías, sustantivo o verbo, y en el caso de *regular* como adjetivo también.

Restan de recoger los verbos que han podido levantar dudas sobre su significado. Entre estos encontramos *afectar, aludir, anonadar, conformar, consensuar, considerar, derogar, desarrollar, desdeñar, dilucidar, discriminar, disociar, escindir, expresar, identificar, implementar, implicar, indagar, jactar, juzgar, lomear, ostentar, referir, repentizar, segregare, valorar*. Estos lemas verbales y los anteriores, entre los que pueden presentar dudas ortográficas o morfológicas, presentan un bonito programa de posibles ejercicios didácticos sobre análisis de significado, adecuación con sus argumentos y, en resumen, para potenciar la precisión léxica.

Entre los 71 adjetivos que presenta este tramo, el adjetivo *mozo*, con un uso mayoritario como sustantivo, permite añadirle los adjetivos que se restringen a personas humanas: *abnegado, amante, arrogante, desconcertado, dotado, sagaz*. Se recogen otros adjetivos que, descontextualizados pueden tener usos sustantivos, como *adjetivo, animal, chulo, médico* y *tópico*. También se encuentran en este tramo adjetivos comunes de alta frecuencia, como *bueno*, cuyo significado sería un bonito ejercicio para constatar las diferentes facetas semánticas con las que se puede interpretar¹⁰, y *culto, estúpi-*

¹⁰ El adjetivo *bueno* (que conceptualmente recibe la siguiente definición en el *DLE*: «De valor positivo, acorde con las cualidades que cabe atribuirle por su naturaleza o destino») muestra sentidos diferentes si su argumento es un sustantivo de alimentación (‘apetecible’), un regalo (‘útil, adecuado’), una bicicleta (‘bien hecha’), una respuesta de examen (‘acertada’), la conducta de una persona (‘recta’), el clima (‘agradable, templado’), etc.

do, frívolo, hostil, lúgubre, siniestro; así como los valores adverbiales y adjetivos de la forma *harto*. A lo largo de materias curriculares habrán salido los adjetivos *antecedente, autodidacta, determinado, empírico, endémico, estoico, homónimo, ortodoxo, retórico, vertical*. Más interesantes desde el punto de vista de su significado y de los sustantivos con los que pueden combinarse son los de la larga lista siguiente, que representan un incentivador proyecto de reflexión didáctica en la escuela secundaria: *abyecto, anodino, arbitrario, asequible, burdo, coercitivo, cognitivo, congruente, conspicuo, diáfano, escatológico, espurio, excelso, explícito, fútil, gravoso, holístico, idóneo, impetuoso, incipiente, inicuo, inmanente, inteligible, lacónico, oneroso, ordinario, paulatino, peculiar, perenne, pertinente, putativo, solemne, sostenible, sublime, susceptible, tácito, torvo, transversal, versátil, vulnerable*; los adjetivos *sostenible* y *vulnerable* han alcanzado hoy una frecuencia notable y muy posiblemente reciben tantas consultas por su presencia en la comunicación. *Sostenible* con un aura positiva sin necesidad de completar su significado con los argumentos que necesita: de **qué** se puede decir *sostenible* y **en qué** se es *sostenible*: *uso, utilización, producción, o en el tiempo, en un lugar, en el ámbito de, en lo económico, etc.*; de qué valores propiamente sociales se dice *vulnerable*: *grupos, sectores, especies vulnerables o población, colectivo vulnerable, o bien, por ser mayores de edad, por padecer enfermedades, por carecer de recursos económicos, por falta de integración en grupos sociales, etc.*

También se recogen en este tramo dudas sobre categoría gramatical: *alguno, aparte, aquel, asimismo, cada, haiga, hasta, lo, mua, mucho, pero, sobre, tanto, te, través, uno*. Lista en la que se sigue encontrando dudas en pronombres personales e indefinidos, preposiciones, conjunciones y elementos de locuciones adverbiales, como *través, a través de*, que se han recogido en los rangos analizados anteriores; o la onomatopeya *mua* (5070)¹¹, frente al *muac* (7050) del tramo anterior.

Los sustantivos, cercanos a 146, presentan un amplio abanico sobre la curiosidad de los usuarios, siempre pensando en un joven estudiante de secundaria obligatoria o postobligatoria. No dejan de sorprender algunas búsquedas que se han hecho tópicas en las consultas del *DLE* como son *almóndiga* y *haiga*, formas de las que el diccionario advierte su condición de vulgarismo por *albóndiga* y de lo que aclara ser denominación de un vehículo ostentoso en la España de la postguerra. Menos explicables son las más de 5000 consultas al alcaloide tóxico *estricnina*, al cultismo *nefelibata*, que Antonio Machado utiliza como ripio en el *Cancionero apócrifo*; el extraño adjetivo *paradojo* por *paradójico*, mucho más usual (después aparece en femenino como sustantivo, *paradoja*) y el cubanismo *tallullo*, con su significado físico, ‘masa de harina de maíz’ o conceptual ‘embrollo’. Estos serían los casos señeros que llaman la atención. Luego se pueden hacer, rehacer, intentar, diversas agrupaciones para presentar los sustantivos siguientes: a) Sustantivos del vocabulario nuclear: *alma, casa, cola, cuerpo, dios, estrella, fuente, hoja, juego, luz, niño, ojo, pico, tierra, varón*, parecido al encontrado en otros tramos en que aparecían extremidades del cuerpo humano como *mano* o *cabeza*. Lllaman la atención ortográfica los casos de *hoja* y *ojo* en este grupo. b) A continuación vendría otro grupo con referentes físicos también, pero de un nivel de frecuencia inferior:

¹¹ Entre paréntesis el número de consultas recibido.

*arroba, ayo*¹², *bullá, carné, caso, celaje, cólera, enfermedad, huso, merluzo, paso, raya, vástago*, en donde se encuentran palabras que pueden presentar dudas semánticas, ortográficas u otras que fácilmente podrían pasar a estar encuadradas en apartados temáticos, como *caso*, un semitérmino claro, o la pareja *cólera* y *enfermedad*. c) El caso de *merluzo*, junto al adjetivo *sexí* y el verbo *joder* y, en un nivel algo más refinado, *concupiscencia*, serían los casos de búsquedas con su pequeño componente morboso.

Hay un grupo importante de sustantivos que se podría presentar como constituido por voces que aparecen siempre en la enseñanza. Tienen que ver más con las Humanidades que con las Ciencias, en general, pero son palabras que, o bien se usan en las disciplinas tradicionales, o se han de memorizar como conceptos de lo que se estudia: *acento, afrenta, ambiente, artículo, blasfemia, carpiano, código, conflicto, cuenta, déspota, dogma, economía, empresa, fenómeno, fin, función, glosario, guerra, héroe, igualdad, información, línea, longitud, metáfora, modo, morfología, oxímoron, problema, proposición, recurso, revolución, sentido, teoría*. Entre ellas se pueden encontrar conceptos prototípicos de alguna disciplina (*acento, artículo, código, dogma, economía, información, teoría*); otros que hemos llamado semitérminos (*ámbito, calidad, capacidad, condición, conjetura, esencia, excepción, experiencia, fenómeno, fin, función, igualdad, instancia, modo, norma, problema, proposición, propósito, recurso, sentido, síntesis*); términos más especializados (*alegoría, antonomasia, carpiano, escatología, eufemismo, glosario, holismo, metáfora, morfología, némesis, oxímoron*); y otras palabras no solo marcadas por la escuela, que se han colocado aquí por su relación con algunos temas de derechos humanos (*racismo, machismo, misoginia*), políticos (*burocracia, dictadura*) y combativos (*afrenta, empresa, conflicto, héroe, nación, patria, república*) y un tipo de persona, *déspota*.

Se encuentran sustantivos relacionados con actitudes y sentimientos humanos: *altruismo, armonía, azar, confianza, cinismo, desidia, egoísmo, esperanza, euforia, felicidad, honor, pena, prudencia, sosiego, vigor, zozobra, con celo y merced* (estos últimos muy probablemente buscados en plural, *celos* y *mercedes*). Junto a estos un gran número de sustantivos abstractos, cuyo significado explica muy probablemente el número de sus búsquedas: *albedrío, coherencia, compromiso, concienciación, concupiscencia, constitución, contingencia, corrupción, coyuntura, crisis, desarrollo, detrimento, disrupción, diversidad, estrategia, expectativo, interacción, juicio, menester, misantropía, necesidad, paradoja, perjuicio, probidad, propiciación, repertorio, reticencia, salud, serendipia, sinergia, violencia*.

Lo mucho y variado encontrado en este tramo 11 con menor número de consultas y con mayor número de voces cuya categoría gramatical puede ser dudosa, variedad en verbos con flexión irregular o dificultad ortográfica de algún tipo y disminución del número de voces buscadas por información sexual primaria, revela intereses de estudiantes más adultos ya. La presencia de gran número de lemas sin referencia física señala con precisión un programa que desarrollar en la enseñanza de la lengua materna.

¹² Muy posiblemente la búsqueda de la forma verbal *haya*, con ortografía errónea de *aya*, se muestra en la consulta electrónica como posibles sustantivos *ayo, aya*.

4. CARACTERIZACIÓN DE LAS FORMAS LÉXICAS BUSCADAS

Hasta aquí se ha expuesto una tentativa de explicación parcial sobre el comportamiento de los usuarios del *DLE*, deducida de los datos cuantitativos de miles de consultas coincidentes en un solo mes, el de abril de 2022¹³. Las líneas dibujadas de este comportamiento son claramente dos: una primera, relativamente limitada a solucionar con el diccionario cuestiones ortográficas y morfológicas; una segunda, sobre el interés o la curiosidad que explicaría la búsqueda de sustantivos, adjetivos y verbos, cuyas consultas parecen principalmente semánticas.

Un intento de caracterizar la primera línea sobre dudas ortográficas y morfológicas podría ser¹⁴:

1. Los lemas buscados responden mayoritariamente a cuestiones ortográficas. Es indicativo de ello los homónimos encontrados, cuya determinación siempre han señalado los manuales de ortografía (bes y uves, haches): *ostia* (7672) / *hostia* (10.385), *uso* (5193) / *huso* (2953), *basto* (8211) / *vasto* (11.538).

2. Lo explicado como simples faltas de ortografía señala en ciertos casos también alguna pronunciación del castellano peninsular central, yeísmo, o bien del seseo general del español: *haber* (37.421) o la forma *haya* en sus varias grafías, la de *hallar* (14.890) / *haya* (8830) / *aya* (45); la forma *haz* (23.333) por duda seseante.

3. El siguiente grupo es el que responde a búsquedas por la categoría gramatical en el que se han señalado palabras puramente gramaticales: *de* (58.460) y *asimismo* (5.899).

4. Otro grupo importante, pero limitado de lemas, para los estudiantes de español como lengua primera, se encuentra en cuestiones explicadas por la flexión verbal del español, muy rica por sus formas, según la concordancia, y por sus diferentes modalidades y matices: *dar* (23.188) y *deducir* (5491).

5. A lo largo de lo expuesto, se han comentado algunos lemas que son habituales por estos años o que son tópicos de características chocantes en el *DLE*. Prácticamente todos ellos responden a búsquedas dirigidas por una opinión pública generalizada que insistentemente recae en algunos lemas: *procrastinar* (21.185) y *almóndiga* (5524).

6. Si damos como muy probable que los usuarios son adolescentes mayoritariamente, otra explicación plausible puede ofrecerse para las búsquedas de lemas con referente sexual o tabuizante, que todo adolescente necesita resolver de una manera u otra, bien por curiosidad o bien por un instinto morboso fácilmente explicable: *follar* (31.816) y *sexi* (5238) o *concupiscencia* (5157).

El número de consultas tan alto sobre un número de formas relativamente limitado que ofrecen estas búsquedas en el diccionario, se contraponen al resultado que dan los

¹³ Los altos números que alcanzan las consultas pueden generar dudas sobre la exactitud de su registro. Sin embargo, en los cómputos estadísticos de otros meses y años aparecen en cantidades similares (Battaner 2020), lo que aseguraría la correspondencia real entre las búsquedas y su representación numérica, aunque esta no fuera exacta.

¹⁴ Se ejemplifica la exposición con casos representativos elegidos entre tramos alejados. La cantidad de consultas recibidas va entre paréntesis.

trabajos de investigación sobre el uso del o de los diccionarios hechos sobre encuestas, en los que, mayoritariamente, se concluye que el diccionario se usa para buscar el significado de las palabras. Esta contraposición explica la diferente visión, subjetiva y objetiva, que representan las respuestas a encuestas, siempre particulares, frente a los datos contables anónimos recogidos electrónicamente. El interesante trabajo de Santamaría y Azorín (2017), último de la serie de artículos que las investigadoras han dedicado al uso de los diccionarios, recoge mayoritariamente, una vez más, que las respuestas mayoritarias de los encuestados es que usan el diccionario para aclarar significados; conclusión que hay que matizar con lo interpretado de los datos electrónicos.

La segunda línea que estas consultas al diccionario permiten extraer incide en el tipo de formas buscadas por dudas semánticas, es decir, lo propiamente esperado de un diccionario general. Según disminuye en los tramos el número de consultas por forma, las identificadas como lemas van creciendo en variedad y pueden empezar a caracterizarse temáticamente en líneas generales a partir del tramo quinto.

Para caracterizar en estas conclusiones los tipos de palabras o lemas por tramos, se limita la elección de ejemplos a una forma del tramo quinto (entre 12.500 y 15.000 consultas) y otra del último analizado, el tramo undécimo (entre 6000 y 5000 consultas).

7. Los datos ofrecen adjetivos precisos, fundamentalmente cultos, lo que señala la necesidad en este nivel de enseñanza por ampliar el vocabulario de la lengua materna: *indulgente* (13.359) y *fútil* (5056).

8. Se recogen abundantes verbos que pueden ser dudosos por su significado o por su ortografía o flexión: *hallar* (14.890) y *anonadar* (5536).

9. Adjetivos y verbos pueden ser buscados por dificultades sintáctico-semánticas para encontrar la adecuación con sus argumentos: *intrínseco* (11.943) y *hostil* (5820); *discernir* (10.634) y *jactar* (5171).

10. Resulta difícil caracterizar los sustantivos recogidos. Entre las búsquedas se encuentra vocabulario nuclear con referente físico indeterminado: *persona* (17.269) y *vástago* (5139).

11. Se recogen sustantivos de significado conceptual, propiamente semitérminos, *estado* (13.704) y *condición* (5968).

12. Hay también sustantivos conceptuales de disciplinas básicas de la enseñanza obligatoria: *historia* (13.043) y *alegoría* (5815).

13. Coinciden sustantivos sin referente físico con adjetivos buscados del mismo tipo: *respeto* (12.807) y *altruismo* (5903).

14. Hay neologismos explicables por su novedad, *sororidad* (7915), *procrastinar* (21.185), *topar* (6412), etc.

15. Se escapan a estos tipos de explicación a mi ver, *idiosincrasia* (9700), *patriotero* (8049), *salbanda* (6322), *estricnina* (5691), *tallullo* (5438) y otras.

16. Se encuentra alguna atención a búsquedas tópicas en el *DLE*, lo que indica que hay un conocimiento extendido socialmente de lo que el diccionario de la Academia puede recoger como no normativo: *bizarro* (13.496) y *haiga* (5786).

No es explicable por qué unos lemas reciben más búsquedas que otros cuando presentan, sin embargo, menor frecuencia. Para ello recurrimos a confrontar el número de consultas hechas al Diccionario con los datos de frecuencia que tienen en el CORPES XXI como referencia (Tabla 2).

Lema	Categoría Nº consultas	Frec. general CORPES	Frec. norml. sin sign.
<i>indulgente</i>	A 13.359	346 casos	0,99 casos por millón
<i>hallar</i>	V 14.890	28.252 casos	81,43 casos por millón
<i>intrínseco, ca</i>	A 11.493	2.212 casos	6,37 casos por millón
<i>discernir</i>	V 10.604	1.370 casos	3,94 casos por millón
<i>persona</i>	N 17.269	283.132 casos	816,13 casos por millón
<i>estado</i>	N 13.704	198.525 casos	572,25 casos por millón
<i>historia</i>	N 13.043	165.633 casos	477,43 casos por millón
<i>respeto</i>	N 12.807	25.407 casos	73,23 casos por millón
<i>bizarro, rra</i>	A 13.496	541 casos	1,55 casos por millón
<i>fútil</i>	A 5.056	248 casos	0,71 casos por millón
<i>anonadar</i>	V 5.536	442 casos	1,27 casos por millón
<i>hostil</i>	A 5.820	2.972 casos	8,56 casos por millón
<i>jactar</i>	V 5.171	1.046 casos	3,01 casos por millón
<i>vástago</i>	N 5.139	841 casos	2,42 casos por millón
<i>condición</i>	N 5.968	98.287 casos	283,31 casos por millón
<i>alegoría</i>	N 5.815	1.248 casos	3,59 casos por millón
<i>altruismo</i>	N 5.903	454 casos	1,30 casos por millón
<i>haiga</i>	N 5.786	67 casos	0,19 casos por millón

TABLA 2. Relación entre el número de consultas recibido por un lema en el DLE en abril 2022 y la frecuencia asignada al mismo lema en el CORPES XXI (consulta, septiembre 2022)

REFERENCIAS BIBLIOGRÁFICAS

- AZORÍN FERNÁNDEZ, D. (2000): «¿Para qué usan el diccionario los escolares? Reflexiones a propósito de una encuesta», en Stephan Rhustaller y Josefina Prado, coords., *Tendencias en la Lexicografía Española Actual*, Huelva, Universidad de Huelva, pp. 75-100.
- BATTANER, P. (1994): «L'ús dels diccionaris escolars», *Caplletra*, 17, pp. 109-124.
- BATTANER, P. (2002): «Investigación en Didáctica de la lengua y de la literatura y la consolidación del Área de Conocimiento», *Revista de Educación*, 328, pp. 59-80.

- BATTANER, P. (2020): «¿Qué sabemos de los usuarios del *Diccionario de la Lengua Española* en línea?», en Real Academia Española, *Crónica de la lengua española 2020*, Madrid, Espasa, pp. 286-301.
- BATTENBURG, J. D. (1991): *English monolingual learners' dictionaries. A user oriented study*, Tübingen, Max Niemeyer.
- IBRAIM, A. H. y M. ZALESKY, coords. (1989): «Enquête: L'Usage des dictionnaires», en *Lexiques*, número especial de *Le Français dans le monde*, pp. 24-30.
- SANTAMARÍA PÉREZ, I. y D. AZORÍN FERNÁNDEZ (2017): «El diccionario monolingüe en el aula», en M. J. Domínguez Vázquez y M. T. Sanmarco Bande, coords., *Lexicografía y didáctica. Diccionarios y otros recursos lexicográficos en el aula*, Peter Lang, pp. 109-131.