

LA INVESTIGACIÓN DE LAS VOCES CULTAS
A TRAVÉS DE LOS CORPUS HISTÓRICOS
(The Investigation of Cultured Lexical Borrowings
(Learned Words) through Historical Corpora)

CHRISTOPHER J. POUNTAIN
ISABEL GARCÍA ORTIZ
Queen Mary, Universidad de Londres

RESUMEN

Los corpus históricos proporcionan valiosas evidencias de la aparición y frecuencia de los cultismos léxicos, y permiten establecer con cierto grado de precisión sus significados y contextos de uso. Es más, pueden ofrecer información secundaria sobre las circunstancias de su difusión y sobre las actitudes de los hablantes. Basándonos en los datos recuperables de los corpus, mostramos cómo es también factible, hasta cierto punto, tener una visión de la difusión diastrática de estas palabras e incluso evaluar cuándo pasan a formar parte del léxico disponible de los hablantes. No obstante, siempre es imprescindible una mediación filológica de los datos en bruto y el ejercicio de un cierto grado de ingenio a la hora de analizar los datos.

PALABRAS CLAVE: Etimología, préstamo léxico, cultismo, léxico disponible, corpus.

ABSTRACT

Historical corpora offer valuable evidence of the appearance and frequency of cultured lexical borrowings (learned words), and allow us to establish with some precision their meanings and contexts of use. In addition, they may also offer secondary evidence on the circumstances of their diffusion and on speaker attitudes towards them. We show how, on the basis of textual evidence recoverable from the corpora, it is also possible to give some account of the diastratic diffusion of these words and even the degree to which they may come to form part of speakers' available vocabulary. However, philological mediation of the raw evidence is always essential, and such conclusions can only be drawn by exercising a degree of ingenuity in the manipulation of the data.

KEY WORDS: Etymology, lexical borrowing, learned vocabulary, available vocabulary, corpora.

1. INTRODUCCIÓN

El presente artículo tiene por objetivo mostrar algunos de los retos que se nos han presentado a la hora de acudir a los corpus en línea del espa-

ñol en los estudios llevados a cabo por la línea de investigación *Loaded Meanings* en el marco del proyecto *Language Acts and Worldmaking*¹ y como parte de la *Open World Research Initiative (OWRI)*², que tiene por objeto de estudio el fenómeno de las voces cultas en español y otras lenguas de Europa occidental.

Una de las más conocidas limitaciones de los corpus lingüísticos es que estos no pueden ofrecer información más allá de su propio contenido (Hunston 2002: 22; Cheng 2012: 175). Dado que la decisión con respecto a los contenidos de cada corpus depende del propio diseñador, este solo puede mostrarnos aquella información que su diseñador quiere que veamos. CORDE, por ejemplo, no permite lematizar búsquedas u obtener datos con respecto a la frecuencia relativa de las consultas realizadas. La búsqueda en corpus lingüísticos requiere de un considerable ingenio heurístico, así como de la mediación filológica de los resultados de búsqueda (para una discusión más detallada de ambos aspectos, Pountain 2011).

2. DEFINICIÓN Y DELIMITACIÓN DE LAS VOCES CULTAS

Uno de los rasgos que caracteriza el estudio de las voces cultas es la confusión terminológica y metodológica que acompaña a este campo de investigación. Se trata de los desafíos que plantean los términos propios de este campo (nos referimos, en concreto, a los reiterados intentos por establecer una clara diferenciación entre los términos *cultismo* y *latinismo*, por no mencionar *semicultismo*³) y de las inconsistencias en los parámetros de identificación. Esto es debido a la existencia de lagunas históricas, así como a las divergencias en los intereses de cada investigador.

A raíz del uso indistinto de los términos *latinismo* y *cultismo*, algunos autores han planteado la necesidad de discernir entre lo que, para ellos, son realidades independientes. Para Bustos Tovar (1974: 23-24), el común denominador de ambos tipos de palabras es su procedencia latina de influjo culto; no obstante, en línea con la propuesta de Américo Castro, propone reservar el término *latinismo* para aquellas palabras que no han sufrido absolutamente ninguna variación, como *maximum*, y el término *cultismo* para aquellas palabras que han sufrido una adaptación fonética mínima, como *máximo*. Clavería Nadal (1991: 39-66), en cambio, parte de la teoría del préstamo y propone denominar *latinismo* a lo que autores como Bustos Tovar denominan *cultismo*; es decir, a aquellos elementos latinos que el castellano ha tomado en préstamo. El *cultismo*, para la autora, es un elemento que no necesariamente se ha tomado del latín y que no puede considerarse

¹ Véase <<https://languageacts.org>>.

² Véase <<https://ahrc.ukri.org/research/fundedthemesandprogrammes/themes/owri/>>.

³ El análisis del polémico término *semicultismo* se encuentra fuera del alcance de esta publicación. Para una explicación más detallada, referimos a Azofra Sierra (2006: 237-238), Clavería Nadal (1991: 14-19) y Wright (1976).

un préstamo en la lengua, pues es más bien un recurso expresivo utilizado de manera puntual y consciente. En palabras de la autora, el *latinismo* es un elemento exclusivamente tomado del latín que ha arraigado en la lengua receptora y se ha convertido en una “propiedad estructural” de la misma (Clavería Nadal 1991: 66). El cultismo, por otra parte, está relacionado “con modos de expresión cercanos a la estilística” (Clavería Nadal 1991: 6). Por tanto, latinismo y cultismo coincidirían únicamente en aquellos casos en los que el empleo consciente de un término extranjero, con fines estilísticos o técnicos, se hiciera a partir de bases latinas. Especifica la autora que “en su generalización posterior, este continuará siendo un latinismo, pero puede que pierda su valor de cultismo” (Clavería Nadal 1991: 40). Propuestas más tardías, como la de García Valle (1992, 1998), orientan el debate hacia la distinción en la vía de introducción de las formas latinas, es decir, oral o escrita. Para la autora, el latinismo engloba todas aquellas formas latinas introducidas por vía escrita y que, en su opinión, probablemente se introdujeron en la escritura a manos de los escribas con el fin de cubrir vacíos léxicos en una época en la que el castellano como vehículo de cultura aún estaba gestándose⁴. El cultismo, por el contrario, engloba todas aquellas formas latinas introducidas por vía oral como alternativas frente a la vacilación de formas de pronunciación de una misma palabra. Reconoce como cultismos tanto las palabras que en la actualidad mantienen dobles (*ánima* frente a *alma*, de lat. *ānīma*) como aquellas formas en pugna durante la época medieval, y de las cuales triunfó la variante más culta en detrimento de las variantes populares. Tal es, en opinión de la autora, el caso de *calumnia*, recogida en documentos de la época bajo distintas formas: *calumpnia*, *calonia*, *callopnia*, *calopnia* (1992: 95). Para la autora, el rasgo que diferencia inequívocamente al latinismo es la inexistencia de alternativas coexistentes:

Los cultismos, por lo tanto, se caracterizarían por introducirse oralmente, como una alternativa culta aceptada por los hablantes con el propósito de acabar con la vacilación que existía entre diversas formas de pronunciación de una misma palabra, desechando así las múltiples variantes populares que coexistían hasta ese momento [...] En cambio, los latinismos se presentarían en la escritura sin variantes populares alternativas o coexistentes con ellos, en esto se diferenciarían de los cultismos [...] (García Valle 1998: 67).

Las tres propuestas mencionadas muestran ciertas limitaciones. En primer lugar, ninguna de ellas es incluyente de las voces procedentes de otras lenguas de cultura, como es el griego. Es este un punto que ya se ha tratado con anterioridad (por ejemplo, Azofra Sierra 2006: 234). Esta exclusión elude uno de los aspectos más interesantes acerca del fenómeno de las voces cultas: a saber, cómo contribuyeron las lenguas de cultura a la creación y enriquecimiento de las vernáculos, ya fuera a través del latín, del griego, del griego a través del latín o incluso de cualquiera de estas dos lenguas a

⁴ Partiendo de los postulados de Wright (1982), la autora sitúa tal período a partir del siglo XI en España (García Valle 1992: 94).

través de las propias vernáculos. Por ejemplo, *problema* era en sus orígenes una palabra poco común en griego que fue tomada en préstamo por el latín e introducida en las lenguas romances y en inglés como resultado de las traducciones vernáculos de la versión latina de la obra de Aristóteles titulada *Problemata* (Pountain 2016). Sin embargo, *problema* se ha convertido en la 145ª palabra más frecuente del español del siglo XX según la lista de frecuencias de Davies y Davies (2017).

En segundo lugar, estas propuestas dificultan la aplicación de estos términos a corpus de palabras más amplios, aun si reducimos el corpus de nuestro estudio únicamente a términos de origen latino. La propuesta de García Valle (1992, 1998) es la que más limitaciones presenta, ya que parece circunscribirse a un período histórico muy concreto y tiene la desventaja añadida de que, con frecuencia, no es posible contrastar mediante un método científico muchas asunciones relativas a la lengua oral, porque dicha realidad histórica no es accesible a través de las fuentes documentales. No obstante, las propuestas de Bustos Tovar (1974) y Clavería Nadal (1991) también resultan limitantes, ya que subordinan la definición de cultismo a una cuestión, en el fondo, sociolingüística. Bustos Tovar (1974: 40-42) especifica claramente que el estudio histórico de la palabra deberá determinar “si se mantuvo como cultismo, o se incorporó al léxico popular”. Clavería Nadal (1991), en realidad, pretende precisamente lo contrario: deslindar la teoría del cultismo (para ella, latinismo) de criterios sociológicos surgidos a raíz del doble significado de los calificativos *culto/restringido* y *popular/común*⁵. Es por ello que propone reservar el término *cultismo* para aquellos rasgos lingüísticos relacionados con cuestiones estilísticas (ya sea de carácter formal, técnico o literario) y referirse a los latinismos como cultismos únicamente en aquellos casos en los que estos se empleen de manera aislada y con fines estilísticos. Al establecer esta división, no obstante, empuja a considerar a los cultismos de uso más frecuente e «inmotivado» como un conjunto de elementos aparte, que recibirán su nombre en función de su lengua de origen (latinismo, grecismo, etc).

Estamos con Malkiel (1957: 18) en que “los términos *culto*, *semiculto*, *patrimonial* representan normas del desarrollo fonético y no deben confundirse con los criterios del análisis sociológico”. Es más, viene implícito y forma parte de nuestro estudio asumir que el cambio está en la naturaleza de la lengua y que las voces cultas que en sus orígenes estuvieron limitadas a una comunidad sociolingüística reducida hoy pueden ser parte de la lengua común. De acuerdo con esta visión, las palabras que conforman nuestro objeto de estudio son aquellos préstamos tomados directamente del latín, o del griego a través del latín, en la Edad Media y épocas posterior-

⁵ La autora (1991: 12) pone como ejemplo el doblete *rápido/raudo*. Desde el punto de vista formal, *rápido* es un cultismo, ya que ha sufrido una menor adaptación fonética, mientras que *raudo* es una palabra patrimonial que sí ha consumado la evolución fonética que le correspondía. No obstante, algunos autores aducen que, desde el punto de vista sociolingüístico, *raudo* es una palabra culta, pues es de uso poco común y está a menudo relegada a contextos formales, mientras que *rápido* es una palabra patrimonial porque es de uso frecuente, i.e. es “patrimonio de todos”.

res. Los casos más claros de lo que constituye esta definición de cultismo son aquellas voces que (a) constituyen excepciones a la evolución fonética regular y se han adaptado en grado mínimo a la lengua receptora, (b) aparecen en una fecha relativamente tardía, sea en el proceso de elaboración de las lenguas vernáculas, conforme estas comienzan a ganar preferencia en detrimento del latín en contextos administrativos, técnicos y artísticos, sea en épocas posteriores (el Renacimiento y el Siglo de las Luces son especialmente fructíferos), (c) tienen el mismo sentido que la forma correspondiente latina, o, por lo menos, lo tenían cuando aparecieron por primera vez, (d) pertenecen originalmente a un campo semántico más bien intelectual y a un nivel de discurso elevado. Por este motivo, el uso de los cultismos se circunscribe inicialmente a una élite social culta y se limita a la lengua escrita. Con el transcurso del tiempo, no obstante, es posible (y mucho más probable de lo que se suele pensar) que algunas de estas palabras emprendan su camino hacia la lengua oral, un fenómeno que la lingüística histórica ha dado en llamar desplazamiento hacia abajo (ing. *downward migration*), e incluso lleguen a ocupar puestos relativamente altos en las listas de frecuencias. Como iremos demostrando, los siglos XIX y XX son decisivos para la fijación de cultismos cuyas primeras documentaciones se remontan siglos atrás en el tiempo.

3. CONSIDERACIONES METODOLÓGICAS

Nos interesa principalmente trazar las dinámicas de lo que se ha denominado el problema de integración (ing. *embedding problem*) (Weinreich *et al.* 1968) con respecto a las voces cultas. Si bien los estudios clásicos de las voces cultas han girado en torno a las primeras documentaciones de estas palabras, la mera aparición de un cultismo en un texto no garantiza que este goce de aceptación o que se haya incorporado plenamente a la lengua, especialmente si el texto en particular pertenece a un registro especializado o técnico (fenómeno este bastante frecuente en el caso de los textos que constituyen la base de nuestro conocimiento con respecto a las lenguas medievales de Europa occidental). Nuestro objeto principal es estudiar las voces cultas que en el español actual se hallan entre las voces más comunes de la lengua y nuestro interés se centra, por tanto, en analizar el proceso de integración de las voces cultas, así como en estudiar qué mecanismos propician su progresivo desarrollo hasta llegar a designar conceptos cognitivos básicos. Es por ello que nuestro corpus base está tomado del diccionario de frecuencias de Davies y Davies (2017). Se trata de un diccionario de frecuencias actualizado que recoge las cinco mil palabras más frecuentes de nuestro idioma. En concreto, nuestra línea de investigación trabaja, por un lado, con las voces de origen culto que figuran entre las 300 palabras más frecuentes en la lista de frecuencias de Davies (remitimos, por poner un ejemplo, a la Gráfica E de este mismo artículo); por

otro lado, para el análisis de la terminología médica de origen culto, las palabras se han extraído de entre las primeras 5.000 palabras más comunes según la misma lista de frecuencias.

La metodología seguida en la investigación se basa en la combinación de diccionarios y corpus, estos últimos históricos en su mayoría. Dado que el objetivo principal de la investigación es el análisis del proceso de difusión de las voces cultas en las lenguas de Europa occidental, se han consultado tanto fuentes del español como de los idiomas hablados en otras naciones europeas, particularmente el inglés, el francés, el italiano y el portugués.

Con respecto a las fuentes de datos del español, la metodología gira en torno al uso del CORDE y el CDE, principalmente⁶. Consideramos que la combinación de ambos corpus permite lograr una mayor fiabilidad en los resultados, ya que el CDE es especialmente útil a la hora de obtener datos estadísticos, mientras que CORDE resulta más práctico a la hora de obtener datos más concretos acerca de las documentaciones seleccionadas. En el ámbito de la lingüística de corpus, las comparaciones entre el CORDE y el CDE han puesto de manifiesto precisamente la utilidad que cada corpus tiene para las tareas señaladas. Destaca el análisis de Rojo Sánchez (2010), cuya conclusión, tras analizar tanto el proceso de codificación como las posibilidades de búsqueda de ambos corpus, demuestra que, si bien el CDE destaca por la rigidez del sistema de consultas (por ejemplo, no permite realizar consultas en una escala de tiempo inferior a un siglo, ni es posible obtener información detallada con respecto al país, autor o área temática) y los datos a gran escala (útiles para obtener datos estadísticos, como el aumento de la frecuencia de una determinada voz a lo largo de varios siglos), CORDE presenta un sistema de búsquedas mucho más flexible, que permite ahondar en la historia de cada ejemplo de uso obtenido. Por el contrario, admite el autor que el CORDE resulta menos eficaz a la hora de obtener datos estadísticos, debido, entre otros factores, a que no ofrece frecuencias relativas.

En cuanto a los problemas particulares que plantea CORDE, relativos principalmente a la fiabilidad de las fechas, se ha utilizado el CORDEMÁFORO (Rodríguez Molina y Octavio de Toledo y Huerta 2017) para aquellos casos en los que la obra referenciada se encontraba en la base de datos. En aquellos casos en los que la fuente no figura en la base de datos, se ha acudido a fuentes fiables, principalmente la *Biblioteca Digital Hispánica*

⁶ Resulta menos provechoso para nuestros propósitos el uso de corpus cuya información esté limitada geográficamente o temporalmente. Los cultismos, dada su naturaleza, trascienden a menudo variedades dialectales y límites regionales. Nuestro estudio es, además, diacrónico, por lo que nos interesan especialmente los corpus académicos, así como el corpus de Mark Davies; tanto CORDE como CDE poseen la ventaja de abarcar un marco cronológico amplísimo, cuyo límite es posible complementar con los datos del CREA. Es por ello que hemos prescindido de corpus como el CORDIAM, enfocado al análisis de las variedades del español de América, o el CODEA, que no ofrece documentaciones posteriores a 1800. No obstante, sí se ha acudido a otros corpus del español para la búsqueda de documentaciones o datos concretos que complementen la investigación en aquellos casos en los que la información proporcionada por los corpus académicos ha resultado insuficiente; es el caso de, por ejemplo, el DICCA-XV (remitimos al ejemplo 26 de este mismo artículo).

(BDH), con el fin de establecer, en la medida de lo posible, si el manuscrito consultado se corresponde con la publicación original o si, por el contrario, se trata de una edición posterior, una introducción o cualquier otro tipo de adición (Lleal Galceran 2013). No se han tenido en cuenta las fechas obtenidas a través del CDE, pues este último corpus se ha reservado para la elaboración de datos estadísticos.

Pese a las medidas tomadas, es inevitable que un estudio basado en corpus encuentre obstáculos difíciles de sortear, que precisamente constituyen la naturaleza de todo corpus lingüístico. La imposibilidad de reflejar la lengua en toda su totalidad da lugar a ciertos vacíos que la documentación textual, por rica que sea, difícilmente podrá cubrir (Kabatek 2016: 3-5). Para paliar esta carencia hemos sometido los datos, en la medida de lo posible, a una interpretación o, si se prefiere, a un proceso de *mediación filológica* (Pountain 2011).

4. MEDIACIÓN FILOLÓGICA: ESTUDIO DE UN CASO

Con el fin de ilustrar el intrincado proceso de mediación filológica que requiere el uso de corpus lingüísticos, proponemos analizar el caso de la palabra *existir*, así como los derivados léxicos *existencia* y *existente*. *Existir* ha experimentado una difusión tal que, según Davies y Davies (2017), es ahora la 194ª palabra más frecuente en español (*existencia* ocupa la 953ª posición en la lista de frecuencias, mientras que *existente* ocupa la 2271ª posición). Quizá el primer dato que interese establecer sea que, tal vez sorprendentemente, las formas *existencia* y *existente* preceden al verbo *existir*. En el contexto iberorromance, la primera documentación de *existencia* se encuentra en Juan Hernández de Heredia (1); *existente*, por su parte, se documenta más tardíamente, esta vez en *Visión deleytable* de Alfonso de la Torre (2). Ambos autores son conocidos innovadores de léxico culto.

- (1) Esto fizieron considerando que segunt que la uentura del hombre, la qual se coço por la jnfluencia de las estrellas segunt el lugar do ellas son & como se gouiernan segunt la senyoria del tiempo de su primera **existencia**, assi fues de las ciudades. (CORDE: Juan Fernández de Heredia, *Traducción de Vidas paralelas de Plutarco*, III, 1379-1384)
- (2) E çierto es que todos los omnes desean aver bien e fuyr el mal, e no es cobdiçada ninguna cosa por ellos que non sea buena o que non tenga alguna espeçie de bondad aparente o **existente** (CORDE: Alfonso de la Torre, *Visión deleytable*, c.1430-1440)

Existencia y *existente* se utilizan frecuentemente como antónimos de *apariciencia* y *aparente*, este hecho queda patente en el contexto de *existente* provisto en el ejemplo (2), mientras que en (3) observamos un uso similar de *existencia*. *Existir* es más difícil de localizar porque es necesario rastrear todo el paradigma de formas verbales. No hay casos fiables en castellano antes del siglo XVI (4).

- (3) Buena gente, grant engaño es tomar omne la apareççencia por la **existencia**. (CORDE: San Vicente Ferrer, *Sermones*, 1411-1412)

- (4) no hallé ninguno que hablase expresamente particularizando esta enfermedad, con la excepción de Velasco de Taranta, en el capítulo propio, que no hizo más que afirmar que algunas veces se produce esta enfermedad o pleuresía epidémica y maravillarse, como yo, de cómo los antepasados no hicieron cura ni capítulo específico puesto que todos, en especial los modernos, reconocen que tal enfermedad **existe**; pero tampoco él, como los pasados, quiso adentrarse en el caso, sino que lo omitió (CORDE: Diego Álvarez Chanca, *Tratado nuevo no menos útil que necesario...*, 1506)⁷

Además, a partir de la búsqueda en los corpus de otras lenguas, es posible deducir que el castellano no ha tomado en préstamo *existir* y sus derivados léxicos directamente del latín, sino que han llegado al castellano por influencia de otras lenguas, especialmente otras lenguas romances, ya que, tanto en italiano como en francés, los primeros ejemplos de uso de *esistenza/existence* y *esistente/existant* anteceden a los ejemplos de uso documentados en castellano. (De hecho, parece ser que Fernández de Heredia fue una figura clave en la introducción de vocabulario culto: Dworkin 2010). La primera documentación del francés *existence* en TLF, confirmada por FRANTEXT, data de 1330 (5), mientras que *existant* se documenta por primera vez en un texto de 1400 (6) en un contexto que imita el uso más frecuente de *existente* en el latín de los documentos legales del período medieval (7).

- (5) Entre non et **existence** / Veul je bien faire difference. (FRANTEXT: Guillaume de Digulleville, *Le Pèlerinage de vie humaine*, 1330-1331)
 (6) en son temps, Ascanius, filz de Eneas, **existant** roy de Ytalie après son pere, fonda la cité de Albe (FRANTEXT: Christine de Pisan, *Le Livre de la Mutacion de Fortune*, II, 1400)
 (7) domino Gondissalvo **existente** episcopo in Salamantina ecclesia (CORDE: Anónimo, Carta de donación [Documentos de los archivos catedralicio y diocesano de Salamanca], 1202)

Aunque se podría argumentar que *existant* ya era una forma verbal en francés, no hay testimonios de uso de otras formas de este mismo verbo hasta el siglo XVI:

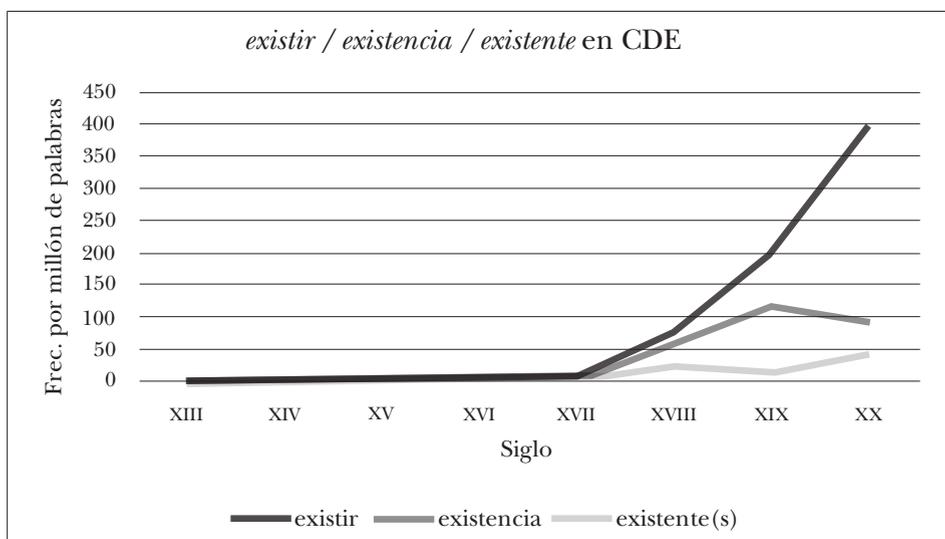
- (8) et c'est la premiere condition generale, absoluë et vniuerselle de son estre, en laquelle elle entre, au mesme moment qu'elle sort de l'estre qu'elle a en Dieu, pour **exister** en soy-mesme. (FRANTEXT: Pierre de Bérulle, *Discours de l'estat et des grandeurs de Jésus par l'union ineffable de la divinité avec l'humanité*, 1623)

Las palabras italianas *esistenza* y *esistente* también se documentan por primera vez en la primera mitad del siglo XIV (9). El verbo *esistere* no se menciona hasta la quinta edición del diccionario de la Accademia della Crusca (1863-1923). En inglés, *existence* se documenta por primera vez en Chaucer en lo que parece ser un préstamo claro del francés (10); *existent* no se documenta hasta el siglo XVI (11):

⁷ Hay un caso del siglo XV (*Pues, concluye aquel sabio que todas las cosas por trabajo e exerciçio del mouimien-to existen, e por consiguiēte la natura humana creze en su vigor corporal e spiritual, seyendo exercitado en tan noble deporte.* (CORDE: Rodrigo Sánchez de Arévalo, *Vergel de los príncipes*, 1454-1457)) en un texto que, según el CORDEMÁFORO (Rodríguez Molina y Octavio de Toledo y Huerta 2017), es de fiabilidad intermedia.

- (9) ed essere beato è sua propria essenza, ed **esistente** proprietà; e per questa cotalle **esistenza**, semplicità, ed incommutabilità conobbero (OVI: Domenico Cavalca (c.1270-1342), *Esposizione del Simbolo degli Apostoli*, antes de 1342)⁸
- (10) Forto se Hym that is freend in **existence** From hym that is by apparence (OED: Chaucer, *Romaunt Rose*, principios del siglo XV.)
- (11) In thynge whiche receyue perfeccyon, the matier is **existent** or hath his beyngre afore the forme (OED: Marsilio de Padua, *The Defence of Peace (Defensor Pacis)*, trad. W. Marshall, 1535 (El texto original en latín es: *materia formae praexistit in actu*))

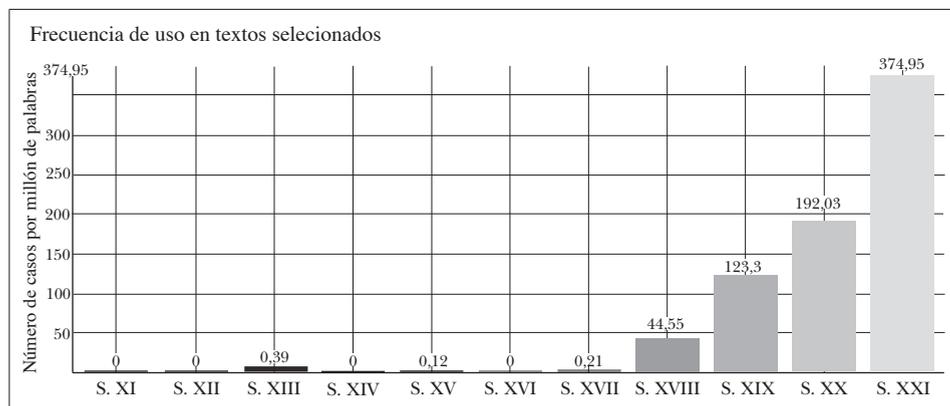
Volviendo a la familia léxica de *existir* en castellano, el CDEGH permite trazar el aumento de la frecuencia de estas palabras siglo por siglo. Tal y como se puede observar en la Gráfica A, *existir* sobrepasó a *existencia* y *existente* en cuanto a frecuencia relativa en el siglo XVIII, y continuó su trayectoria ascendente hasta el siglo XX. Los datos disponibles en *Enclave* (véase la Gráfica B) sugieren que esta tendencia continúa en el presente. Volveremos sobre la importancia de estas tendencias en 6.



Gráfica A: *Existir, existencia y existente* en el CDEGH⁹

⁸ Cavalca fue el autor de numerosos tratados en lengua vernácula derivados de fuentes originales latinas.

⁹ En la mayoría de las gráficas, la información cronológica (por siglos) se sitúa en el eje horizontal sin mayor subdivisión; el eje vertical muestra la frecuencia por millón de palabras, ya que una frecuencia menor dificultaría la apreciación de las tendencias.

Gráfica B: *Existir en Enclave*

Las documentaciones históricas de *existir* y sus derivados léxicos nos permiten trazar la evolución semántica de estas palabras con cierto grado de precisión. *Existencia* ha ampliado su significado desde lo que puede considerarse su sentido filosófico original de “realidad concreta”, en oposición a *aparencia* “realidad aparente”, aunque se ha mantenido en gran medida como sustantivo abstracto. Como tal, no obstante, es una nominalización conveniente de la noción existencial y se utiliza frecuentemente con este significado. Sí desarrolló, en cambio, un uso como sustantivo contable “tipo de existencia” (12) a partir del siglo XVI. Dicho significado está más estrechamente relacionado con el significado concreto de “vida, experiencia”, que se documenta a partir del siglo XVIII (13). Un sentido más restringido de la palabra se encuentra en la acepción castellana “mercancía, producto” (14), que data igualmente del siglo XVIII.

- (12) Y ponderad más cómo goza el Redentor con estas dos **existencias** de mayor excelencia que con sola la natural (CORDE: Juan de Pineda, *Diálogos familiares de la agricultura cristiana*, 1589)
- (13) Mil grandiosas esperanzas / Eran mi **existencia** toda (CDEGH: Manuel José Quintana (1772-1857), *Poesías*, 1814)
- (14) Y si a esto agrega usted el valor de las **existencias** indicadas [...] (CDEGH, Gaspar Melchor de Jovellanos (1744-1811), *Correspondencia*, 1778)

El aumento de la frecuencia de *existir* parece guardar relación, no tanto con su ampliación de significado o con el desarrollo de nuevos contextos sintagmáticos, como con lo que parece ser una creciente disposición con respecto a su uso, un fenómeno que en ocasiones se ha denominado cambio de marcación (ing. *shift of markedness*) (Trask 1996: 139-143). En el siglo XIX encontramos ejemplos de uso de *existir* fundamentalmente en tres sentidos: (a) en el sentido que quizá podría ser considerado como su sentido literal u original de “ser real y verdadero”, que se utiliza para enfatizar la realidad de algo (15); (b) otro sentido plenamente léxico como equivalen-

te de “estar vivo, vivir” (16); y (c) lo que parece ser una noción existencial semigramaticalizada (17):

- (15) un impedimento secreto; una dispensación condicional de un impedimento que acaso no **existe** (CORDE: Francisco Navarro Villoslada, *Doña Urraca de Castilla*, 1849)
- (16) ¿Y mi hijo? ¡Oh! Calla; si ya no **existe**, no me lo digas ahora. (CORDE: Francisco Navarro Villoslada, *Doña Urraca de Castilla*, 1849)
- (17) esa misteriosa relación que entre Ramiro y Bermudo **existe** (CORDE: Francisco Navarro Villoslada, *Doña Urraca de Castilla*, 1849)

Resulta conveniente pararse a analizar los pormenores del uso existencial de la palabra. En (18) puede observarse que el español utiliza una oración de relativo que involucra al verbo *existir*, mientras que en inglés tal construcción no siempre es un requisito imprescindible (cf. Stockwell *et al.* 1965: 95-6; Wonder 1979). En el español actual, existe cierta reticencia a la hora de calificar un sustantivo mediante una expresión locativa, por lo que esta se sustituye por una oración de relativo. Así, en inglés, la oración *the man on the roof* se corresponde, en español, con la oración *el hombre que está en el tejado*. El sentido existencial de *existir* se ve favorecido particularmente por este tipo de contexto:

- (18) la desconfianza que **existe** fuera de España hacia la tan promovida “salud del sistema bancario español” (CDEWD: <http://decrecimientocordoba.blogspot.co.uk/2013/08/en-ocho-meses-hemos-consumido-lo-que-la.html> [9 de enero 2017])
- (19) Romário de Souza Farías ha asegurado que **existe** en Brasil una campaña para difamarle (CREA: *La Vanguardia*, 30.8.1995)

No parece ser una coincidencia que *existente* se utilice, aparentemente con mucha más anterioridad según los registros textuales, como equivalente de *que existe* en su sentido existencial. Mientras que el sentido original de *existente* como adjetivo atributivo utilizado en oposición a *aparente* se mantuvo (20), desde finales del siglo XV observamos el desarrollo de una función existencial, en la cual *existente* introduce una frase locativa en forma de atributo nominal (21):

- (20) todo prójimo ha de ser amado según el grado de su bondad **existente** o posible (CORDE: Juan de Pineda, *Diálogos familiares de la agricultura cristiana*, 1589)
- (21) por el qual açercamiento el espíritu animal **existente** en los tales nervios & carne sea inmundado por la propiedad de la cosa tocada (CORDE: Fray Vicente de Burgos, traducción de *El Libro de Proprietatibus Rerum de Bartolomé Anglicus*, 1494)

A este respecto, es interesante un ejemplo de finales del siglo XVIII en el que se ofrece el equivalente en castellano de una oración escrita originalmente en latín: en (22), *existentes en el reino* corresponde a *qui in illo regno erant*. Este uso tiene continuación hoy día en registros formales.

- (22) convocó Cortes generales para esta capital llamando a los arzobispos, obispos, abades, condes, príncipes y duques **existentes** en el reino según refiere el escritor de la Crónica latina de don Alonso VII, autor coetáneo. “Constituit diem celebrandi concilium apud Legionem, civitatem regiam, IV nonas junii in die Sancti Spiritus

cum archiepiscopis, et episcopis, abbatibus, comitibus, principibus qui in illo regno erant.” (CDEGH: Francisco Martínez Marina (1754-1833), *Teoría de la Cortes o Grandes Juntas Nacionales*, 1794)

Hay indicios que demuestran que *existir* se está utilizando cada vez con más frecuencia en ciertas circunstancias como alternativa directa del propio verbo existencial *hay*, tal vez porque *hay* presenta una estructura fonética débil¹⁰. Cuando se da este uso, *existe/existía* suele ocupar la posición más cercana al extremo izquierdo de la oración, al igual que *hay* (23). Las oraciones subordinadas parecen favorecer particularmente este uso (24, 25).

- (23) Riera asegura que no hay problemas lingüísticos ni de convivencia: “**Existe** un auténtico bilingüismo” (CREA: *La Vanguardia*, 2.5.1995)
- (24) Tal hecho parece indicar que no **existe** en nuestro país nadie que pueda ejercer sus funciones (CREA: *ABC Cultural*, 29.11.1991)
- (25) Es cierto que en Barcelona **existe** un verdadero amor por la música (CREA: *La Vanguardia*, 2.11.1995)

El paralelo desarrollo de este fenómeno en el portugués de Brasil parece ser particularmente acusado¹¹.

Lo que demuestra esta breve incursión en la trayectoria del verbo *existir* y sus derivados léxicos es que, mientras que los corpus en línea tienen la capacidad de ofrecernos evidencias estadísticas en bruto con respecto al aumento de frecuencia, los datos pormenorizados con respecto al proceso de integración del léxico –cambio de significado, ampliación de significado, funciones sintácticas e incluso colocaciones frecuentes– solo pueden obtenerse mediante un procesamiento manual de dichas evidencias a cargo del propio investigador. Un corpus contiene la base de esta información, pero esta no viene expresamente descodificada, por lo que permanece oculta hasta el momento en que se lleva a cabo este procedimiento analítico.

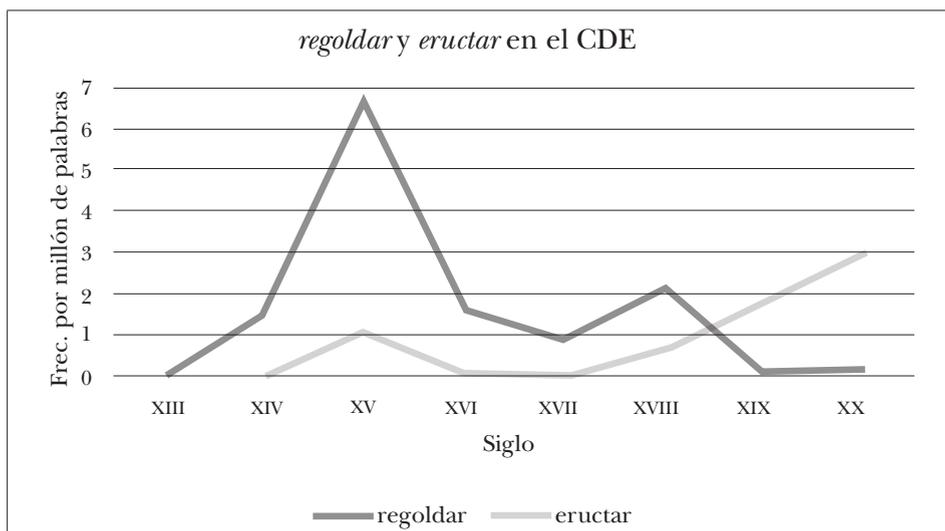
5. EL CORPUS COMO FUENTE DE INFORMACIÓN SOBRE LA PERCEPCIÓN LINGÜÍSTICA

Los corpus lingüísticos también pueden revelar información sobre otras cuestiones lingüísticas, como pueden ser los juicios en torno a los actos lingüísticos (Bloomfield 1944), un fenómeno que recientemente se ha denominado “dialectología perceptiva” (Preston 1999). Uno de los casos más conocidos es el de la palabra *eructar* y su competición con la palabra

¹⁰ NGLÉ 1114 menciona *existir*, *ocurrir* y *tener* como verbos con un significado similar al verbo existencial *haber*. No obstante, *existir* presenta un campo de significación mucho más restrictivo que *haber*: (a) no es invariable, al contrario de lo que sucede con *haber*, sino que concuerda en número con el sustantivo que introduce (*existen muchas posibilidades de ayuda*); (b) no puede asumir un objeto pronominal (–¿Hay tigres en la India? –Sí, los hay. Por el contrario –¿Existen tigres en la India? –*Sí, los existen.); (c) no posee el significado de “acontecimiento” que sí tiene *haber* (*Hubo/Existió un silencio prolongado*).

¹¹ En el caso del portugués de Brasil, Whitlam (2011: §38.1.3) incluye formalmente *existir* como una alternativa a *há* y (en el portugués de Brasil) a *tem* en ciertas circunstancias: *Existem pessoas que acreditam em reencarnação* “existen personas que creen en la reencarnación”.

regoldar. Aunque se trate de palabras poco frecuentes en los registros escritos, es posible atisbar en cierta medida el proceso de competición entre ambas palabras gracias a los datos extraídos del CDEGH¹² (véase la Gráfica C).



Gráfica C. *Regoldar y eructar en el CDEGH*¹³

La palabra *regoldar* posee un origen incierto. No obstante, se trata, con toda probabilidad, de una palabra patrimonial: Corominas y Pascual (1980-91: IV, 851-3) sugieren que su procedencia etimológica es **regũrgĩtare*, que, pese a no estar documentada en latín clásico, es una palabra que sí está generalizada como cultismo en el discurso médico (esp. *regurgitar*, fr. *rẽgurgiter*, it. *rigurgitare*, ing. *regurgitate*) (*OED*). Aunque las documentaciones de la palabra *regoldar* son escasas, el DICCA-XV, por ejemplo, ofrece el siguiente ejemplo de uso procedente del tratado gastronómico de Enrique de Villena, fechado en 1423:

- (26) [...] porque non tenga mal gesto segunt fazen los beuedores e deshordenados comedores e porque non **reguelde** o escupa o tosa o bosteze o estornude o le huela mal el resollo antes deue vsar salsas [...] (DICCA-XV: Villena, Enrique de, *Arte cisorĩa o Tractado del arte de cortar del cuchillo*, 1423)

Eru(c)tar es claramente un cultismo; el sustantivo *eructuación* parece antecederlo y, de hecho, la primera documentación de esta palabra en CORDE, que data del siglo XV, nos ofrece información secundaria sobre el grado de desconocimiento que provocaba inicialmente esta palabra. En el

¹² De los datos obtenidos en el CDEGH con respecto a la palabra *eructar*, se han excluido aquellos ejemplos que eran claramente referencias a la palabra latina *eruct(u)o*.

¹³ Esta gráfica incluye tanto las formas verbales como los sustantivos correspondientes (*regüeldo*, *regoldación* y *eructo*, *eruct(u)ación*).

siguiente ejemplo, se muestra cómo la palabra *eructación* aparece glosada junto a *regoldamiento*:

- (27) Las señales de la diafragma llagada son espesso & grand aliento. & tos sonante & dolorosa alienacion. & escupimiento cardeno sed & fastio de la vianda con **eructacion**: o regoldamiento & rigor pungitiuo. & el lugar cerca las faltas costillas. (CORDE: Anónimo, *Traducción del Tratado de cirugía de Guido de Cauliaco*, 1493)

Tal y como señalan Corominas y Pascual (1980-91: II, 660), aproximadamente un siglo más tarde Cervantes nos ofrece en *Don Quijote* información relevante con respecto al proceso de integración de *eructar* en español:

- (28) “Esso de *erutar* no entiendo”, dixo Sancho.

Y don Quixote le dixo:

“*Erutar*, Sancho, quiere dezir *regoldar*; y este es vno de los mas torpes vocablos que tiene la lengua castellana, aunque es muy sinificatiuo; y, assi, la gente curiosa se ha acogido al latin, y al *regoldar* dize *erutar*, y a los *regueldos*, *erutaciones*; y quando algunos no entienden estos terminos, importa poco, que el vso los yra introduziendo con el tiempo, que con facilidad se entiendan, y esto es enriquezer la lengua sobre quien tiene poder el vulgo y el vso”.

“En verdad, señor”, dixo Sancho, “que vno de los consejos y auisos que pienso lleuar en la memoria ha de ser el de no *regoldar*, porque lo suelo hazer muy a menudo”.

“*Erutar*, Sancho, que no *regoldar*”, dixo don Quixote.

“*Erutar* dire de aqui adelante”, respondió Sancho, “y a fee que no se me oluide”. (Miguel de Cervantes, *Don Quijote* II, 43, texto electrónico por Fred F. Jehle a base de la edición de Rodolfo Schevill y Adolfo Bonilla (Madrid: Gráficas Reunidas, 1928) <<https://users.ipfw.edu/jehle/cervante/dqsb1615.htm>> [30 de junio 2019]; cursiva nuestra)

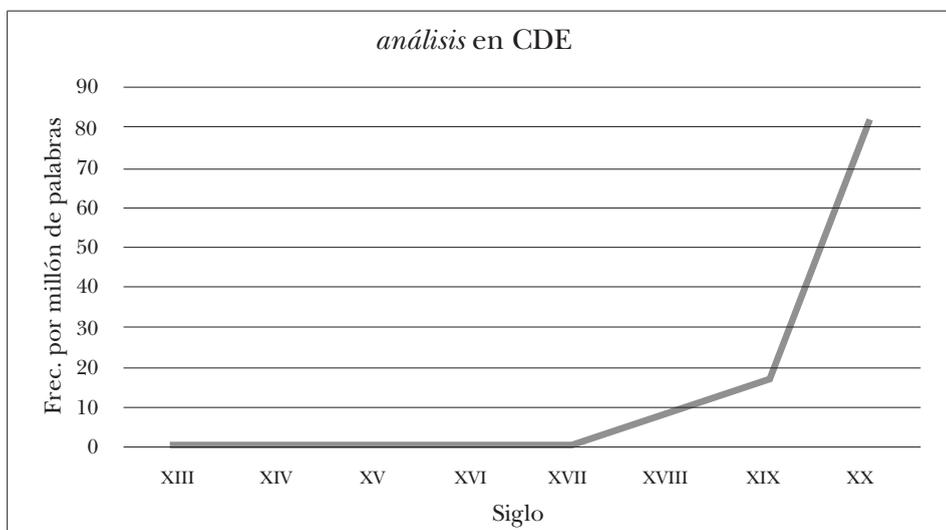
Según el escenario que plantea el ejemplo anterior, “la gente curiosa” (por lo que se puede entender la gente culta, o la que presumía de serlo) recurría a la palabra *eructar* con el fin de evitar la palabra tabú *regoldar* en aras de la cortesía (ya que la propia acción también es censurable), mientras que la gente común aún no usaba ni comprendía el significado de la palabra *eructar*. Aunque tal escenario es fácilmente imaginable (hace tiempo que la influencia del tabú se constató como uno de los factores más potentes del cambio semántico), esta evidencia secundaria que nos ofrece Cervantes con respecto a la diferenciación social de ambas palabras sirve como corroboración de la existencia de este fenómeno. Al mismo tiempo, el anterior pasaje sirve como evidencia de que la palabra *eru(c)tar* se encontraba, a comienzos del siglo XVII, integrada en el vocabulario de un grupo social culto, pero aún no era reconocida por la población general. En el español actual, *eructar* se encuentra totalmente integrada en el vocabulario activo de los hablantes, mientras que *regoldar* es hoy día una palabra obsoleta. Se trata, pues, de un caso evidente de cambio desde arriba (ing. *change from above*).

Podemos obtener información similar con respecto a la palabra *análisis*, un préstamo tardío cuyos primeros ejemplos de uso en CORDE datan de comienzos del siglo XVII. Entre ellos, el siguiente ejemplo extraído de

CORDE y documentado en Lope de Vega muestra un rechazo rotundo al uso de esta palabra:

- (29) Si cumplo con la lengua castellana,
resolución diciendo, ¿qué conceto
 es llamarla **análisis**, o a qué efeto,
tópica a la *invención*, cosa tan vana?
 Ampliar la lengua propia es cosa urbana,
 adulterarla es bárbaro defeto,
 porque su idioma y cándido dialecto
 con voces peregrinas se profana. (CORDE: Lope de Vega Carpio, *Rimas humanas y divinas del licenciado Tomé de Burguillos*, 1634)¹⁴

El rechazo que suscitaba en Lope de Vega el uso de la palabra parece confirmar que *análisis* era aún un neologismo, pese a que ya gozaba de aceptación por parte de algunos sectores de la sociedad (de lo contrario, Lope, como cualquier otro purista, no hubiera mostrado tal desacuerdo con el uso de la palabra), aunque a ojos de Lope esta se considerara una palabra indeseada por rebuscada. Esta evidencia textual confirma nuevamente que el uso de *análisis* es un cambio elitista promovido desde las capas sociales más altas; los datos estadísticos extraídos del CDEGH demuestran que la palabra *análisis* no llegó a integrarse en la lengua hasta el siglo XVIII, momento a partir del cual esta muestra un aumento drástico de la frecuencia, con un patrón similar al experimentado por el verbo *existir* (véase Gráfica D).



Gráfica D: *análisis* en el CDEGH

¹⁴ La palabra *tópica*, que Lope también menciona con reprobación, no ha sobrevivido con esta función, pese a que su uso como adjetivo, tanto en masculino como en femenino, ha experimentado una trayectoria estadística similar a la de *análisis*.

Es igualmente interesante señalar la comparación con el sustantivo *resolución*: parece ser que *análisis* se introdujo como un término técnico en el ámbito de la retórica cuyo equivalente latino era *resolutio*, tal y como lo demuestra el siguiente ejemplo:

- (30) El segundo ejercicio de la rhetorica es la **analysis**, en Latin resolutio. Es examen, y como vna anatomia de la obra hecha. (CORDE: Baltasar de Céspedes, *Del uso y ejercicio de la Retórica*, 1607)

De hecho, de los 22 ejemplos de uso de *análisis* del siglo XVII que se registran en CORDE, 19 proceden de esta misma fuente, lo que demuestra que su uso era limitado. Bartolomé Jiménez Patón, amigo de Lope, utiliza la palabra de una manera similar y en un contexto que nuevamente demuestra su uso especializado: en el siguiente ejemplo, de hecho, su uso podría considerarse más bien una cita (es decir, *la análisis que dice Aristóteles* equivale a *lo que Aristóteles llama "análisis"*).

- (31) huyase este vicio, y saquen del tema y principal pensamiento del Evangelio la **análisis** que dice Aristóteles (CORDE: Bartolomé Jiménez Patón, *Elocuencia española en arte*, 1604-1621)

Resolución, también un cultismo, tiene una trayectoria más larga en español que parece empezar a fines del siglo XV, especialmente en textos médicos (32):

- (32) Deuedes entender que en las apostemas de los miembros de dentro no se deuen poner maduratiuos: mas deuedes fazer sangria & purgacion & solucion & **resolucion**. (CORDE: Anón., *Gordonio*. BNM I315, 1495)

En la lengua actual, *resolución* y *análisis* presentan dominios semánticos diferenciados, hasta el punto de que ninguno de ellos se menciona en la definición propuesta para el otro en la edición más reciente del *DLE*.

En cuanto a otros cultismos técnicos de acuñación más reciente, encontramos, por ejemplo, que la primera documentación de *apendicitis* en CORDE procede de la novela de Blasco Ibáñez *Cañas y barro*, publicada en 1902 (33). Este ejemplo de uso es relevante no solo porque confirma la existencia de un cierto nivel de familiarización con un término médico de invención reciente y documentado en literatura creativa¹⁵, lo cual demuestra que la palabra se ha difundido más allá de la comunidad lingüística de origen, sino también porque Blasco Ibáñez describe con precisión el tipo de contexto social que pudo haber propiciado la difusión de la palabra hacia estratos sociales más bajos (a saber, el primer contacto con el término de una comunidad rural). En este texto, la primera aparición de *apendicitis* se corresponde con el diagnóstico realizado por el médico ambulante, mientras que la segunda aparición de la palabra se produce en boca de los aldeanos, que repiten las palabras del médico:

- (33) Llegó el médico de Sollana, en una de sus visitas semanales, y lo llevaron á la barraca de Sangonera. El jornalero de la ciencia movió la cabeza negativamente.

¹⁵ La primera documentación de ing. *apendicitis* en el *OED* es del *American Journal of Medical Science* de 1886.

Nada quedaba que hacer. Era una **apendicitis** mortal: la consecuencia de un abuso extraordinario que llenaba de asombro al médico. Y por el pueblo repetían lo de la **apendicitis**, recreándose las mujeres en pronunciar una palabra tan extraña para ellas. (CORDE: Vicente Blasco Ibáñez, *Cañas y barro*, 1902)

Los corpus modernos disponibles para la lengua oral también ofrecen la posibilidad de explorar las percepciones y actitudes de los hablantes con respecto a la lengua. *Genial* es un cultismo de gran interés, ya que originalmente posee un significado bastante restringido que, no obstante, en líneas generales ha seguido la misma trayectoria semántica que el sustantivo del que procede, *genio*, de manera que se ha convertido en el español actual en un término genérico para expresar aprobación y que, de hecho, tiene un abanico de aplicación bastante amplio (véase Pountain 2018). El siguiente ejemplo extraído del CDEGH muestra el rechazo que suscita el uso de *genial* como término genérico en los padres del encuestado, lo que constituye una evidencia de cambio en tiempo aparente (ing. *apparent time*):

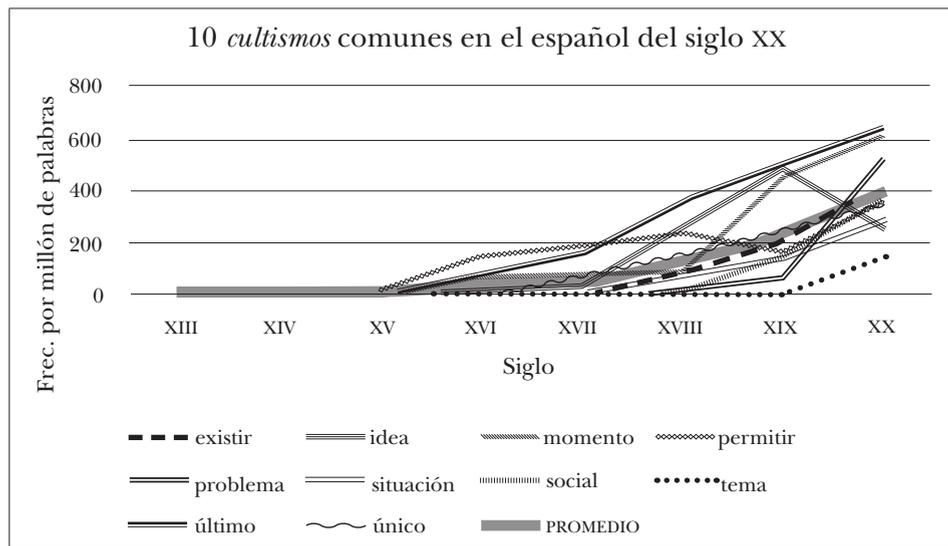
- (34) Ellos, digamos, nos corregían. Ahorita, hoy día yo hablo mucho peor de lo que hablaba cuando ellos vivían, porque se le pegan a uno las palabras aunque uno no quiera, se le pegan. No, allá en esa época, no. Que el “**genial**”...y uno decía... El “chevere” también se puso de moda, en una época. Y todo nos lo quitaban a punta de esas correcciones. No era no diga, no, sino, por ejemplo: po, po, pú. Fue tan **genial** de verdad la cosa que se nos quitó el “**genial**”, decir el “**genial**”. Eh... todas las cosas era así. Mi abuelita, por ejemplo... hay otra cosa que uno decía... no sé, hay otra... otra... otro como extranjerismo que usaba uno mucho y ella nos lo corregía también. No me acuerdo ahorita cuál es. (CDEGH: *Habla Culta: Bogotá*: M42)

6. LA DIFUSIÓN DIASTRÁTICA DE LAS VOCES CULTAS

Extraer información con respecto a la frecuencia relativa de una unidad léxica a través del CDEGH y de *Enclave* (cuyos datos son los de CORDE) es relativamente sencillo. La tendencia que se observa en las Gráficas A y B para *existir* y sus derivados resulta ser común a todos los cultismos que han logrado integrarse plenamente en la lengua. La Gráfica E muestra la trayectoria combinada de diez cultismos, incluido *existir*, que se encuentran entre las 300 palabras más frecuentes del español actual según la lista de frecuencias de Davies y Davies (2017). La lista se muestra a continuación por orden alfabético, con el rango de cada palabra entre paréntesis: *existir* (194), *idea* (247), *momento* (121), *permitir* (218), *problema* (145), *situación* (285), *social* (220), *tema* (240), *último* (188), *único* (213)¹⁶.

¹⁶ Estas palabras también figuran entre las 1.000 palabras más comunes en la totalidad del CREA (desafortunadamente, este corpus no está lematizado y, por tanto, las palabras aparecen listadas individualmente; cabe esperar que la posición en el listado de frecuencias sea superior en el caso de los lemas). Las posiciones son: *existir* (*existe* 456, *existen* 677 [y *existencia* 668]), *idea* (*idea* 414, *ideas* 765), *momento* (*momento* 131, *momentos* 712), *permitir* (*permite* 679), *problema* (*problemas* 258, *problema* 280), *situación* (*situación* 210), *social* (*social* 189, *sociales* 507 [y *socialista* 891]), *tema* (*tema* 393, *temas* 874), *último* (*último* 254, *últimos* 358, *última* 399), *único* (*único* 370, *única* 628).

Aunque se pueden observar ciertas diferencias en detalles particulares, la tendencia general es clara: el momento crítico para la difusión de estas palabras se produce a partir del siglo XVIII, con un repunte pronunciado en el siglo XX.



Gráfica E: Aumento de frecuencia de diez de los cultismos más comunes (CDEGH)

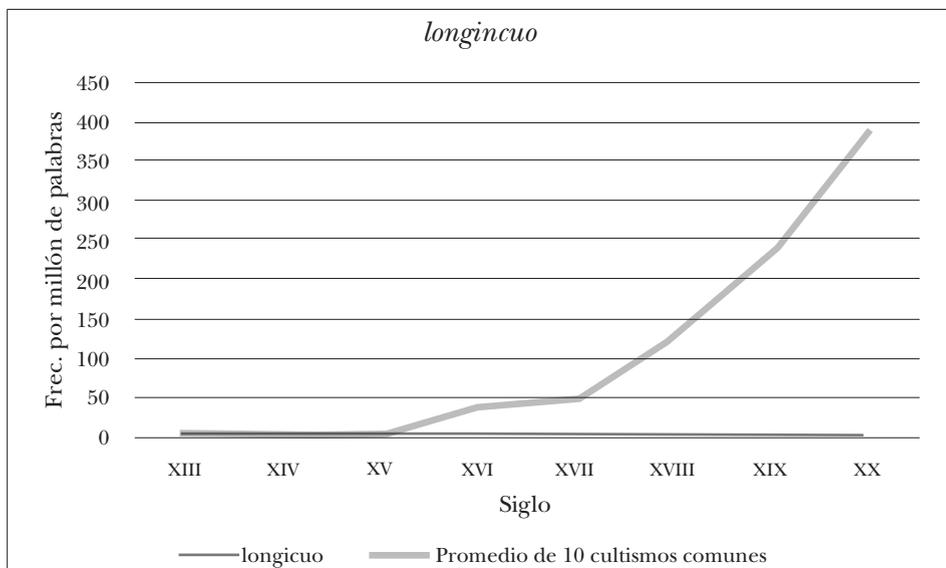
Esta tendencia contrasta con otros patrones observables en los corpus en línea con respecto a las voces cultas. La Gráfica F muestra el perfil de frecuencia en CDEGH de la palabra *longinquo*, comparado con aquel experimentado por los diez cultismos más comunes expuestos arriba. Es famoso el tono irónico en el que Cervantes emplea este cultismo:

(35) Don Quixote le dixo que no tuuiesse pena del desamparo de aquellos animales; que el que los lleuaria a ellos por tan longinquos caminos y regiones tendria cuenta de sustentarlos.

“No entiendo esso de *logicuos*”, dixo Sancho, “ni he oydo tal vocablo en todos los dias de mi vida”.

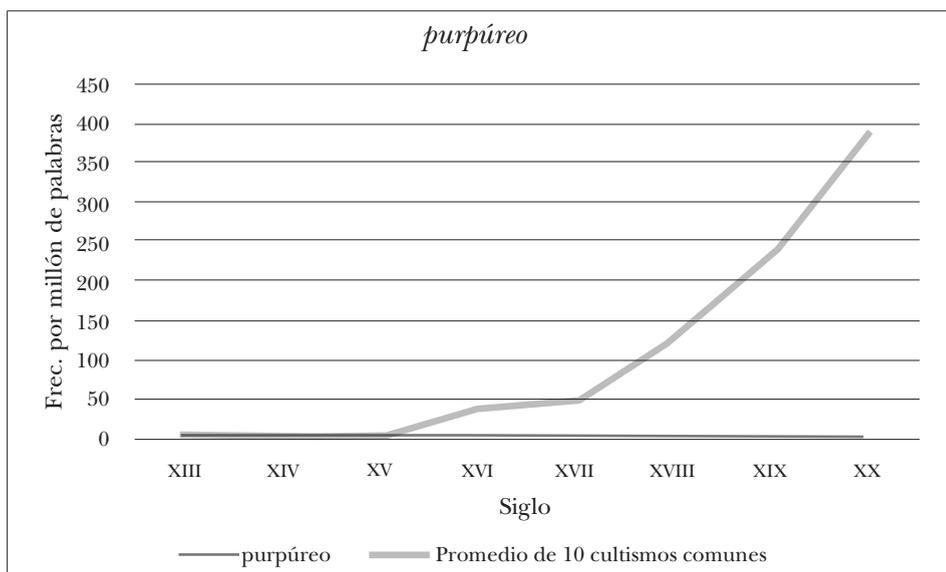
“*Longinquos*”, respondió don Quixote, “quiere dezir apartados, y no es marauilla que no lo entiendas, que no estás tu obligado a saber latin, como algunos que presumen que lo saben, y lo ignoran”. (Miguel de Cervantes, *Don Quijote de la Mancha*, II, 29; texto electrónico por Fred F. Jehle a base de la edición de Rodolfo Schevill y Adolfo Bonilla (Madrid: Gráficas Reunidas, 1928) <<https://users.ipfw.edu/jehle/cervante/dqsb1615.htm>> [30 de junio 2019]; cursiva nuestra)

Longinquo no llegó a difundirse a otras capas sociales y, de hecho, parece haber desaparecido completamente del español moderno, pese a que figura en todas las ediciones del *DLE* desde 1780 (como cabe esperar, ninguna edición registra cambios semánticos en la palabra). La frecuencia relativa de *longinquo* es extremadamente baja y en ella puede apreciarse una tendencia en declive (Gráfica F).



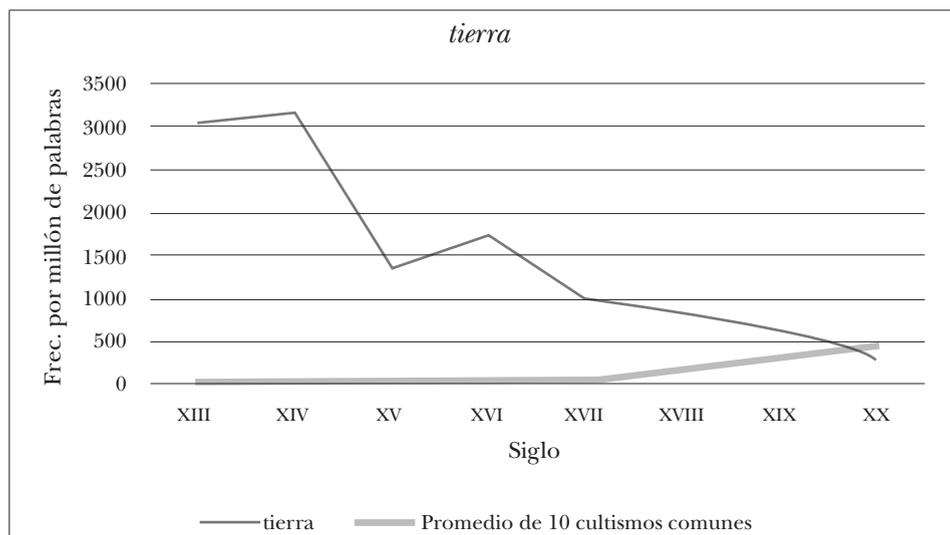
Gráfica F: Frecuencia histórica de *longinquo* (CDEGH)

La Gráfica G muestra una frecuencia estadística similar de la palabra *purpúreo*, un cultismo que, si bien no ha desaparecido por completo, tampoco se ha integrado completamente en la lengua oral: *purpúreo* muestra un declive constante, de nuevo con una frecuencia relativa extremadamente baja a lo largo de todo el eje cronológico.



Gráfica G: Frecuencia histórica de *purpúreo* (CDEGH)

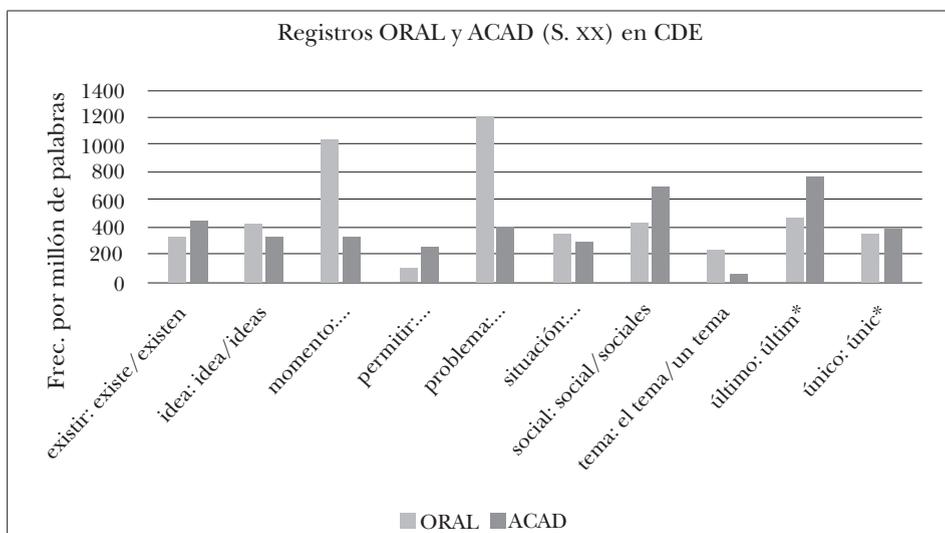
Los corpus también nos permiten establecer una comparación interesante entre las voces cultas y las palabras patrimoniales. En estas se puede apreciar un declive en su frecuencia: la gráfica H muestra el perfil estadístico de *tierra*, donde se observa que, pese a que su frecuencia global permanece alta, hay una tendencia global descendente, hasta tal punto que en el siglo XX su frecuencia se iguala a la de los cultismos más comunes. Se puede suponer que este proceso está provocado por el continuo refinamiento de la lengua, que tiene por efecto la diversificación del vocabulario, sobre todo en registros escritos, que constituyen la mayor parte de los datos de los corpus.



Gráfica H: Frecuencia histórica de *tierra* (CDEGH)

Si bien la frecuencia global que se refleja en las estadísticas obtenidas de los corpus históricos se puede considerar una evidencia de la popularización de las voces cultas, otro indicador de popularización que es necesario tener en cuenta es el de registro lingüístico. La frecuencia de una palabra en un registro técnico o literario no es necesariamente indicativa de la integración de una palabra en la lengua oral. Como es bien sabido, los datos orales en los corpus corresponden a fechas relativamente recientes. En algunas ocasiones es posible discernir lo que parecen ser representaciones de la lengua hablada en los corpus históricos, pero estos no permiten realizar una distinción rigurosa y, de hecho, el estatus de las mismas ha sido objeto de debate (cf. Oesterreicher 2005). Por lo tanto, establecer el proceso de transmisión de las voces cultas desde la lengua escrita hacia la lengua oral es una cuestión muy compleja. En la Gráfica I mostramos la comparación de las frecuencias de los diez cultismos analizados en la Gráfica E

entre los registros ORAL y ACAD[émico] del siglo XX según el CDEGH¹⁷. Lo que resulta interesante con respecto a estas estadísticas es que, aunque *a priori* pudiera esperarse que estos cultismos tuvieran mayor presencia en el registro académico que en la lengua oral, en realidad tal tendencia no se produce, sino que, de hecho, dos palabras (*momento* y *problema*) son más frecuentes en el registro oral. Esto supone una evidencia adicional con respecto al grado de integración de estas palabras en el español actual.

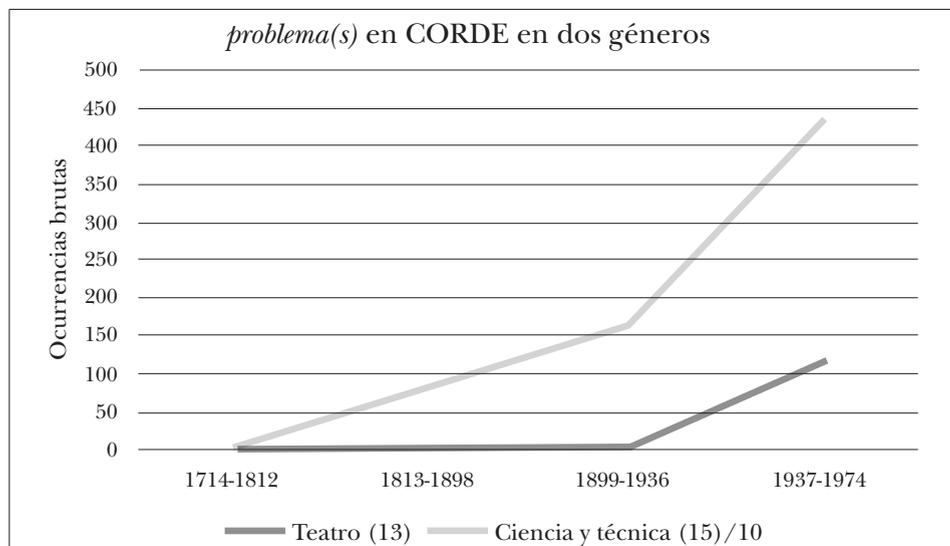


Gráfica I: Los diez cultismos más comunes en los registros oral y académico del CDEGH

Resulta todavía más complicado ofrecer una explicación histórica más exhaustiva de las estadísticas basadas en registros lingüísticos, aunque es posible extraer, con un poco de ingenio, algunas conclusiones interesantes de los corpus. La Gráfica J muestra las cifras en bruto de *problema(s)* en dos de los géneros reconocidos por CORDE: ciencia y técnica, que se puede considerar un registro escrito elitista, y teatro, que es más probable que refleje en cierta medida la lengua oral, incluso si se trata de diálogos convencionalizados¹⁸. El corpus que conforma el género ciencia y técnica se ha dividido por diez para facilitar la comparación. Lo que nos interesa en este caso no es una comparación estadística, que, además, este corpus no puede facilitar, sino una comparación del progreso de *problema* en dos géneros diferentes.

¹⁷ Se trata del único período para el que disponemos de tal información.

¹⁸ Hay que tener en cuenta que estas cifras no indican frecuencia, ya que esta es imposible de determinar a partir de la información que ofrece CORDE: la población del corpus solo se ofrece para la totalidad del período contemporáneo (1714-1974), mientras que las subdivisiones más amplias obviamente no se corresponden con las subdivisiones disponibles como criterios de búsqueda.



Gráfica J: Resultados en bruto de *problema(s)* en CORDE en dos registros lingüísticos

A la luz de estos datos, parece factible concluir que esta palabra se integró y mostró un aumento de frecuencia antes en el registro científico que en el registro dramático, y que el período crucial para su difusión en el registro dramático no comenzó hasta finales del siglo XX.

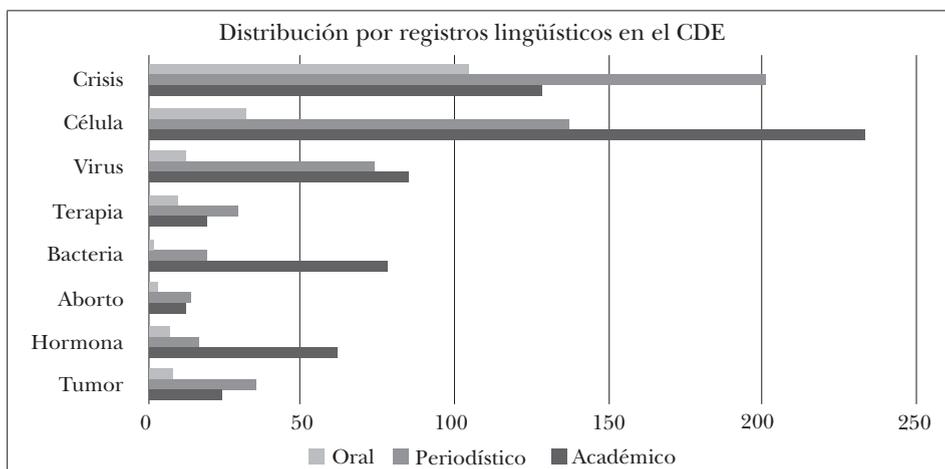
7. LA REPRESENTACIÓN DEL VOCABULARIO PRODUCTIVO Y RECEPTIVO EN LOS CORPUS LINGÜÍSTICOS

Si bien es cierto que la culminación del proceso de integración de una voz en la lengua sucede cuando esta se transfiere a la lengua oral, los estudios sobre adquisición de lenguas han constatado que el aprendizaje del vocabulario es un proceso complejo que consta de varias etapas intermedias. Es decir, los hablantes de una determinada lengua atraviesan distintos grados de familiarización con una unidad léxica hasta que esta se transfiere a su vocabulario activo. Pese a la escasez de evidencias empíricas, los estudios sobre la evaluación del vocabulario generalmente representan este proceso como un continuo cuyo punto de partida es el conocimiento nulo de una unidad léxica y cuyo término es el total entendimiento. El avance desde un extremo (que se denomina *conocimiento pasivo* o *receptivo*) hasta el otro extremo (denominado *conocimiento activo* o *productivo*) se produce a medida que los hablantes aprenden nuevos significados, colocaciones o limitaciones de registro asociadas con la palabra (Schmitt 2010: 36-37)¹⁹.

¹⁹ Hasta la fecha, existen dos planteamientos teóricos distintos con respecto a la relación entre el vocabulario pasivo y activo, denominados enfoque dimensional (ing. *dimensions approach*) y enfoque evolutivo (ing. *developmental approach*). Pese a que el segundo enfoque goza de mayor aceptación, ninguno ha sido validado hasta la fecha (Schmitt 2010 para un debate más exhaustivo sobre el tema). Ade-

El estudio estadístico de las voces cultas demuestra que los corpus históricos tienen la capacidad de ofrecer nuevos planteamientos metodológicos a las investigaciones sobre adquisición y profundidad del léxico. Dado que los datos de los principales corpus históricos del español se limitan al conjunto de palabras que ya se han utilizado, de una forma u otra, en la escritura, o, para el siglo XX, en el habla, cabría esperar que estos documentan únicamente el uso activo de la lengua, sea este escrito u oral. No obstante, queremos plantear la posibilidad de que los corpus también ofrezcan información sobre los conocimientos pasivos de los hablantes.

A este respecto resultan muy ilustrativos los ejemplos extraídos de la terminología médica. Por lo general, el léxico de la medicina hunde sus raíces en el griego y el latín²⁰, lo que convierte a esta parcela del léxico en un área muy fructuosa para el estudio de las voces cultas. En principio, el léxico de la medicina forma parte del léxico de especialidad (Arntz y Picht 1995: 27); sin embargo, es frecuente que muchas palabras que inicialmente pertenecían a un registro técnico se transfieran a la lengua oral, de ahí que el proceso de difusión del léxico de especialidad y sus relaciones con la lengua común sea una cuestión fundamental en los estudios terminológicos. Un claro ejemplo de este fenómeno son los cultismos médicos que figuran entre las cinco mil palabras más comunes del español actual, de acuerdo con los parámetros de Davies y Davies (2017). A continuación (la Gráfica K), se presenta la dispersión de una selección de estas palabras por tres de los cuatro registros lingüísticos diferenciados para el siglo XX (ORAL, PER[iodístico], ACAD[émico]) en el CDEGH:



Gráfica K: Distribución por registros lingüísticos de una selección de cultismos médicos en el CDEGH

más, estas hipótesis plantean ciertos problemas metodológicos, ya que tampoco es posible establecer el orden en el que se supone que se aprende cada elemento, así como cuántas fases se pueden identificar dentro del continuo. Como resultado, se han propuesto distintos modelos de escala (Palapanidi 2013).

²⁰ Gutiérrez Rodilla (2014: 89) afirma contundentemente que un diccionario moderno de medicina en español incluye aproximadamente un 93% de términos procedentes de las lenguas clásicas, cuyo desglose sería más o menos el siguiente: 48% griego, 35% latín y 10% híbridos greco-latinos.

En términos generales, la Gráfica K muestra una evidente variación en la frecuencia de estas voces a través de estos tres registros lingüísticos. De acuerdo con los resultados, la tendencia general muestra una tendencia al alza en los registros académicos y periodísticos frente a una frecuencia relativamente baja en la lengua oral. Esta constatación no es sorprendente, pues estos términos se originaron precisamente en el seno del discurso técnico-científico. En este sentido, merece la pena comentar el caso de *bacteria*, que muestra una muy acusada diferenciación entre el registro académico y el oral. En todos los ejemplos de cultismos “de éxito”, como son los representados en la Gráfica E, encontramos que a su aumento de frecuencia corresponde una ampliación de acepciones, tal como observamos para *existir* en 4. *Bacteria*, en cambio, según el *DLE*, no tiene sino una acepción sola, y tal vez por eso no se haya difundido más en la lengua hablada. Por el contrario, *célula*, que muestra una frecuencia mucho más alta en el registro oral, no solo tiene acepciones en otros campos técnicos como son la física (p. ej. *una célula fotoeléctrica*) sino que se ha aplicado metafóricamente para denominar a un pequeño grupo político o religioso. *Crisis*, la palabra más frecuente de las estudiadas aquí, ha alcanzado una frecuencia impresionante en el registro oral en comparación con el académico, sin duda por ya no estar limitado al campo médico.

La influencia de la prensa y los medios de masas sobre la lengua es un fenómeno que ya ha sido ampliamente debatido. Desde finales del siglo XX, los expertos afirman que los medios de comunicación son responsables de gran parte del lenguaje que se escucha en la calle (Bell 1997[1991]: 1-3). Este fenómeno se ha vuelto más acusado en las últimas décadas, debido, sobre todo, a la popularización y continuo desarrollo de los medios de comunicación, a un mayor acceso a la enseñanza y “a las necesidades del ciudadano de a pie de entablar contacto con ciertos ámbitos científicos” (Martín Camacho 2004: 39; Pérez Pascual 2012: 207-208). Se puede pensar que la influencia de los medios de comunicación sobre la lengua contribuya principalmente al aumento del vocabulario pasivo, bien sea porque la información lingüística que se ofrece de estas unidades léxicas es usualmente parcial (por ejemplo, porque sea frecuente el uso de una única acepción en detrimento de otras), bien sea porque se trata de un término que hace referencia a un aspecto alejado de los intereses generales de la población, como puede suceder con términos menos comunes como *espectrograma* (Martín Camacho 2004: 39-40). Con respecto a la Gráfica K, la palabra *crisis* es ilustrativa del primer factor, ya que esta se utiliza frecuentemente en prensa para cuestiones de índole política y económica, mientras que su acepción médica es más inusual. De hecho, la alta frecuencia de esta palabra en prensa puede deberse precisamente a que se trata de una voz plenamente integrada en el ámbito de la política, que a su vez es uno de los temas que más cobertura recibe por parte de los medios de comunicación. De la misma manera, la confusión que generan las palabras *bacteria* y *virus*, utilizadas en ocasiones de manera indistinta y a menudo con preferencia

por la segunda, puede interpretarse como una señal del desconocimiento total o parcial de los hablantes respecto a sus significados médicos, como demuestra el siguiente texto, valioso por la información secundaria que proporciona sobre estas palabras (véase 5), del propio CDEWD:

Es uno de nuestros errores más comunes: ir a ver al médico y pedirle antibióticos a la primera señal de un resfrío. “Muchos, por falta de información, piensan que las bacterias causan los resfríos, y pretenden medicarse con antibióticos; sin embargo, debemos recordarles que el 100% de los resfríos son producidos por un virus, y que las bacterias no hacen nada”, comenta el doctor Acosta Altamirano (CDEWD: <<https://www.selecciones.com.ar/salud/alimentacion/estas-resfriado>>, Argentina, 2011)

Al mismo tiempo, este desconocimiento o falta de precisión da lugar al fenómeno llamado “determinologización” (Arntz y Picht 1995: 40), fenómeno este usual en términos técnicos y científicos divulgados por los medios de comunicación, lo que deja suponer que estos vocablos, al empezar a penetrar en el habla cotidiana, vayan rebasando sus acepciones puramente técnicas.

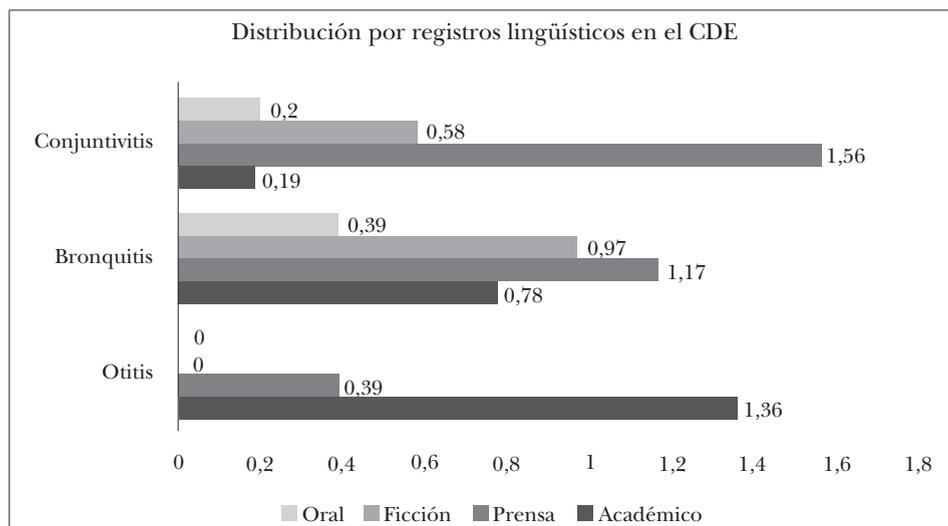
En suma, los datos con respecto a la distribución por registros de las voces cultas que nos ofrece el CDEGH pueden servir como fuente de información para la evaluación de lo que se puede denominar la “profundidad” léxica de determinados grupos de palabras. Al mismo tiempo, el estudio del conocimiento pasivo del vocabulario se posiciona como un elemento clave en el análisis de la integración de los cultismos, pues, en líneas generales, puede decirse que el proceso de difusión que experimentan estas voces en la lengua general se inicia en la fase de recepción por parte de los hablantes y culmina en la fase de producción, es decir, su plena integración a la lengua oral.

Es igualmente preciso adoptar una perspectiva crítica a la hora de evaluar el contenido de carácter productivo en los corpus históricos. Hasta ahora se ha asumido que las listas de frecuencia utilizadas en los estudios sobre integración del léxico culto parten de un corpus de palabras que constituya un fiel reflejo de la lengua oral, esto es, del vocabulario más común de una lengua. No obstante, desde mediados del siglo XX son comunes las críticas con respecto a las deficiencias de las listas de frecuencia, a pesar del continuo refinamiento de las metodologías. En el marco de las investigaciones sobre disponibilidad léxica, las investigaciones en torno a la composición del *Français Élémentaire* a mediados del pasado siglo demostraron la ausencia en las listas de frecuencias léxicas de términos tan cotidianos como *autobús* (López Morales 1995-96: 245-259). Pese a que los métodos de elaboración de los diccionarios de frecuencias modernos son más innovadores, en listas recientes basadas en corpus²¹, como la elaborada por Davies y Davies (2017), aún se percibe la ausencia de palabras cotidianas como *sartén* o *lavadora*. La falta de representación de palabras tan comunes no resulta sorprendente, pues estos diccionarios se componen, no solamente de las palabras más habituales, sino también de las que “aparecen

²¹ El diccionario de frecuencias del español de Davies (2017) utiliza los datos de CDEGH y CDEWD.

distribuidas por todos los tipos de textos y situaciones de comunicación” (Alvar Ezquerro 2005: 22).

Las implicaciones que estos resultados tienen en la representación del vocabulario activo en los corpus lingüísticos plantean cuestiones de interés filológico con respecto a la propia naturaleza de la lengua y a la importancia de la situación comunicativa. Si tomamos como objeto de estudio los cultismos médicos, cabe pensar en varios términos que en español resultan relativamente frecuentes pero cuyo uso se circunscribe a contextos comunicativos muy concretos, como sucede con los nombres de algunas de las enfermedades infantiles más comunes (Gráfica L).



Gráfica L: Frecuencia en corpus de una selección de cultismos médicos

La baja frecuencia en los corpus históricos de palabras como *conjuntivitis*, *bronquitis* y *otitis* demuestra que estas voces, de aplicación restringida, apenas tienen representación en ellos. Sin embargo, estas palabras no pueden descartarse como parte del vocabulario activo de determinados grupos sociales, y en realidad son muy comunes en círculos que comparten intereses o necesidades muy concretas (basta consultar cualquier revista sobre maternidad y paternidad). A este respecto sugiere Pérez Pascual (2012: 211) que un criterio complementario al de frecuencia y dispersión en la elaboración de los corpus lingüísticos podría ser la inclusión de material procedente de fuentes alternativas, como pueden ser los programas divulgativos o los manuales de enseñanza de las escuelas. Con este criterio, quizá, se lograría una mejor correspondencia entre la frecuencia reflejada en los corpus y la representación del vocabulario productivo de la lengua.

8. CONCLUSIONES

Los datos recuperables de los corpus nos permiten trazar la trayectoria histórica de los cultismos desde su fecha de primera documentación hasta su posterior difusión u olvido en la lengua. En ocasiones, nos dejan entrever el mecanismo de difusión y las actitudes lingüísticas de los hablantes frente al uso de las palabras tomadas en préstamo y a los cambios semánticos. De la misma manera, la información que los corpus son capaces de proveer con respecto a los géneros textuales permite en ocasiones cuantificar el proceso de difusión del léxico, al mismo tiempo que provee de una posible base para la investigación de los mecanismos de adquisición de vocabulario en el continuo receptivo-productivo. Sin embargo, por útil que resulte a nuestros propósitos la investigación basada en corpus, esta solo puede desarrollarse mediante la mediación filológica de los datos en bruto que se obtienen en los resultados de búsqueda. La primera aparición de un cultismo no quiere decir necesariamente que este esté plenamente integrado en la lengua, ya que el contexto lingüístico puede dejar claro que todavía se trata de una palabra desconocida que necesita explicación; en los cálculos de frecuencia siempre hay que tener en cuenta la dispersión, ya que puede ser que la palabra en cuestión se vea favorecida por un autor determinado. También hay que reconocer las inevitables limitaciones de los datos, sobre todo la dificultad práctica de obtener datos sobre registros lingüísticos a través de la clasificación de fuentes según su género textual. Dicho esto, queda claro que estas herramientas de investigación son capaces de arrojar luz sobre el fascinante proceso diacrónico que es la incorporación de determinados cultismos al habla cotidiana.

CORPUS Y DICCIONARIOS

ACCADEMIA DELLA CRUSCA (1863-1923): *Vocabolario degli accademici della Crusca*, 5ª ed. <<http://www.lessicografia.it>> [2 de diciembre 2016].

BDH = Biblioteca Nacional de España: *Biblioteca Digital Hispánica* [en línea]. <<http://www.bne.es/es/Catalogos/BibliotecaDigitalHispanica/Inicio/index.html>> [3 de septiembre 2019].

CDE = DAVIES, MARK (2002-): Corpus del Español <<http://www.corpusdelespanol.org>> [28 de junio 2019]. Se trata de una colección de corpus, de los que nos referimos aquí a dos, el Género/Histórico (CDEGH) y el Web/Dialectos (CDEWD).

CODÉA = SÁNCHEZ-PRIETO BORJA, PEDRO (coord.): *Corpus de Documentos Españoles Anteriores a 1700* [en línea]. <<http://demos.bitext.com/codea>> [3 de septiembre 2019].

CORDE = REAL ACADEMIA ESPAÑOLA: Banco de datos (CORDE) [en línea]. *Corpus diacrónico del español*. <<http://www.rae.es>> [15 de mayo 2018].

CORDEMÁFORO = RODRÍGUEZ MOLINA, JAVIER y OCTAVIO DE TOLEDO Y HUERTA, ÁLVARO (2017): “Acceso a la base de datos CORDEMÁFORO”, *Scriptum Digital*, 6. <<http://www.scriptumdigital.org/documents/Octavio-Molina-Base-de-datos-Scriptum.xlsx>> [3 de septiembre 2019].

CORDIAM = BERTOLOTTI, VIRGINIA y CONCEPCIÓN COMPANY COMPANY: *Corpus Diacrónico y Diatópico del Español de América* [en línea]. <<http://www.cordiam.org>> [3 de septiembre 2019].

COROMINAS, JOAN y JOSÉ A. PASCUAL (1980-91): *Diccionario crítico etimológico castellano e hispánico*, Madrid: Gredos.

CREA = REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [en línea]. *Corpus de referencia del español actual*. <<http://www.rae.es>> [15 de mayo 2018].

DAVIES, MARK y KATHY HAYWARD DAVIES (2017): *A Frequency Dictionary of Spanish. Core Vocabulary for Learners*, Londres: Routledge.

DICCA-XV = Grup d'Història i Contacte de Llengües (GHCL) y COLOMA LEAL (dir.) (2013): *Diccionario del castellano del siglo XV en la Corona de Aragón*. <<http://stel.ub.edu/diccxv>> [3 de septiembre 2019].

DLE = ASOCIACIÓN DE ACADEMIAS DE LA LENGUA ESPAÑOLA (2014-): *Diccionario de la lengua española*, 23ª ed., Madrid: Espasa, y <<http://www.rae.es/obras-academicas/diccionarios/diccionario-de-la-lengua-espanola>> [28 de junio 2019].

Enclave = REAL ACADEMIA ESPAÑOLA: *Enclave* <<https://enclave.rae.es>> [15 de mayo 2018].

FRANTEXT = Base textuelle FRANTEXT, ATILF-CNRS & Université de Lorraine <<http://www.frantext.fr>> [21 de noviembre 2016].

OED = *Oxford English Dictionary* <<http://www.oed.com>> [21 de noviembre 2016].

OVI = Le Banche dati dell' Opera del Vocabolario Italiano Antico <www.ovi.cnr.it> [22 de noviembre 2016].

TLF = *Trésor de la Langue Française informatisée* (2004), Paris: CNRS Éditions y <<http://atilf.atilf.fr>> [4 de octubre 2016].

REFERENCIAS BIBLIOGRÁFICAS

ALVAR EZQUERRA, MANUEL (2005): "La frecuencia léxica y su utilidad en la enseñanza del español como lengua extranjera", en María Auxiliadora Castillo Carballo *et al.* (coords.), *Las gramáticas y los diccionarios en la enseñanza del español como segunda lengua: deseo y realidad. Actas del XV Congreso Internacional de ASELE, Sevilla 22-25 de septiembre de 2004*, Sevilla: Universidad de Sevilla, 19-39.

ARNTZ, REINER y HERIBERT PICHT (1995 [1982]): *Introducción a la terminología*, Madrid: Fundación Sánchez Ruipérez-Pirámide.

AZOFRA SIERRA, M^a ELENA (2006). "Consideraciones sobre el concepto de cultismo", *Revista de Filología Románica*, 23, 229-240.

BELL, ALLAN (1991): *The Language of News Media*, Oxford: Blackwell.

BLOOMFIELD, LEONARD (1944): "Secondary and Tertiary Responses to Language", *Language*, 20, 45-55.

BUSTOS TOVAR, JOSÉ JESÚS (1974): *Contribución al estudio del cultismo léxico medieval*, Madrid: Real Academia Española.

CHENG, WINNIE (2012): *Exploring Corpus Linguistics: Language in action*, Londres: Routledge.

CLAVERÍA NADAL, GLORIA (1991): *El latinismo en español*, Barcelona: Departament de Filologia Espanyola, Facultat de Lletres, Universitat Autònoma de Barcelona.

DWORKIN, STEVEN N. (2010): "Thoughts on the Re-Latinization of the Spanish Lexicon", *Romance Philology*, 64, 173-184.

GARCÍA VALLE, ADELA (1992): "Otra vez sobre los conceptos de 'latinismo', 'cultismo' y 'semicultismo' a la luz de los nuevos datos", *Anuario de Estudios Filológicos*, 15, 81-96.

— (1998): *La variación nominal en los orígenes del español*, Madrid: Consejo Superior de Investigaciones Científicas.

GUTIÉRREZ RODILLA, BERTHA MARTÍA (2014): "El lenguaje de la medicina en español: cómo hemos llegado hasta aquí y qué futuro nos espera", *Panace@*, XV (39), 86-94.

HUNSTON, SUSAN (2002): *Corpora in Applied Linguistics*, Cambridge: Cambridge University Press.

KABATEK, JOHANNES (2016): "Un nuevo capítulo en la lingüística histórica iberorrománica: el trabajo crítico con los corpus. Introducción a este volumen", en Johannes Kabatek (ed.), *Lingüística de corpus y lingüística histórica iberorrománica*, Berlín/Boston: De Gruyter, 1-17.

LLEAL GALCERAN, COLOMA (2013): "Rigor metodológico e investigación filológica", *Scriptum Digital*, 2, 107-121.

LÓPEZ MORALES, HUMBERTO (1995-96): "Los estudios de disponibilidad léxica: pasado y presente", *Boletín de Filología*, 35, 245-259.

MALKIEL, YAKOV (1957): "Préstamos y cultismos", *Revue de Linguistique Romane*, 21, 1-61.

MARTÍN CAMACHO, JOSÉ CARLOS (2004): *El vocabulario del discurso tecnocientífico*, Madrid: Arco/Libros.

NGLE = REAL ACADEMIA ESPAÑOLA / ASOCIACIÓN DE LAS ACADEMIAS DE LA LENGUA ESPAÑOLA (2009): *Nueva gramática de la lengua española. Morfología / Sintaxis I*, Madrid: Espasa.

OESTERREICHER, WULF (2005): "Textos entre inmediatez y distancia comunicativas: El problema de lo hablado escrito en el Siglo de Oro", en Rafael Cano Aguilar (ed.), *Historia de la lengua española*, Barcelona: Ariel, 729-769.

PALAPANIDI, KIRIAKI (2013): "Conocimiento receptivo y productivo del vocabulario de LE", *Revista Nebrija de Lingüística Aplicada*, 7(6) [14] <<http://www.nebrija.com/revista-linguistica/conocimiento-receptivo-y-productivo-del-vocabulario-de-le.html>> [5 June 2018].

PÉREZ PASCUAL, JOSÉ IGNACIO (2012): "El léxico de especialidad", en Luis Luque Toro *et al.* (eds.), *Léxico español actual III*, Venecia: Università Ca'Foscari di Venezia, 189-219.

POUNTAIN, CHRISTOPHER J. (2011): "Putting Philology back into Linguistics", *Hispanic Research Journal*, 12, 99-117.

— (2016): "Learned borrowing as a transnational phenomenon", ponencia leída en "Transnational Modern Languages", Italian Cultural Institute, Londres, 2-3 de diciembre, 2016.

— (2018): "Los cultismos de cada día", ponencia leída en XI Congreso Internacional de Historia de la Lengua Española, Lima, 6-10 de agosto, 2018.

PRESTON, DENNIS R. (1999): *Handbook of Perceptual Dialectology*, Amsterdam / Philadelphia: Benjamins.

RODRÍGUEZ MOLINA, JAVIER y OCTAVIO DE TOLEDO y HUERTA, ÁLVARO (2017): "La imprescindible distinción entre texto y testimonio: el CORDE y los criterios de fiabilidad lingüística", *Scriptum Digital*, 6, 5-68.

ROJO SÁNCHEZ, GUILLERMO (2010): “Sobre codificación y explotación de corpus textuales: otra comparación del *Corpus del Español* con el *CORDE* y el *CREA*”, *Lingüística*, 24, 11-50.

SCHMITT, NORBERT (2010): *Researching Vocabulary: A Vocabulary Research Manual*, Basingstoke: Palgrave Macmillan.

STOCKWELL, ROBERT P., J. DONALD BOWEN y JOHN W. MARTIN (1965): *The Grammatical Structures of English and Spanish*, Chicago/Londres: University of Chicago Press.

TRASK, R.L. (1996): *Historical Linguistics*, Londres: Arnold.

WEINREICH, URIEL, WILLIAM LABOV y MARVIN I. HERZOG (1968): “Empirical Foundations for a Theory of Language Change”, en Winfred P. Lehmann y Yakov Malkiel (eds.), *Directions for Historical Linguistics: A Symposium*, Austin: University of Texas Press, 95-189.

WHITLAM, JOHN (2011): *Modern Brazilian Portuguese Grammar. A Practical Guide*, Londres / Nueva York: Routledge.

WONDER, J. P. (1979): “Expresiones locativas modificadoras de sustantivo”, *Hispania*, 62, 679-84.

WRIGHT, ROGER (1976): “Semicultismo”, *Archivum Linguisticum*, 7, 13-28.

— (1982): *Late Latin and Early Romance in Spain and Carolingian France*, Liverpool: Francis Cairns.