



*RILEX*

*REVISTA SOBRE INVESTIGACIONES LÉXICAS*

VOLUMEN II - NÚMERO 2

JUNIO, 2019

Sven Tarp

Ana Salgado

Rute Costa

María Pilar Valero Fernández

Ivana Lončar

Kelly Linares

José Armando San Martín Gómez

REVISTAS CIENTÍFICAS DE LA UNIVERSIDAD DE JAÉN

Los estudio e investigaciones que se recogen en esta revista están sujetos a una licencia de reconocimiento de *Creative Commons*. Esta licencia permite **compartir** (copiar y redistribuir el material en cualquier medio o formato) y **adaptar** (remezclar, transformar y construir a partir del material para cualquier propósito, incluso comercialmente) el material siempre que se indique adecuadamente el origen y los cambios

© 2019

Universidad de Jaén

*RILEX. Revista sobre investigaciones léxicas*

ISSNe: 2605-3136

**EQUIPO EDITORIAL**

***DIRECCIÓN***

Dra. M.<sup>a</sup> Águeda Moreno Moreno (Universidad de Jaén)

***SUBDIRECCIÓN Y SECRETARÍA***

Dra. Marta Torres Martínez (Universidad de Jaén)

***EDITOR DE SECCIÓN: LEXICOLOGÍA***

Dra. M.<sup>a</sup> Águeda Moreno Moreno (Universidad de Jaén)

***EDITOR DE SECCIÓN: LEXICOGRAFÍA***

Dr. Jesús Camacho Niño (Universidad de Jaén)

***EDITOR DE SECCIÓN: ENSEÑANZA DEL ESPAÑOL***

Dr. Narciso M. Contreras Izquierdo (Universidad de Jaén)

***COEDITOR***

Dra. Elisabeth Fernández Martín (Universidad de Jaén)

***ASISTENCIA TÉCNICA***

Alicia Arjonilla Sampedro (Universidad de Jaén)

Inmaculada Ruiz Sánchez (Universidad de Jaén)

## **COMITÉ CIENTÍFICO**

Dr. Ángel López García-Molins, Universidad de Valencia, España

Dr. Cecilio Garriga Escribano, Universidad Autónoma de Barcelona, España

Dra. Concepción Maldonado González, Universidad Complutense de Madrid, Ediciones SM, España

Dra. Dolores Azorín Fernández, Universidad de Alicante, España

Dr. Giuseppe Trovato, Universidad de Venecia, Italia

Dra. Gloria Clavería Nadal, Universidad Autónoma de Barcelona, España

Dr. Humberto Hernández Hernández, Universidad de La Laguna, España

Dra. Josefina Prado Aragonés, Universidad de Huelva, España

Dr. José Ignacio Pérez Pascual, Universidad de A Coruña, España

Dr. José Ramón Carriazo Ruiz, Universidad Nacional del Educación a Distancia, España

Dra. Mar Campos Souto, Universidad de Santiago de Compostela, España

Dra. Mar Cruz Piñol, Universidad de Barcelona, España

Dra. M.<sup>a</sup> Luisa Calero Vaquera, Universidad de Córdoba, España

Dra. Marta Higuera García, Instituto Cervantes, España

Dr. Pedro Fuertes-Olivera, Universidad de Valladolid, España

Dr. Stefan Ruhstaller, Universidad Pablo de Olavide, España

Dr. Sven Tarp, Universidad de Aarhus, Dinamarca

## ÍNDICE

---

### **Sven Tarp**

La ventana al futuro: Despidiéndose de los diccionarios para abrazar a la Lexicografía ..... 5

### **Ana Salgado y Rute Costa**

Marcas temáticas en los diccionarios académicos ibéricos: estudio comparativo ..... 37

### **María Pilar Valero e Ivana Lončar**

Nivelación del artículo lexicográfico en un diccionario monolingüe de aprendizaje: el caso de las locuciones adverbiales ..... 64

### **Kelly Linares**

Tratamiento de las definiciones de los conectores discursivos en la lexicografía didáctica cubana. El caso del *Diccionario básico escolar* ..... 95

### **José Armando San Martín Gómez**

Aproximación al indigenismo léxico en el español de El Salvador (siglos XVII-XVIII) ..... 121



## LA VENTANA AL FUTURO: DESPIDIÉNDOSE DE LOS DICCIONARIOS PARA ABRAZAR LA LEXICOGRAFÍA

### THE WINDOW TO THE FUTURE: GOODBYE TO DICTIONARIES AND HELLO TO LEXICOGRAPHY

Sven Tarp  
*Universidad de Aarhus*  
[st@cc.au.dk](mailto:st@cc.au.dk)

#### RESUMEN

Esta contribución discute la situación actual de la lexicografía describiéndola como una crisis de identidad que se debe a un modelo de negocios obsoleto. En su lugar, propone un nuevo modelo de financiación que se apoya en las nuevas tecnologías y una visión diferente de la misma esencia de la lexicografía como disciplina. La visión es subir en la cadena de valor integrando los datos lexicográficos en herramientas, servicios y plataformas de alta tecnología. A modo de ejemplo, el artículo discute la materialización de esta idea en un asistente de escritura publicado en 2019 que todavía se encuentra en plena fase de desarrollo. Al final, y para comprobar el uso y funcionamiento de esta herramienta, se recomienda un método diferente a los que tradicionalmente se aplican para estudiar el uso de diccionarios.

**Palabras clave:** origen de la lexicografía, diccionarios, Cuarta Revolución Industrial, innovación disruptiva, modelo de negocios, cadena de valor, datos lexicográficos, asistentes de escritura, estudios sobre el uso, desarrollo guiado por pruebas

#### ABSTRACT

This contribution discusses the current situation within lexicography and describes it as an identity crisis due to an obsolete business model. Instead, it proposes a new financial model based on the new technologies and a different understanding of the very essence of lexicography as a discipline. The idea is to go upstream in the value chain by integrating the lexicographical data into high-tech tools, services and platforms. As a matter of example, the article then discusses the materialization of this idea in a writing assistant published in 2019 and under continuous improvement. Finally, a method different from those traditionally applied to study the use of dictionaries, is recommended in order to test the use and functionality of this tool.

**Keywords:** birth of lexicography, dictionaries, Fourth Industrial Revolution, disruptive innovation, business model, value chain, lexicographical data, writing assistants, user research, test-driven development

## 0. INTRODUCCIÓN

En un artículo reciente, Maldonado (2019) comenta el creciente número de estudios académicos sobre el uso de diccionarios. Esta iniciativa es, según ella, “loable y necesaria” aunque “no deja de sorprender en el mundo de la lexicografía empresarial” donde no se aborda ningún trabajo sin “un estudio previo” que avala “tanto la viabilidad económica del proyecto como el acierto a la hora de cubrir las necesidades de los clientes potenciales” (Maldonado, 2019, p. 103). La autora acierta, quizás más de lo que ella piensa. En la literatura académica sobre lexicografía escasean las contribuciones que, de una u otra forma, abordan el aspecto económico. Las reflexiones de Rundell (2007), Winestock & Jeong (2014), Nesi (2015), Simonsen (2017), Fuertes-Olivera (2019) y algún que otro autor son excepciones dentro de una disciplina que en gran medida hace caso omiso del mismo modelo de negocios que la sostiene.

Esta omisión quizás podría justificarse hace dos o tres décadas cuando todavía vivíamos en el mundo feliz de los diccionarios impresos con una gran clientela dispuesta a pagar por ellos. Hoy en día, la situación ha cambiado por completo. La introducción de nuevas tecnologías disruptivas ha parado casi por completo la producción y comercialización de diccionarios impresos en Europa y América del Norte. Las ventas que algunas casas lexicográficas siguen registrando en África, América Latina y parte de Asia parecen más bien una estrategia de salida diseñada para exprimir las últimas gotas de una aventura que ya no tiene futuro.

La mayoría de los usuarios de los nuevos diccionarios en línea esperan que su uso sea gratuito. Como consecuencia, las casas editoriales están luchando por sobrevivir y muchas de ellas ya han abandonado la producción y comercialización de los diccionarios. En Europa occidental se calcula que el número de casas lexicográficas que ganan dinero con sus diccionarios en línea puede contarse con los dedos de una sola mano. No hace falta ser un Einstein para comprender que el modelo de negocios usado hasta ahora ha quedado obsoleto. Este problema debiera despertar mucho más interés entre los lexicógrafos universitarios, pues es difícil imaginarse una solución que sea puramente económica y que no cuente con una activa colaboración

interdisciplinaria de lexicógrafos, informáticos y otras especialistas relevantes. En este contexto se puede interpretar el creciente interés por estudiar el uso de diccionarios, sean estos existentes o prototipos, como una técnica de avestruz que sirve para ocultar una sorprendente falta de visiones y responsabilidad académica.

A continuación, discutiremos este problema y esbozaremos un nuevo modelo de negocios que corresponde a las tecnologías que actualmente configuran la llamada Cuarta Revolución Industrial, un término acuñado por Schwab (2015: s.p.), según el cual este fenómeno “se caracteriza por una fusión de tecnologías que está borrando las líneas entre las esferas física, digital y biológica”. A este respecto, defenderemos la idea de que la lexicografía solo puede renacer y sobrevivir si se integra en herramientas de alta tecnología desarrolladas mediante la mencionada colaboración interdisciplinaria. A modo de ejemplo, mostraremos cómo esta idea puede materializarse en un asistente de escritura diseñado para hispanohablantes que requieren ayuda para escribir en inglés.

## 1. CRISIS DE IDENTIDAD

No cabe duda de que la lexicografía se encuentra en su más profunda crisis desde que viera la luz hace más de cuatro mil años. La debacle de las ventas afecta seriamente la producción de nuevos diccionarios. En España, si se excluye la 23.<sup>a</sup> edición del *DLE*, no se ha publicado ningún gran diccionario general desde hace más de una década, sea este en papel o formato digital. Simultáneamente, una fauna de proyectos universitarios, a menudo basados en principios poco convincentes, se han desarrollado con financiación pública o comunitaria. Mucho ruido y pocas nueces. Cuando la financiación termina después de 2-3 años, estos proyectos suelen defraudar con solo un puñado de artículos acabados (los llamados prototipos) que rápidamente pasan al olvido. Ninguna editorial ha mostrado interés en continuar con los experimentos. Los estudios que se hayan hecho sobre su viabilidad económica no habrán sido muy alentadores. El mismo fenómeno se reproduce en cada vez más países, especialmente los tecnológicamente más desarrollados.

De todo esto se desprende que el tradicional modelo de negocios constituye el detonante visible de la crisis que actualmente sufre la lexicografía. Sin embargo, la causa

en última instancia de esta crisis no es el mencionado modelo sino el cambio de paradigma tecnológico. La relación entre lexicografía y tecnología ha sido el sujeto de varias reflexiones, entre ellas las de De Schryver (2003), Hanks (2010, 2013), Rundell & Kilgarriff (2011), Rundell (2012), Nielsen (2013) y Fuertes-Olivera *et al.* (2018). A lo largo de la historia, la tecnología ha influido enormemente en el desarrollo de las cinco fases principales del proceso de compilación de los productos lexicógrafos:

- 1) *determinación* de los tipos de datos que se ofrecerán a los futuros usuarios;
- 2) *recuperación* de datos brutos a partir de las fuentes empíricas;
- 3) *preparación* de los datos lexicográficos;
- 4) *almacenamiento* de los datos lexicográficos; y
- 5) *presentación* de los datos lexicográficos en el producto final.

Cada una de estas fases ha conocido cambios sísmicos y mejoras considerables cuando se han introducido nuevas tecnologías en diferentes momentos de la historia. La forma particular en que nació la lexicografía europea es inconcebible sin la invención de la pluma y el pergamino (McArthur, 1986). El crecimiento exponencial de la producción de diccionarios durante los últimos cinco siglos no habría sido posible sin el desarrollo continuo de las tecnologías de la imprenta y la encuadernación (Hanks, 2010, 2013). La cada vez mejor calidad de los datos lexicográficos que ha caracterizado la disciplina durante los últimos cincuenta años es inseparable de la introducción de los corpus digitales y del internet como fuentes empíricas (Hanks, 2012; Tarp & Fuertes-Olivera, 2016, 2018). Los actuales métodos de trabajo de los lexicógrafos son inimaginables sin los ordenadores y bases de datos, etc.

Se puede constatar que la introducción de nuevas tecnologías en la lexicografía implica grandes cambios en el medio de presentación de los diccionarios, el diseño y contenido de los artículos, la distribución y uso de los diccionarios, los temas tratados, la cantidad de usuarios, el papel social del lexicógrafo y el modelo de financiación. Es decir, históricamente las tecnologías disruptivas entrañan una revolución casi total de la disciplina. El presente no es menos dramático. La creciente aplicación de las nuevas tecnologías disruptivas tiene grandes consecuencias para la lexicografía como práctica cultural y afecta la disciplina en las cinco dimensiones mencionadas más arriba.

Un nuevo modelo de negocios solo puede nacer de una profunda reflexión que rompa con el paradigma tradicional y busque nuevas soluciones a los problemas tanto viejos como nuevos. Esto implica que un nuevo empuje lexicográfico debe basarse en un profundo conocimiento de lo que realmente está pasando ante nuestros ojos, aunque a menudo no lo veamos. Pues bien, si volvemos al tema que nos interesa en este artículo, o sea, la presentación del producto lexicográfico al usuario destino, podemos observar cuatro grandes tendencias que actualmente se están desarrollando en el marco de la Cuarta Revolución Industrial, o sea:

1. La transformación del tradicional *diccionario impreso* en *diccionario digital*, especialmente el de Internet, proceso que todavía no ha culminado ya que muchos diccionarios digitales siguen con marcadas características de su antepasado impreso (Fuertes-Olivera & Tarp, 2014).
2. La transformación del *diccionario estandarizado*, que ha caracterizado la disciplina desde la introducción de la tecnología de la imprenta, en *diccionario personalizado* que se adapta cada vez más a las necesidades de su usuario en cada consulta concreta (Rundell 2010 & Tarp, 2011).
3. La transformación del *diccionario solitario* de consulta en *diccionario integrado* en otras herramientas, p.ej. asistentes de escritura (Tarp *et al.*, 2017, 2018).
4. Finalmente, se observa una cierta sustitución del *diccionario como tal* por *datos lexicográficos* que se comercializan y usan para diferentes fines sin adquirir la forma de un diccionario (Maldonado, 2019; Tarp, 2019).

Estas grandes transformaciones que se entrelazan y todavía están en plena fase de desarrollo requieren una profunda reflexión por parte de los lexicógrafos pues terminarán sin duda dando forma a un nuevo producto lexicográfico que, de una u otra manera, decidirá el futuro de la disciplina. Es difícil predecir qué va a pasar exactamente puesto que la continua introducción de nuevas y sorprendentes tecnologías, especialmente el advenimiento de la inteligencia artificial, sigue cambiando las reglas del juego. No obstante, con la información que tenemos ahora parece que las mencionadas tendencias transformativas apuntan a “una herramienta integrada,

basada en una plataforma digital y prestando un servicio personalizado por medio de datos lexicográficos” (Fuertes-Olivera *et al.*, 2018, p. 157).

Históricamente, la lexicografía se convirtió en un negocio exitoso a partir de la introducción de la imprenta. Como indicamos más arriba, hoy vivimos una situación bien contraria con las casas editoriales quedando sin ingresos y los lexicógrafos cazando recursos financieros para implementar sus ideas, en la mayoría de los casos sin éxito. Aparte de cierta perplejidad ante la aplicación de las tecnologías puestas a su servicio, parece que una visión anticuada de lo que es la lexicografía obstaculiza la solución de sus problemas actuales. En este sentido, la crisis por la que pasa la disciplina también toma la forma de una crisis de identidad que vuelve a plantear la vieja pregunta: ¿qué es la lexicografía? (Nomdedeu-Rull *et al.*, 2019).

En su clásico *Dictionary of the English Language*, Johnson (1755) definió la lexicografía como el “arte y práctica de escribir diccionarios” (art and practice of writing dictionaries). Más recientemente y de forma casi idéntica, Martínez de Sousa (1995), en su *Diccionario de Lexicografía Práctica*, define la disciplina como el “arte y técnica de elaborar diccionarios” aunque también ofrece una segunda definición que no es relevante aquí (“estudio del léxico de una lengua”).

Las definiciones de Johnson y Martínez de Sousa son clásicas y se repiten en un sinfín de obras académicas y de referencia. No sorprende, pues nadie puede negar que la compilación de diccionarios ha sido fundamental para la lexicografía. Ahora bien, la pregunta que debemos plantearnos es si esta definición no es demasiado estrecha y excluye otras actividades que no hayan sido tan visibles, aunque no dejan de ser, desde siempre, elementos constitutivos que hoy en día vuelven a cobrar relevancia para la disciplina.

Una pequeña excursión a los albores de la lexicografía europea nos ayudará a aclarar esta duda.

## 2. REGRESANDO A LAS RAÍCES

Hace más de 40 años, el estudioso egipcio Al-Kasimi (1977, p. 1) observó que los “principales motivos detrás del nacimiento de la lexicografía varían de una cultura a otra”, por lo que cada cultura desarrolla los “diccionarios que corresponden a sus requerimientos particulares”. Con esta perspectiva, y apoyándose en un estudio de Stathi (2006), Hanks (2013, p. 507) rastrea el origen de la lexicografía europea en la época griega clásica y explica cómo los escribas griegos, en el siglo V antes de nuestra era, “solían introducir glosas en las copias manuscritas de las obras de Homero y otros autores tempranos” para explicar “palabras obsoletas e inusuales”. Dos siglos más tarde, los estudiosos de la biblioteca de Alejandría recopilaron estas glosas en glosarios independientes (ver figura 1). Según McArthur (1986, p. 76), los historiadores de la lexicografía consideran que esta práctica constituye el mismo “origen del diccionario como tal”.

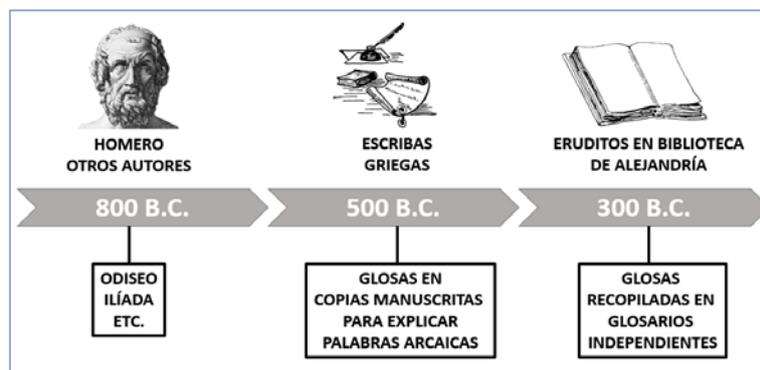


Figura 1. Cronología del nacimiento de la lexicografía europea

Desgraciadamente, ni Hanks ni McArthur explican claramente en qué siglo exactamente nació la lexicografía como disciplina. ¿Ocurrió en el siglo III o el V a. C.? Esta pregunta no es retórica pues tiene gran relevancia para la discusión de la sección anterior. Aunque la etimología de una palabra no necesariamente define su significado moderno, puede ser relevante conocerla cuando se trabaja en una línea de tiempo. El término “lexicografía” es originalmente griego y significa algo como “escribir sobre la lengua”, o sea, exactamente lo que hicieron los escribas. Ellos produjeron datos lexicográficos (glosas) asignados a palabras difíciles que más adelante fueron

recopilados en glosarios que representan prototipos de diccionarios en la tradición europea.

Si el trabajo de los escribas no se incorpora como un componente integral de la lexicografía, una parte creciente de lo que hacen los lexicógrafos del siglo XXI tampoco puede categorizarse como tal pues estos últimos también producen datos lexicográficos que a menudo no acaban estrenándose en diccionarios, ni impresos ni digitales. Sus datos se almacenan en bases de datos y pueden utilizarse para varios fines. Cuando, en abril de 2019, la empresa danesa Ordbogen A/S presentó la primera versión de su asistente de escritura español-inglés (Fisker, 2019), tenía que extraer datos lexicográficos de seis bases de datos diferentes para alimentar esta herramienta. Es muy probable que los lexicógrafos que originalmente prepararon estos datos no tengan ni idea de este uso de sus datos y que tampoco puedan reconocerlos en su nuevo entorno. Este fenómeno ya no es la excepción sino el inicio de una nueva era. En este sentido, Maldonado (2019) predice que:

El futuro de nuestro trabajo como lexicógrafos está abocado a convertirnos en suministradores de unos datos que se integrarán en herramientas y servicios lingüísticos más amplios: buscadores, asistentes de escritura, programas de traducción, correctores de estilo, etc. (p. 104)

Es más, los griegos del siglo V a. C. experimentaban obviamente necesidades de información cuando se encontraban con palabras “obsoletas e inusuales” en las obras de Homero y otros autores tempranos. Estas necesidades no fueron ni generales ni abstractas, sino que fueron muy concretas y estaban directamente relacionadas con una actividad específica (lectura) y un lugar específico (página, línea, posición) en el texto. Cuando los escribas insertaron glosas en las copias de las mencionadas obras, lo hicieron en el contexto específico donde se producía (o esperaban que se produjera) una necesidad de información. Por consiguiente, estas glosas o datos protolexicográficos constituían, desde el principio, *datos contextualizados* (Tarp & Gouws, 2019).

Los libros manuscritos eran valiosos y solo poca gente podía permitirse el lujo de tener sus propias copias. Por eso acudían a lugares públicos (el nacimiento de las bibliotecas) para disfrutar de las obras de Homero y demás autores tempranos. Los

escribas que trabajaban en estos lugares conocían probablemente algunos de estos lectores y podían observar sus problemas cuando se encontraban con las palabras difíciles. También es posible que discutieran los problemas con ellos. En este sentido, las glosas insertadas también representaban un *servicio personalizado* basado en un conocimiento personal entre los escribas (protolexicógrafos) y lectores (usuarios protolexicográficos).

Resulta, por lo tanto, que la lexicografía europea nace de una tradición en que los escribas ofrecían datos personalizados y contextualizados a lectores con necesidades de información sobre palabras que no entendían bien. Este fenómeno comenzó a desintegrarse dos siglos después cuando los bibliotecarios alejandrinos empezaron a compilar glosarios. Estos glosarios representaban sin duda alguna una invención práctica y útil de gran proyección histórica. Sin embargo, también tenían un efecto secundario menos deseable, pues hicieron que la satisfacción de las necesidades de información que se producían en un contexto específico fuera descontextualizada y externalizada a una fuente de información por separado.

La introducción de obras de referencia independientes creó una distancia entre el surgimiento de una necesidad de información y su solución lexicográfica. Esta distancia no dejó de crecer a lo largo de los siglos con el lamentable resultado de que el tiempo de consulta se iba alargando, el proceso de consulta se complicaba y el riesgo de no encontrar una solución adecuada aumentaba. De esta manera, el uso de datos contextualizados fue abandonado por la rama de la lexicografía práctica que se dedicaba a la compilación de diccionarios (Tarp & Gouws, 2019).

La elaboración de glosas por los escribas griegos y su inserción en copias manuscritas tienen muchas afinidades con el trabajo que deben hacer los lexicógrafos del siglo XXI. Ellos también deben prestar un servicio cada vez más personalizado que se dé directamente en el contexto donde surgen las necesidades de sus usuarios. En este sentido no hay nada nuevo bajo el sol. Pero a diferencia de los escribas clásicos que trabajaban con la pluma y el pergamino, los escribas modernos tienen a su disposición las nuevas tecnologías punteras que les permiten integrar los datos lexicográficos en herramientas sofisticadas como los lectores digitales y los asistentes

de escritura. Se trata de familiarizarse con estas herramientas y repensar la disciplina sobre esta base.

### 3. SUBIENDO LA CADENA DE VALOR

Los nuevos diccionarios en línea tienen grandes ventajas en comparación con el diccionario tradicional en lo que se refiere a su accesibilidad, calidad, cantidad de datos y actualización constante. Con todo, resulta que los internautas no están dispuestos a pagar el servicio, a menos que este les sea imprescindible para su trabajo o estudio. Además, podemos observar cómo cada vez más personas con necesidades de información acuden a fuentes alternativas, probablemente porque los diccionarios que conocen no se adaptan suficientemente a sus expectativas en la era digital. Estas tendencias contradictorias indican que los diccionarios van quedando obsoletos como fuentes de ingreso. Sin embargo, esto no necesariamente significa que los diccionarios también queden obsoletos como fuentes de información. Los diccionarios, por las características mencionadas, seguirán sin duda siendo útiles y hasta imprescindibles como obras de consulta para usuarios que buscan información especializada para resolver varios tipos de problemas. En una sociedad donde la comunicación y el acceso a información fiable son requisitos cada vez más importantes es difícil imaginarse que no haya una demanda social, al menos por algún tiempo, de diccionarios de calidad, rápida consulta y actualización constante. No debemos, por lo tanto, dejar de producir diccionarios con estas características.

Como vemos, existe una situación compleja con tendencias contradictorias. Por un lado, ya no se puede contar con los diccionarios como fuente principal para financiar el trabajo lexicográfico. Por otro lado, tampoco se puede dejar de desarrollar diccionarios que se aprovechen plenamente de las nuevas tecnologías para satisfacer las complejas necesidades de información que surgen en la sociedad contemporánea. De esto se desprende que la supervivencia de la lexicografía como práctica cultural requiere un modelo de negocios completamente nuevo que cumpla con dos requisitos fundamentales:

1. Generar ingresos mediante la comercialización de productos lexicográficos.
2. Financiar la compilación de diccionarios cada vez más avanzados, aunque estos no generen ingresos.

Parece un nudo gordiano. Y como cualquier otro nudo gordiano solo se puede resolver rompiéndose. La herramienta a aplicar en este caso debe moldearse mediante una integración de la lexicografía, la tecnología y la economía. Fuertes-Olivera (2019, p. 25) declara con toda razón que “el único camino es subir en la cadena de valor”. ¿Qué significa esto? Pues, significa que la lexicografía antes de todo debe tomar consciencia de que su producto por excelencia son los datos lexicográficos y que estos pueden usarse para un sinnúmero de aplicaciones que van desde su presentación en diccionarios hasta su integración en diferentes tipos de herramientas, asistentes, plataformas, servicios, etc. (ver figura 2).

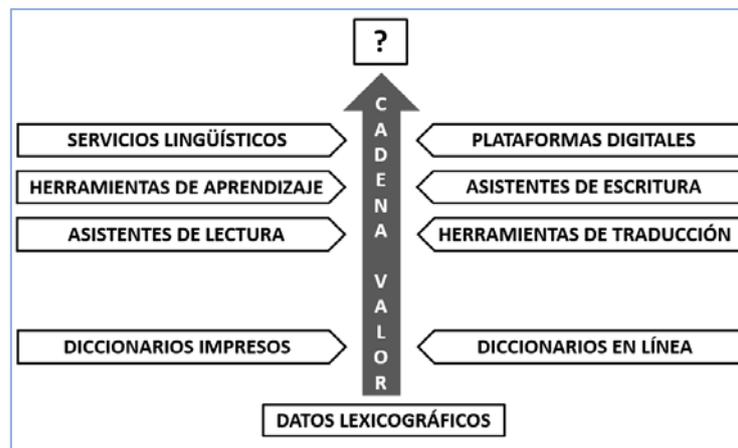


Figura 2. Cadena de valor con datos lexicográficos integrados

Entre la confluencia de avances tecnológicos en curso, Schwab (2016, s.n.) menciona explícitamente “la inteligencia artificial (IA), la robótica, el internet de las cosas (IoT), los vehículos autónomos, la impresión 3D, la nanotecnología, la biotecnología, la ciencia de materiales, el almacenamiento de energía y la computación cuántica”. El fundador del Foro Económico Mundial agrega:

Muchas de estas innovaciones están en sus albores, pero ya están llegando a un punto de inflexión en su desarrollo a medida que se construyen y amplifican mutuamente en una fusión de tecnologías a través de los mundos físico, digital y biológico. (Schwab, 2016, s.p.)

Ante esta perspectiva, no hay que cerrar la puerta a ninguna opción tecnológica. Al mismo tiempo, cabe señalar que la subida en la cadena de valor no se puede hacer de forma aleatoria y chapucera, pues para que la lexicografía prospere en el marco de la Cuarta Revolución Industrial hay que reunir, por lo menos, cuatro requerimientos fundamentales:

- Tener, o buscar a alguien que tenga, *músculos financieros*. Los proyectos generalmente requieren grandes inversiones y los productos necesitan comercializarse a escala internacional para generar suficientes ingresos.
- Comprometerse en una *colaboración interdisciplinaria* con informáticos, diseñadores industriales y otros expertos. Todo debe realizarse a una escala industrial con los agentes más cualificados.
- Mostrar *creatividad*. Como bien dijo Steve Jobs, no se trata de “encontrar soluciones lineales sin una perspectiva amplia”, sino que “sintetizar cosas nuevas”; cf. Wolf (1996, s.n.).
- Actuar con *autoestima y confianza*. La lexicografía no es una advenediza, sino una tradición cultural sumamente rica de más de cuatro mil años que tiene muchísimo que aportar a la actual avalancha tecnológica.

El camino hacia adelante no será un lecho de rosas. La lexicografía, en su estado actual, corre el riesgo de convertirse en una disciplina colonizada que depende de disciplinas ajenas con mayor poder tecnológico. Por ello, cabe evitar en la medida de lo posible que la lexicografía se conforme con un papel secundario como mero productor de materias primas (datos lexicográficos).

Para ello es preciso que los profesionales de la disciplina —o sea, los lexicógrafos— se integren plenamente en el nuevo mundo feliz. No deben dedicarse únicamente a la producción de datos lexicográficos, pues también deben preocuparse por el diseño de las nuevas herramientas tecnológicas que ellos mismos utilizan para extraer, elaborar y almacenar sus datos. Y ante todo, es necesario que se interesen mucho más por la presentación de estos datos, sean cuales sean los medios en que se presenten a los usuarios, pues para este tipo de trabajo tienen sobre los especialistas de otras disciplinas

más jóvenes la gran ventaja de ser los herederos de una cultura milenaria que ha acumulado grandes conocimientos de este aspecto.

Por su parte, las tradicionales casas lexicográficas tienen tres opciones:

- Invertir en el desarrollo de las mencionadas herramientas, plataformas y servicios. Esto también les permitiría emplear los ingresos para desarrollar y mejorar sus diccionarios.
- Producir datos lexicográficos a medida y venderlos a otros agentes económicos que los incorporan en sus herramientas, plataformas y servicios. En este último caso correrían el riesgo de convertirse en maquiladoras lexicográficas.
- Seguir el ejemplo de un creciente número de casas editoriales y abandonar el negocio lexicográfico por completo. En este caso, habrá otros agentes económicos que les sustituyen como ya hemos visto en algunos países (Simonsen, 2017).

No hay que ser ingenuo. En el actual sistema capitalista, ninguna empresa, a menos que reciba subvenciones públicas, tiene interés en invertir en algo que no genera beneficios de uno u otro tipo. Pues, para financiar una nueva generación de diccionarios digitales, estos podrían concebirse como promotores de los demás productos de la empresa además de fuente de última instancia a la que se puede referir desde estos productos para que se conviertan en productos *premium* mediante la optimización de la información suministrada. Con creatividad y espíritu emprendedor también se podrá resolver este desafío y sacar la lexicografía del pantano actual.

#### 4. PRESENTACIÓN DE *WRITE ASSISTANT*

*Write Assistant* es un buen ejemplo de tanto las nuevas posibilidades como los nuevos desafíos que se plantean a la lexicografía en este momento. Consiste en una serie de asistentes de escritura bilingües desarrollada por la empresa danesa Ordbogen A/S (cuyo nombre traducido en español sería "Diccionario S.A."). Esta empresa, que fue fundada por informáticos en 2001, irrumpió poco después en el mercado lexicográfico de Dinamarca conquistando en solo dos o tres años una participación en el mercado de diccionarios en línea de más del 90 % y sustituyendo a casas editoriales de prestigio que tenían décadas y hasta siglos de experiencia en la producción y venta de

diccionarios impresos. Hoy también se ha convertido una empresa líder en servicios lingüísticos y material didáctico en línea. Debido a su peculiar modelo de negocios, es uno de las muy pocas casas lexicográficas en Europa que gana dinero con sus diccionarios en línea. Sin embargo, este modelo que se introdujo en el momento apropiado en su país de origen ya no se puede reproducir en otros países europeos. Para expandirse en los mercados internacionales es preciso ofrecer otro producto tecnológico puntero. *Write Assistant* es una respuesta a este desafío.

Las principales características y la funcionalidad de *Write Assistant* han sido tratadas ampliamente por Tarp *et al.* (2017, 2018). Aquí, nos limitamos a tratar algunos aspectos relevantes para el tema de este artículo.

*Write Assistant* ha sido diseñado con solo una función principal según los criterios de la Teoría Funcional de la Lexicografía, o sea, prestar ayuda a usuarios que tengan problemas cuando escriben en una lengua extranjera (Tarp, 2008). La idea de desarrollar esta herramienta se inspiró en dos observaciones. La primera es que la mayor parte de la escritura de textos hoy en día se realiza en móviles, tabletas y ordenadores, en tanto que la escritura a mano se restringe cada vez más a la esfera personal. Esto también significa que la consulta de diccionarios para resolver problemas relacionados con la producción de textos se hace casi exclusivamente cuando se trabaja en esos aparatos. *Primera conclusión:* la asistencia lexicográfica debe ofrecerse directamente en móviles, tabletas y ordenadores.

La segunda observación es que la gente generalmente gasta demasiado tiempo consultando fuentes externas, incluso cuando estas fuentes están disponibles en los mismos aparatos. Esto puede interrumpir la fluidez de escritura e influir negativamente en su concentración. *Segunda conclusión:* La asistencia lexicográfica y de otro tipo debe integrarse en otras herramientas y prestarse directamente en los documentos y textos que los usuarios están escribiendo en programas como Word, Outlook, Gmail, PowerPoint, WeChat, WhatsApp, Messenger, Twitter, y un largo etcétera.

A partir de estas reflexiones, *Write Assistant* se ha diseñado como una aplicación que puede descargarse en el aparato del usuario ofreciéndole ayuda instantánea cuando redacta un texto en una segunda lengua. La aplicación conecta al usuario con *big data*

importado de dos fuentes empíricas, o sea, un corpus digital de L2 y una base de datos lexicográfica. Su fuerza motriz es un modelo lingüístico que ha sido entrenado sobre el corpus. Hasta ahora, el procesamiento se ha hecho por medio de una programación estadística, pero actualmente se está preparando la incorporación paralela de la inteligencia artificial para conseguir un modelo lingüístico aún mejor. De esta manera, el asistente de escritura se alimenta de datos lexicográficos importados de diferentes fuentes e, incluso, generados por medio de diferentes técnicas (ver figura 3).

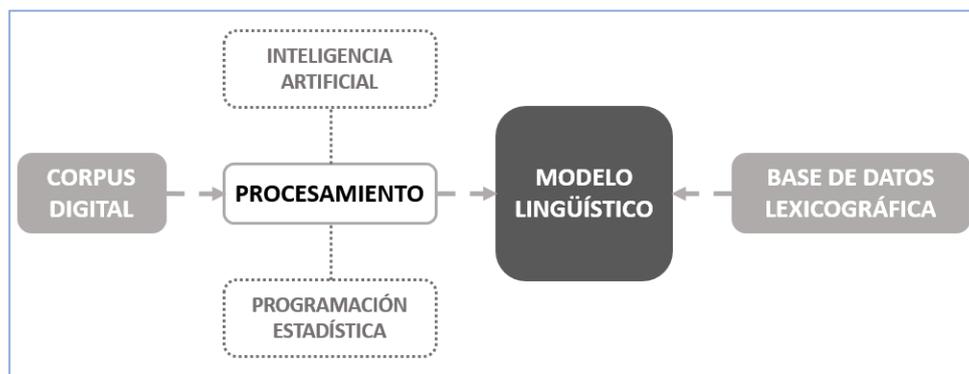


Figura 3. El modelo lingüístico y sus fuentes empíricas

Aunque *Write Assistant* fue publicado en 2019 para usuarios de seis lenguas maternas diferentes, todavía está en pleno proceso de desarrollo con nuevas opciones técnicas que se van incorporando paulatinamente. Una vez desarrollada plenamente, se prevé que la asistencia sea suministrada en tres ventanas emergentes diferentes que se abren dentro del mismo documento en que trabaja el usuario, o sea, las ventanas de sugerencias, consultas y alertas, respectivamente (ver figura 4).

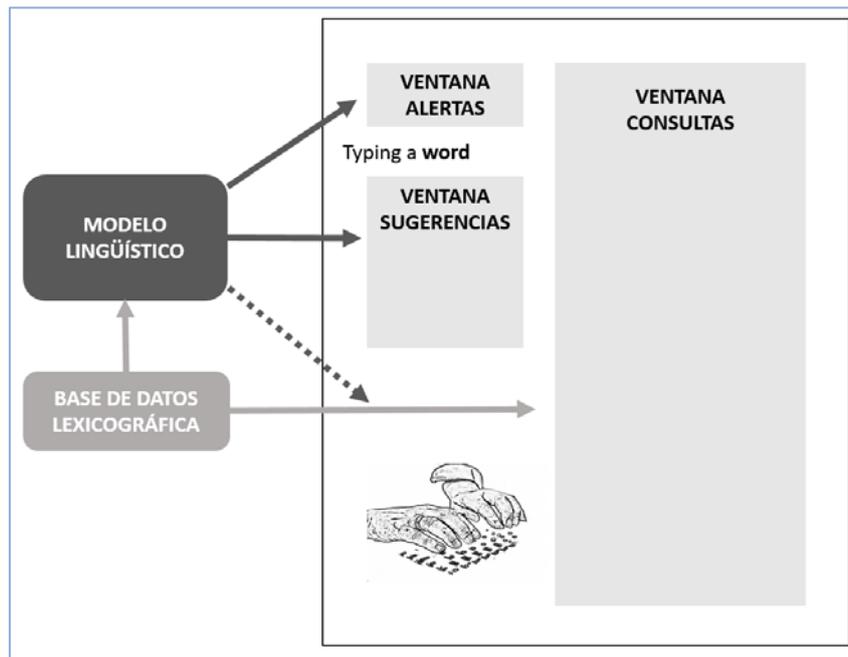


Figura 4. Las tres ventanas de *Write Assistant*

Aparte de su contenido, las tres ventanas tienen diferentes funcionamientos. La ventana de sugerencias aparece por defecto en la pantalla del usuario una vez que este empieza a escribir un texto. Esta ventana ofrece en primer lugar *terminaciones de palabras de L2*, y posibles *palabras siguientes*, cuando el escritor teclea una o más letras o una palabra completa de L2. Estas sugerencias se generan de forma automática por medio del modelo lingüístico y son efímeras en el sentido de que solo se pueden reproducir si el usuario vuelve a escribir exactamente la misma secuencia de palabras. Además, la ventana de sugerencias también ofrece *equivalentes de L2* cuando el usuario escribe una palabra de L1. Estos equivalentes son suministrados por la base de datos lexicográfica y presentados en un orden priorizado y sensible al contexto por medio del modelo lingüístico.

La ventana de consulta solo se abre cuando el usuario la activa. Si este no conoce las palabras propuestas por el modelo lingüístico, o no sabe cuál de ellas debe usar y cómo usarla, podrá activar esta ventana haciendo clic en una de las palabras o sencillamente tocando la pantalla sobre ellas. La ventana enseguida ofrece datos lexicográficos como significado, flexión, gramática, sintaxis, sinónimos, colocaciones, ejemplos, etc., todos ellos importados de la base de datos lexicográfica. La línea de

puntos en la figura 4 indica que está previsto que estos datos sean cada vez más adaptados al contexto específico a medida que mejore el modelo lingüístico con la introducción de la inteligencia artificial. En cierto sentido, esta ventana constituye *la ventana al futuro* de la lexicografía.

La tercera ventana todavía no ha sido diseñada, pero se prevé que ofrezca alertas que se activan cuando el usuario escriba una palabra que los diseñadores de la herramienta consideran como desafiante desde el punto de vista lingüístico o cultural.

## 5. ALGUNOS DESAFÍOS LEXICOGRÁFICOS

Los tres tipos diferentes de ventanas discutidos en el capítulo anterior plantean una serie de nuevos desafíos para la lexicografía, entre ellos:

- ¿Qué tipos de datos lexicográficos requieren los asistentes de escritura con las características mencionadas?
- ¿Cuáles son las palabras que requieren alertas lingüísticas y culturales?
- ¿Cuáles son los datos lexicográficos requeridos para sostener estas alertas?
- ¿Qué tipos de datos lexicográficos son necesarios para que los asistentes de escritura, según el deseo expresado por varios usuarios, también se diseñen como herramientas de aprendizaje de L2?
- ¿Cómo pueden los datos lexicográficos tradicionales interactuar con datos que se generen automáticamente por medio de la inteligencia artificial?
- ¿Qué consecuencias tienen los asistentes de escritura para el diseño de nuevas bases de datos lexicográficos?
- ¿Cómo se puede presentar la cantidad relativamente grandes de datos lexicográficos en la pequeña ventana de consultas sin incurrir en una sobrecarga de información que complique el acceso a los datos requeridos en cada caso?

Todas estas preguntas merecen un profundo análisis que incorpore a lexicógrafos, informáticos, diseñadores y especialistas en la enseñanza de lenguas, entre otros. A continuación intentaremos dar respuesta a la última pregunta.

## 6. LA VENTANA AL FUTURO

La propuesta que sigue forma parte de un proceso de intercambio de ideas entre lexicógrafos e informáticos. Como tal, no representa necesariamente la versión que terminará publicándose y presentándose a los usuarios de *Write Assistant* pero ofrece una instantánea de un método de trabajo que quizás puede aplicarse en otros proyectos similares. Este carácter inacabado de la propuesta también significa que los diseñadores todavía no se han incorporado al proyecto, por lo que la presentación gráfica deja mucho por desear.

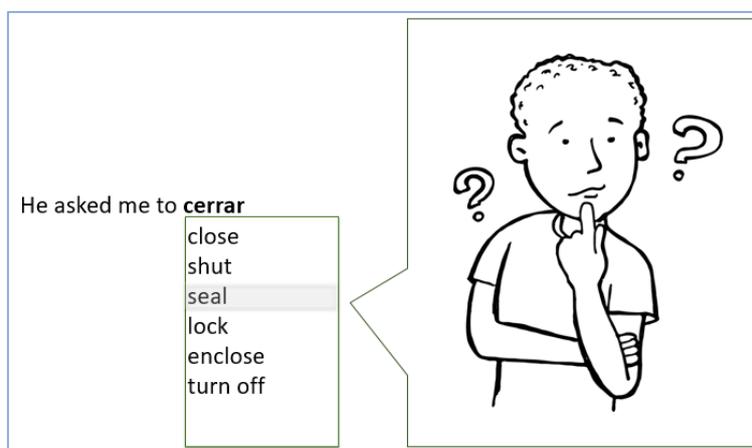


Figura 5. El desafío: ¿Cómo llenar la ventana de consultas?

En la figura 5, vemos el punto de partida. Un usuario con el español como lengua materna escribe “He asked me to” y a continuación teclea la palabra española *cerrar* porque no sabe qué palabra inglesa debe usar. Enseguida aparece en la ventana de sugerencias una lista de posibles equivalentes ingleses. El usuario no conoce el significado de “seal” por lo que activa la ventana de consultas mediante un clic en “seal”. El desafío ahora consiste en llenar esta ventana para satisfacer las necesidades concretas del usuario. Una posible solución se puede ver en figura 6.

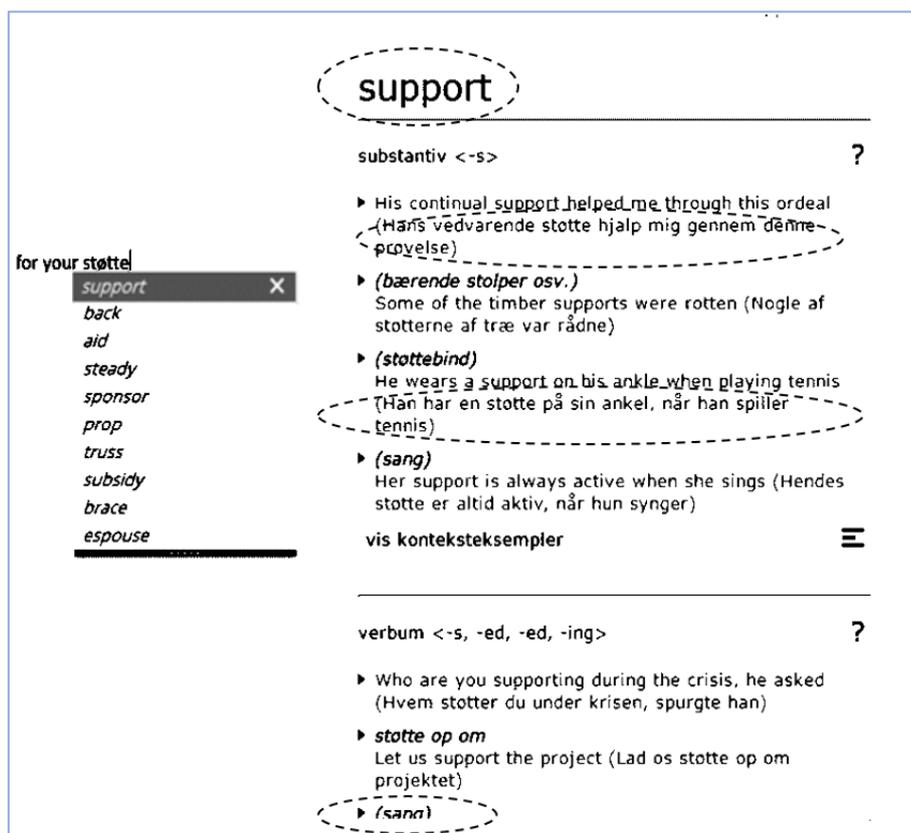


Figura 6. Versión beta del asistente de escritura danés-inglés

La figura 6 es una captura de pantalla de la original versión beta del asistente de escritura danés-inglés en la que un artículo de uno de los diccionarios en línea de Ordbogen se ha usado para rellenar la ventana de consultas. Esta solución no parece la más apropiada para una herramienta como *Write Assistant* si se tiene en cuenta que sus usuarios deben encontrar la información requerida sin dar demasiadas vueltas que les hacen perder el foco en el mensaje que están escribiendo. Para evitar este problema hace falta una selección meticulosa de los datos que se presenten por defecto. No debe haber datos superfluos y de difícil acceso e interpretación. El diseño del artículo reproducido en la figura 6 no ayuda en este sentido. Entre los principales problemas se puede mencionar:

- El lema, aparte del tamaño de sus letras, parece completamente redundante ya que esta palabra está marcada en la ventana de sugerencias y es difícil imaginarse que el usuario no sabe desde qué palabra ha accedido el artículo.

- Las explicaciones se ofrecen por medio de oraciones que han sido seleccionadas como ejemplos. Esta solución es problemática, pues cuesta tiempo y requiere un proceso mentalmente complejo deducir el significado de estas oraciones.
- Los ejemplos se ofrecen por defecto en tanto L1 como L2. Esta solución ocupa espacio con el resultado de que varios ejemplos quedan ocultos y requieren otro paso para visualizarse (“vis konteksteksempler”). Habría sido mejor suministrar por defecto una oración de L1 con la opción de expandirla con su traducción en inglés.
- La cantidad de datos es excesiva, por lo menos para esta fase de la consulta, ocupa espacio y obliga al usuario a desplazarse hacia abajo para acceder otros datos relevantes sobre el significado de la palabra.

Como ya mencionamos, fue necesario importar datos de seis fuentes diferentes para alimentar la primera versión del asistente de escritura español-inglés, si bien este resultado tampoco es satisfactorio. Para resolver este problema, Ordbogen A/S está preparando una base de datos lexicográfica inglés-español-inglés adaptada a los requerimientos de *Write Assistant* y que actualmente se está produciendo en el marco del proyecto Valladolid-UVa (Fuertes-Olivera *et al.*, 2018).

En esta base de datos, los datos asignados a la palabra “seal” ocupan nueve páginas si se imprimen en papel de tamaño A4, y comprenden tres lemas (un verbo y dos nombres con diferentes paradigmas de flexión) con un total de 19 sentidos y una gran cantidad de datos asignados a cada acepción (ver figura 7). El reto ahora es cómo se puede presentar y estructurar todos estos datos de manera elegante en la ventana de consultas sin incurrir en una sobrecarga de datos y sin obligar al usuario a tomar demasiados pasos hasta obtener la información deseada (Gouws & Tarp, 2017).

**seal** verbo  
seal, seals, sealed, sealing  
<https://howjsay.com/search?word=seal> pronunciación

1. cerrar algo, por ejemplo un sobre; se refiere a la actividad física con la que alguien mete un documento en un sobre y lo cierra

valencia

- someone seals something
- alguien sella algo

Nota de uso

- Verbo transitivo, es decir usado con un complemento directo.

sinónimo

- seal up
- close up

equivalente

- sellar verbo

colocaciones

- she folded it, sealed the envelope, and ran to the postbox
- lo dobló, cerró el sobre y corrió al buzón de correos

2. cerrar algo y colocar un precinto como medida de seguridad, es decir para informar que alguien ha abierto ese algo sin permiso; ese precinto es un sello de cera o algo similar

valencia

- someone seals something
- alguien lacra algo

Nota de uso

- Verbo transitivo, es decir usado con un complemento directo.

sinónimo

- fasten
- secure

equivalente

- lacrar verbo

colocaciones

- the civil servant sealed it with the king's seal
- el funcionario lo lacró con el sello real

3. cerrar algo, por ejemplo un contenedor, una puerta, una tumba, etc. para que no pueda abrirse; se hace como medida de seguridad

valencia

- someone seals something
- alguien sella algo

Nota de uso

- Verbo transitivo, es decir usado con un complemento directo.

sinónimo

- shut
- close
- stop up

colocaciones

- the container was sealed and nobody could open it up again
- the container was sealed and nobody could open it up again
- sealed the wood before applying paint
- sellaron la madera antes de pintar

ejemplo

- She merely filled the containers, sealed them with a cork, and pasted on labels.
- Simplemente cerró los contenedores, los selló con un corcho y puso las etiquetas.

4. cerrar algo de forma hermética para que no entre ni agua ni aire

valencia

- someone seals something
- alguien impermeabiliza algo

Nota de uso

- Verbo transitivo, es decir usado con un complemento directo.

sinónimo

Figura 7. Extractos de la base de datos inglés-español del proyecto Valladolid-Uva

A continuación, presentaremos un diseño alternativo de la ventana de consultas que se basa en el criterio de relevancia aplicado a la lexicografía (Bothma & Tarp, 2014). Los datos que se estiman más relevantes para el usuario se ofrecen primero —en la medida que sea posible— en tanto que los que se consideran menos relevantes deben

accederse después. Este criterio implica, entre otras cosas, que técnicas de navegación como desplazamientos hacia abajo, que requieren más pasos de parte del usuario, se reduzcan a un mínimo y solo se usen cuando sean inevitables. Los principios rectores son “menos es más”, el famoso lema formulado por Ludwig Mies van der Rohe en 1947, y “simplicidad es la máxima sofisticación”, una frase también famosa y atribuida a Leonardo da Vinci aunque no hay documentación de que la haya usado. La figura 8 muestra, humildemente, una primera tentativa de diseño inspirada por este espíritu.

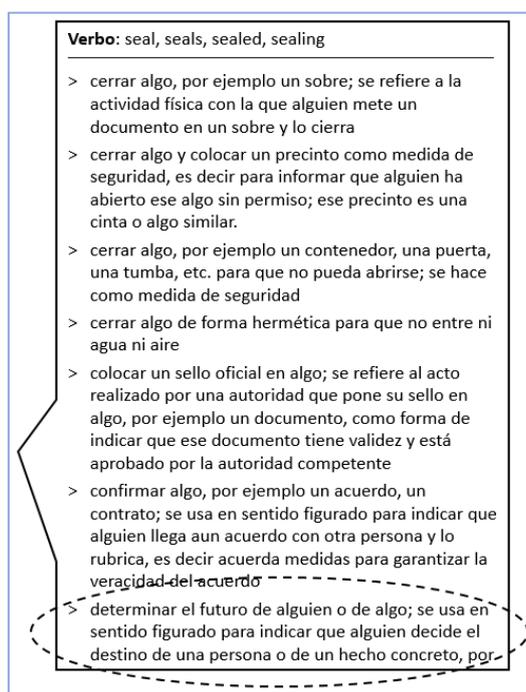


Figura 8. Diseño de la ventana de sugerencias según el criterio de relevancia

El artículo que se expone en la figura 8 ha sido elaborado a partir de lo que se considera más relevante para un usuario que ha iniciado una consulta por no conocer, o tener dudas sobre el uso de las palabras que *Write Assistant* le ha ofrecido en la ventana de sugerencias. Se estima que lo que ante todo importa a un usuario en esta situación es conocer el *significado* de las palabras propuestas, pues sin saber lo que significan resulta impropio y hasta arriesgado usarlas. Además, si el usuario ya conoce el significado de “seal”, podría ser relevante la conjugación de este verbo aunque la flexión en inglés generalmente no constituye el mismo problema como en otras lenguas, por ejemplo el español.

Las definiciones de los diferentes sentidos del verbo “seal” que se han visualizado en la figura 8 han sido importadas de la base de datos Valladolid-UVa mostrada en la figura 7. Ahora bien, esta base de datos contiene nueve acepciones de este verbo, por lo que el usuario se ve obligado a desplazarse hacia abajo para acceder a las demás acepciones. Algunas palabras comunes contenidas en la base de datos tienen 20, 30 o más acepciones, lo que significa que el usuario en algunos casos tendría que activar el botón de desplazamiento hacia abajo varias veces para asegurarse de que la palabra efectivamente tiene el significado que desea expresar. Como el resultado es un proceso de consulta largo y bastante complicado, habrá que encontrar otra solución que reduzca el número de veces en que sea necesario.

La imagen a la izquierda en la figura 9 muestra una solución alternativa. En este caso, el tamaño de todas las definiciones se ha reducido a solo una línea permitiendo la visualización inmediata de todos los sentidos del verbo “seal” (e incluso algunos de los que pertenecen al nombre “seal”). La idea es que el usuario de este modo puede tener una idea preliminar del significado de cada sentido. Si le “suena” o “huele” bien podrá desplegar todo el artículo mediante un clic en “>” como se ve en la imagen de la derecha en la misma figura 9.

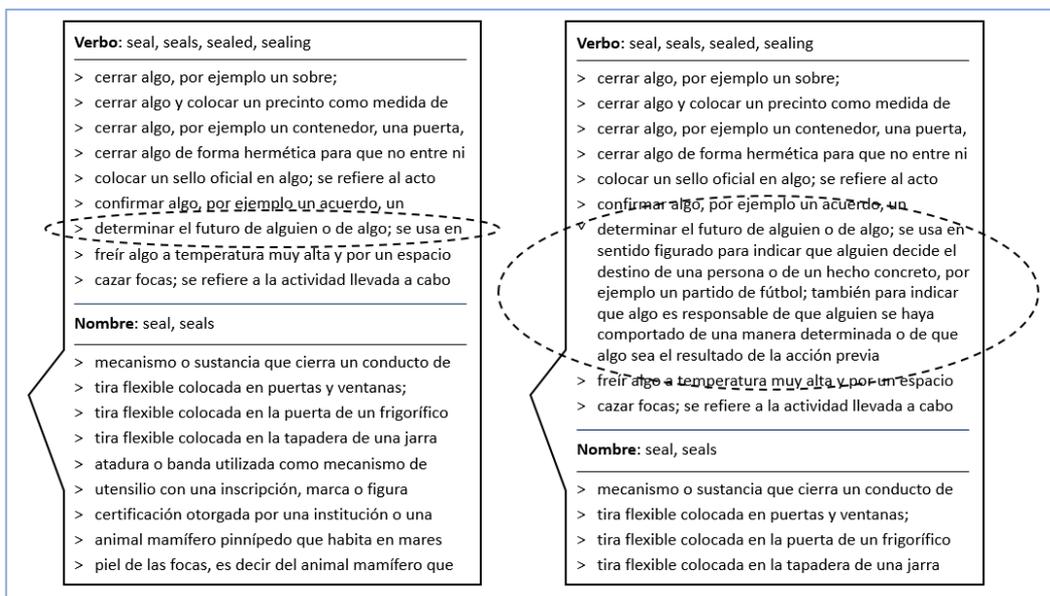


Figura 9. Definiciones compactas y desplegadas

La solución esbozada en la figura 9 crea otros problemas, pues si las breves definiciones de una sola línea no están previstas en la base de datos, muchas oraciones resultarían recortadas y sin terminarse. Si bien este procedimiento puede recomendarse como solución interina, en algunos casos podría ser difícil para el usuario deducir el significado preliminar a partir de las definiciones abreviadas a la izquierda en la figura. Por tanto, a largo plazo la inclusión de definiciones tanto cortas como largas en la base de datos deberá plantearse desde el principio como ya discutieron Fuertes-Olivera *et al.* (2018, pp. 159-160).

La disposición de las definiciones desplegadas entre las demás definiciones abreviadas supone otra inconveniencia, pues deja poco espacio para otros tipos de datos que podrían invitar al usuario a tomar otro paso en el proceso de consulta. Además, si la definición desplegada se encuentra en la parte inferior de la ventana, habría que desplazarse hacia abajo para leerla en su totalidad. Una posible solución a estos problemas sería que las demás definiciones abreviadas se ocultan y la desplegada se coloca en la parte superior de la ventana de consultas como se indica en la figura 10.

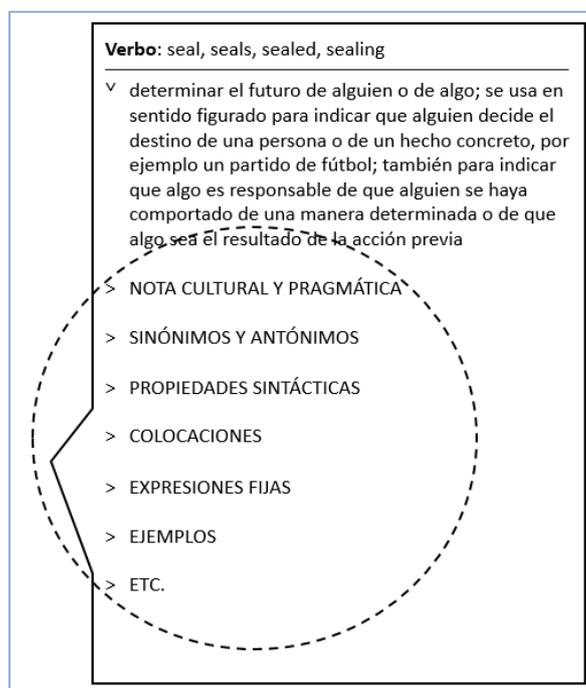


Figura 10. Artículo con metatextos

La propuesta recogida en la figura 10 también incluye una lista de llamados metatextos que todavía precisan de una elaboración ulterior. La idea es que el usuario mediante un clic en uno de estos metatextos puede acceder a datos adicionales que le hagan falta en una consulta concreta. A diferencia de la conjugación que está asignada al verbo como tal, todos estos datos están asignados a una acepción específica del verbo como también recomiendan Nomdedeu-Rull & Tarp (2018) por motivos didácticos.

Esta forma de estructurar los datos lexicográficos no es la más común en los diccionarios que ya existen, pero a nuestro juicio es la más natural. Tarp (2018), por ejemplo, discute cómo muchos verbos españoles tienen diferentes propiedades sintácticas según la acepción específica, en tanto que Dam-Jensen & Tarp (2019) abordan el mismo fenómeno en el caso de los adjetivos, por lo que los autores aconsejan que estos tipos de datos siempre se estructuren en conexión con la acepción pertinente. El mismo procedimiento es también aconsejable en el caso de notas culturales y pragmáticas, sinónimos, antónimos, colocaciones, expresiones fijas no idiomáticas y ejemplos aunque hay muchos casos en los que algún que otro dato lexicográfico de este tipo vale para dos o más acepciones, e incluso todas las acepciones, de una palabra específica.

Aparte de la flexión, los únicos datos relevantes para el usuario de *Write Assistant* y que no pueden asignarse a una acepción particular son las expresiones idiomáticas. Estas expresiones constituyen un caso especial puesto que su significado no se establece por la suma de los significados de las palabras que las integran. Por consiguiente, en la base de datos Valladolid-UVa están almacenadas en cartas por separado con su propio significado y demás datos, y la única relación que tienen con las cartas que recogen sus partes integrantes es una referencia (enlace) de estas a aquellas. Todavía se está discutiendo cómo pueden presentarse en *Write Assistant* e interactuar con las sugerencias generadas por la inteligencia artificial, por lo que no se discutirán aquí.

De todas formas, resulta inevitable que la materialización de la idea expuesta en este párrafo eleve el coste de producción de la base de datos lexicográfica. No obstante, creemos que hay buenos argumentos para estructurar los datos como hemos indicado

si se tiene en cuenta que el objetivo es un producto lexicográfico de calidad que pueda sostener una herramienta integrada de gran potencial como *Write Assistant*.

## 7. DESARROLLO GUIADO POR PRUEBAS

Los desafíos planteados por *Write Assistant* y otras herramientas similares no pueden resolverse por medio de los tradicionales estudios sobre el uso de diccionarios a los que se refiere Maldonado (2019) al inicio de este artículo. Estos estudios suelen interesarse por el uso de los diccionarios existentes (o los prototipos que luego no se publican) y en este sentido representan un tipo de *investigación retrospectiva*. La generalización de la mayoría de sus conclusiones también resulta difícil ya que cada nuevo proyecto lexicográfico tiene su *propia personalidad*.

Esta última observación tiene especial relevancia para un proyecto innovador como *Write Assistant* que se apoya en tecnologías completamente nuevas y disruptivas. Pero esto no significa que no se deba estudiar el uso de esta herramienta e interesarse por la reacción de sus usuarios. Al contrario, significa que hay que aplicar otros métodos de investigación que se ajusten mejor a las nuevas tecnologías. El objetivo no es publicar artículos y libros meses o años después del estudio sino hacer un *esfuerzo continuo* para desarrollar y perfeccionar el nuevo producto con el fin de conseguir la máxima satisfacción de sus usuarios.

No se trata de comprobar un producto acabado como los diccionarios impresos que solían tardar varios años en salir en nuevas ediciones. Se trata de confirmar, modificar y perfeccionar el funcionamiento de un producto que, por su tecnología, es mucho más flexible que los diccionarios tradicionales, y que, por la competencia que sufre, se encuentra en un estado permanente de desarrollo y perfeccionamiento. Todo esto requiere otro enfoque metodológico diferente del que tradicionalmente se ha usado dentro de la lexicografía.

La respuesta a este desafío es el llamado *desarrollo guiado por pruebas* (TDD por su denominación en inglés). Es una metodología que ya se aplica en muchas empresas de alta tecnología.

Pongamos un ejemplo. En la Figura 10 se presentó una propuesta con una lista de metatextos en orden arbitrario. Pero tal orden es relevante para el buen uso de la

herramienta. Una lista de enlaces de este tipo que invita al usuario a tomar pasos adicionales debe de estructurarse de manera tal que el proceso de consulta le sea lo más fácil posible. También se puede prever que muchos de los futuros usuarios no manejen términos especializados como colocaciones y propiedades sintácticas, por lo que los metatextos debieran formularse en un lenguaje más sencillo. Pues las dos preguntas del millón son:

1. ¿Cuál será el orden más idóneo de los metatextos?
2. ¿Cómo deben formularse para que sean más fáciles de comprender?

Las respuestas a estas preguntas pueden buscarse por medio de procesos iterativos con el mencionado desarrollo guiado por pruebas. Funciona más o menos de la siguiente manera: Una versión beta de *Write Assistant* con metatextos se suministra a dos o tres pequeños grupos focales, cada uno de unos 10-15 aprendices españoles del inglés. Los verificadores los observan mientras utilizan la herramienta durante cierto tiempo, p.ej. una hora, y luego les entrevistan para aclarar dudas y recoger sus opiniones y sugerencias. Sobre esta base se prepara una nueva versión con los metatextos ajustados. Esta versión la prueban otros grupos un poco más grandes, p.ej. de unos 20-25 aprendices, a los que también se observan y entrevistan. Esto conduce a otra modificación que se ofrece a otros tantos grupos focales, etcétera, etcétera.

De forma similar, se realizan estudios sobre todos los grandes y pequeños aspectos que son relevantes para el buen uso y máximo provecho de *Write Assistant*. Podría ser el diseño general de la interfaz o de las diferentes ventanas, la funcionalidad de los diversos botones y enlaces, la utilidad de las definiciones de una línea, la organización de los datos adicionales accesibles mediante los metatextos, etc.

Todos los estudios de este tipo los conducen los verificadores profesionales. A diferencia de los profesores universitarios que tienen cada vez menos tiempo para una investigación que se realiza entre clase y clase, los verificadores industriales son técnicos que dedican todo su tiempo a testar el producto. Trabajan en equipo junto con informáticos y diseñadores que enseguida implementan las modificaciones y ajustes que se estiman necesarios como resultado de su trabajo. Como se desprende, es un método adaptado a las nuevas tecnologías y bastante rápido. El proceso inicial

suele durar unas pocas semanas hasta que se consiga un resultado satisfactorio. Como ya indicamos, este desarrollo guiado por pruebas constituye en todo sentido el método más adecuado para mejorar y perfeccionar el producto lexicográfico en el marco de la Cuarta Revolución Industrial.

## 8. CONCLUSIÓN

El subtítulo de este artículo, si bien contiene un importante grano de verdad, es ante todo una provocación que intenta despertar a una disciplina que últimamente viene sufriendo el síndrome de la Bella Durmiente. Con este punto de partida, constatamos inicialmente que la publicación de nuevos diccionarios generales está estancándose, especialmente en los países tecnológicamente más avanzados, y que este fenómeno se debe a un modelo de negocios obsoleto. Al mismo tiempo, argumentamos que la sociedad actual sigue necesitando diccionarios de calidad, de rápido acceso y actualización constante, contradiciendo en cierto modo el enunciado hecho en el subtítulo. Sacamos la conclusión de que, para salvar los diccionarios como *fuentes de información*, debemos despedirnos de ellos como *fuentes de ingreso* y buscar otro modelo de financiación. De ahí el subtítulo.

Para detectar este nuevo modelo cuestionamos la visión tradicional de que la esencia de la lexicografía consiste en la compilación de diccionarios. Una breve excursión a la Época Griega Clásica ha desvelado que la vertiente europea de la disciplina nació, no con la compilación de diccionarios, sino con la producción de datos (glosas) que solo se recopilaban en los primeros diccionarios (glosarios) siglos después. Volviendo a nuestro tiempo, observamos cómo los lexicógrafos modernos –consciente o inconscientemente– cada vez más siguen los pasos de sus predecesores griegos sin que sus productos tomen la forma de un diccionario. Constatamos que esta tendencia irreversible se debe a la introducción de nuevas tecnologías disruptivas que están transformando la lexicografía en todos sus aspectos principales.

A partir de estas reflexiones esbozamos un nuevo modelo de negocios cuyo éxito requiere que se suba en la cadena de valor integrando los datos lexicográficos en herramientas, servicios y plataformas que se basan en las mencionadas tecnologías. Entre los requerimientos para que este modelo prospere, destacamos una estrecha

colaboración interdisciplinaria con informáticos, diseñadores y otros especialistas. A continuación, mostramos cómo esta idea puede plasmarse en un asistente de escritura enfocándose en la presentación de los datos lexicográficos en su ventana de consultas. En este contexto, introducimos el desarrollo guiado por pruebas como un método eficaz que sirve para realizar estudios sobre el uso de las herramientas sofisticadas de alta tecnología.

El título principal dado a este artículo se formuló para transmitir optimismo. Es innegable que la lexicografía vive un momento de dificultades acompañadas de cierta confusión y pesimismo. Esta reacción es comprensible ante lo que Fuertes-Olivera (2016) ha denominado una explosión cámbrica dentro de la disciplina. Sin embargo, un estudio metódico de las actuales tendencias tecnológicas en la sociedad en que vivimos no justifica el mencionado pesimismo. El asistente de escritura discutido en los párrafos anteriores es un ejemplo de que hay un futuro para la lexicografía. Como indicamos en la Figura 2, hay y habrá muchos más ejemplos. Todo depende de los lexicógrafos que deben liberar la mente y adaptarse a las nuevas condiciones. Si no se adaptan y se incrustan en una visión anticuada de lo que es su disciplina seguirá la confusión. En este caso, es muy probable que les sustituyan otros agentes con la mente más abierta.

De una u otra forma, la lexicografía sobrevivirá. La ventana al futuro ya está allí.

Nota: Este artículo es una adaptación y desarrollo ulterior de un artículo en inglés (Tarp 2019).

### **AGRADECIMIENTOS**

Agradecemos al Ministerio de Economía y Competitividad de España (Proyecto con Ref. FFI2014-52462-P) la financiación aportada para la realización de este trabajo.

## BIBLIOGRAFÍA

- Bothma, T.J.D. y S. Tarp (2014). Why Relevance Theory Is Relevant for Lexicography, *Lexicographica*, 30, pp. 350-378.
- Dam-Jensen, H. y S. Tarp (2019). El tratamiento sintáctico del adjetivo en gramáticas y diccionarios de español, *Revista Internacional de Lenguas Extranjeras*, 10, pp. 13-47.
- De Schryver, G.-M. (2003). Lexicographers' Dreams in the Electronic-dictionary Age, *International Journal of Lexicography*, 16(2), pp. 143-199.
- Fisker, K. (Ed.) (2019). *Write Assistant. Spanish-English*. Odense: Ordbogen A/S. Available at: [www.writeassistant.com/en/](http://www.writeassistant.com/en/).
- Fuertes-Olivera, P.A. (2016). A Cambrian Explosion in Lexicography: Some Reflections for Designing and Constructing Specialised Online Dictionaries, *International Journal of Lexicography*, 29(2): 226-247.
- Fuertes-Olivera, P.A. (2019). Designing and Making Commercially Driven Integrated Dictionary Portals: the Diccionarios Valladolid-UVa, *Lexicography*, 6(1), pp. 21-41.
- Fuertes-Olivera, P.A. y S. Tarp (2014). *Theory and practice of specialised online dictionaries: Lexicography versus terminography*. Berlín, Boston: De Gruyter.
- Fuertes-Olivera, P.A., S. Tarp y P. Sepstrup (2018). New Insights in the Design and Compilation of Digital Bilingual Lexicographical Products: the Case of the Diccionarios Valladolid-UVa, *Lexikos*, 28, pp. 152-176.
- Gouws, R.H. y S. Tarp (2017). Information Overload and Data Overload in Lexicography, *International Journal of Lexicography*, 30(4), pp. 389-415.
- Hanks, P. (2010). Lexicography, Printing Technology, and the Spread of Renaissance Culture. En A. Dykstra y T. Schoonheim (Eds.), *Proceedings of the XIV Euralex International Congress*. Leeuwarden: Fryske Akademy. pp. 988-1016
- Hanks, P. (2012). The corpus revolution in lexicography, en *International Journal of Lexicography*, 25(4), pp. 398-436.
- Hanks, P. (2013). Lexicography from Earliest Times to the Present. En A. Keith (Ed.), *The Oxford Handbook of the History of Linguistics*. Oxford: Oxford University Press. pp. 503-536.
- Johnson, S. (1755). *Dictionary of the English Language*. London: J. & P. Knapton.
- Maldonado, M.C. (2019). Las investigaciones de mercado en lexicografía comercial: un aprendizaje para el mundo académico e investigador, *Revista Internacional de Lenguas Extranjeras*, 10, pp. 101-118.
- Martínez de Sousa, J. (1995). *Diccionario de lexicografía práctica*. Barcelona: Bibliograf.
- McArthur, T. (1986). *Worlds of Reference*. Cambridge: Cambridge University Press.
- Nesi, H. (2015). The Demands of Users and the Publishing World: Printed or Online? Free or Paid for? En P. Durkin (Ed.), *The Oxford Handbook of Lexicography*. Oxford: Oxford University Press, pp. 579-589.
- Nielsen, S. 2013. The Future of Dictionaries, Dictionaries of the Future. En H. Jackson (Ed.), *Bloomsbury Companion to Lexicography*. London, New York: Bloomsbury. pp. 355-372.
- Nomdedeu-Rull, A. y S. Tarp. (2018) Hacia un modelo de diccionario en línea para aprendices de español como LE/L2, *Journal of Spanish Language Teaching*, 5(1), pp. 50-65.

- Nomdedeu-Rull, A. (Coord.), H. Hernández Hernández, C. Maldonado Gonzáles y S. Tarp. 2019. El estatus de la lexicografía. Nuevas y variadas respuestas a una antigua cuestión. En C. Calvo Rigual y F. Robles i Sabater (Eds.), *La investigación en lexicografía hoy: diccionarios bilingües, lingüística y uso del diccionario. Quaderns de Filologia*. Anejo 85: pp. 699-733.
- Rundell, M. (2007). The Dictionary of the Future. En S. Granger (Ed.), *Optimizing the Role of Language in Technology-enhanced Learning. Proceedings of the expert workshop organized in Louvain-la-Neuve* (Belgium), 4-5 October 2007. pp. 49–51. <https://hal.archives-ouvertes.fr/hal-00197203/document/>. (Consultado 12 de septiembre, 2019)
- Rundell, M. 2010. What Future for the Learner's Dictionary? En I. Kernerman y P. Bogaards (Eds.), *English Learners' Dictionaries at the DSN4 2009*. Jerusalem: Kdictionaries. pp. 169-175.
- Rundell, M. 2012. The Road to Automated Lexicography: an Editor's Viewpoint. En S. Granger y M. Paquot (Eds.), *Electronic Lexicography*. Oxford: Oxford University Press. pp. 15-30.
- Rundell, M. y A. Kilgarriff (2011). Automating the Creation of Dictionaries: Where Will It All End? En F. Meunier, S.D. Cock, G. Gilquin y M. Paquot (Eds.), *A Taste for Corpora. In Honour of Sylviane Granger*. Amsterdam, Philadelphia: John Benjamins. pp. 257-282.
- Schwab, K. (2015). *The Fourth Industrial Revolution: What It Means and How to Respond. Foreign Affairs*. 12th December, 2015.
- Schwab, K. (2016.) *La Cuarta Revolución Industrial*. Ginebra: World Economic Forum.
- Simonsen, H.K. (2017). Lexicography: What is the Business Model? En I. Kosem, C. Tiberius, M. Jakubiček, J. Kallas, S. Krek y V. Baisa (Eds.), *Electronic Lexicography in the 21st Century. Proceedings of the eLex 2017 Conference, 2017*. Brno: Lexical Computing CZ. pp. 395-415.
- Stathi, E. (2006). Greek Lexicography, Classical. En K. Brown (Ed.), *Encyclopedia of Language and Linguistics*. Second Edition, Vol. 5. Oxford: Elsevier. pp. 145-146.
- Tarp, S. (2008). *Lexicography in the Borderland between Knowledge and Non-knowledge*. Tübingen: Niemeyer.
- Tarp, S. (2011). Lexicographical and Other e-Tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction. En P.A. Fuertes-Olivera y H. Bergenholtz (Eds.), *e-Lxicography: The Internet, Digital Initiatives and Lexicography*. Londres, Nueva York: Continuum. pp. 54-70.
- Tarp, S. (2018). El desafío de los datos sintácticos en los diccionarios para aprendices extranjeros del español, en *Revista de Filología de la Universidad de La Laguna*, 36, pp. 367-391.
- Tarp, S. (2019). Connecting the Dots: Tradition and Disruption in Lexicography, en *Lexikos*, 29, pp. 224-249.
- Tarp, S. y P.A. Fuertes-Olivera (2016). Advantages and Disadvantages in the Use of Internet as a Corpus: The Case of the Online Dictionaries of Spanish Valladolid-Uva, en *Lexikos*, 26, pp. 273-296.
- Tarp, S. y P.A. Fuertes-Olivera (2018). Métodos e técnicas para usar a Internet diretamente como corpus: o caso dos Dicionários on-line de Espanhol Valladolid-UVA, en *Cadernos de Tradução*, 43: pp. 10-32.
- Tarp, S. y R.H. Gouws (2019). Lexicographical Contextualization and Personalization: a New Perspective, en *Lexikos*, 29, pp. 250-268.
- Tarp, S., K. Fisker y P. Sepstrup (2017). L2 Write Assistants and Context-aware Dictionaries: New Challenges to Lexicography, en *Lexikos*, 27, pp. 494-521.

- Tarp, S., K. Fisker y P. Sepstrup (2018). Dicionários sensíveis ao contexto e integradas a assistentes de escrita em L2: Novos desafios para a lexicografia, en *Cadernos de Tradução*, 43, pp. 33-62.
- Winestock, C. y Y. Jeong (2014). An analysis of the smartphone dictionary app market, *Lexicography*, 1(1), pp. 109-119.
- Wolf, G. (1996). Steve Jobs: The Next Insanely Great Thing, en *Wired*, 1st February, 1996. [www.wired.com/1996/02/jobs-2/](http://www.wired.com/1996/02/jobs-2/). (Consultado 12 de septiembre, 2019)



# MARCAS TEMÁTICAS EN LOS DICCIONARIOS ACADÉMICOS IBÉRICOS: ESTUDIO COMPARATIVO

## DOMAIN LABELS IN IBERIAN ACADEMY DICTIONARIES: A COMPARATIVE STUDY

Ana Salgado

Rute Costa

*Universidade Nova de Lisboa*

[anasalgado@campus.fcsh.unl.pt](mailto:anasalgado@campus.fcsh.unl.pt)

[rute.costa@fcsh.unl.pt](mailto:rute.costa@fcsh.unl.pt)

### RESUMEN

La actual revolución digital abre nuevos caminos en el ámbito de la elaboración de recursos lexicográficos, más concretamente en los diccionarios de lengua general que buscan dar respuesta, tanto en forma como en contenido, a nuevas necesidades de sus usuarios y de la sociedad en general. Como también lo hacen para el léxico general, estas obras registran, describen y definen léxico especializado de diferentes áreas del conocimiento. Así, el número de unidades terminológicas que forman parte de la nomenclatura de estos recursos tiende a aumentar debido al auge tecnológico, la evolución de la sociedad y los fenómenos de globalización, en la medida en que estas unidades constituyen fuentes privilegiadas de renovación y enriquecimiento lexicales de los sistemas lingüísticos. La marca temática, expresada mediante abreviaturas, además de apuntar al léxico especializado en los diccionarios de lengua general, sirve también como mecanismo de control de la terminología para los lexicógrafos. De este modo, las marcas temáticas que etiquetan el léxico especializado en diccionarios monolingües son objeto del estudio del presente trabajo cuya finalidad es contribuir a mejorar la interoperabilidad de los recursos. Por consiguiente, pretendemos describir y analizar las marcas temáticas en los diccionarios académicos de lengua general de referencia en la Península Ibérica: esto es, en el *Diccionario de la lengua española* (DLE) de la Real Academia Española (RAE) y en el *Dicionário da Língua Portuguesa Contemporânea* (DLPC) de la Academia das Ciências de Lisboa (ACL), con el fin de repensar los presupuestos teóricos y metodológicos de la tradición lexicográfica en torno a esas marcas. Con los resultados de nuestra investigación, nuestro propósito es destacar la necesidad de un acuerdo entre Academias para, partiendo de una armonización de las marcas temáticas, sugerir una marcación sistemática del uso especializado de las entradas lexicográficas. En la era digital, es necesaria una línea de trabajo común y abierta para optimizar el proceso de marcación temática y su codificación en los diccionarios, en tanto se revelan como bases de datos lexicales estructuradas, organizadas, con un formato estándar, accesibles e interoperables.

**Palabras clave:** lexicografía, metalexicografía, diccionarios de lengua general académicos, marca temática

### ABSTRACT

The current digital revolution has opened new possibilities into the construction of new lexicographic resources. This is especially true for general language dictionaries, whose purpose is to meet the needs, in form and content, of users and of society in general. Side by side with the general lexicon, these works register, describe, and define specialised lexicons used in different areas of knowledge. The number of terminological units in the entry lists of these resources has increased as a result of these technological changes, as result of the evolution of society, and a result of globalisation. This is because since these units are privileged sources of renewal and lexical enrichment of the linguistic systems. Domain labels are expressed using abbreviations. These labels identify specialised lexicons in general language dictionaries and serve as terminological control mechanisms for lexicographers. Domain labels that identify this specialised lexicon in monolingual dictionaries are the focus of the present paper. The aim of this is to help improve the interoperability of the resources being analyzed. Therefore, we intend to analyse, describe, and compare domain labels in the following general language dictionaries: the *Diccionario de la lengua española* (DLE), published by Iberian Academies, the *Diccionario de la lengua española* (DLE) published by Real Academia Española (RAE), and the *Dicionário da Língua Portuguesa Contemporânea* (DLPC), published Academia das Ciências de Lisboa (ACL). Our aim is to review the theoretical and methodological assumptions of the lexicographic tradition behind the use of domain labels. Using the findings from our research, we intend to propose an agreement between Academies in order to harmonize domain labels with the goal of instituting a consistent labelling system to identify the specialised usage of lexicographic entries. In the digital age, a common and open line is needed to optimise the domain labelling process and how it is encoded in dictionaries. A structured, organised, standardised, accessible, and interoperable lexical databases is needed.

**Keywords:** lexicography, metalexicography, general language Academy dictionaries, domain label.

Recibido: 23-05-2019  
Aceptado: 10-12-2019

## 0. INTRODUCCIÓN

Los diccionarios de lengua general reúnen, describen y consagran el léxico de la lengua. A la par del léxico general, estas obras registran y definen unidades terminológicas de diferentes áreas del conocimiento, por lo que estos dos grupos de unidades componen la nomenclatura de los diccionarios. Boulanger y L’Homme (1991), Wiegand (1984) y Ahumada (2002) afirman que las unidades terminológicas en los diccionarios de lengua general rondan el cuarenta y cincuenta por ciento. El número tiene tendencia a aumentar debido al auge tecnológico, la evolución de la sociedad y los fenómenos de la globalización. Las unidades terminológicas constituyen fuentes privilegiadas de renovación y enriquecimiento léxico de los sistemas lingüísticos y su identificación, estructuración y almacenamiento son fundamentales para la organización de datos cuya aplicación redefine a las Humanidades Digitales, entre otras. De este modo, la marcación que apunta al léxico especializado en diccionarios monolingües –marcas temáticas– es el centro del presente trabajo.

Con este artículo pretendemos averiguar la cantidad de marcas temáticas y la consistencia de su empleo a través de la descripción y análisis comparativo de las mismas –ciñéndose el presente trabajo tan solo a la marcación temática– en los diccionarios de lengua general académicos ibéricos: el *Diccionario de la lengua española (DLE)* de la Real Academia Española (RAE) y en el *Dicionário da Língua Portuguesa Contemporânea (DLPC)* de la Academia das Ciências de Lisboa (ACL), para repensar los presupuestos teóricos y metodológicos de la tradición lexicográfica en torno a las marcas lexicográficas.

Como la interoperabilidad<sup>1</sup> está en el centro de las preocupaciones de los lexicógrafos contemporáneos, nos detendremos en las marcas compartidas y las que no lo son en estos diccionarios y nos plantaremos la posibilidad de superar las

---

<sup>1</sup> La interoperabilidad se ha convertido en un concepto clave en las Humanidades Digitales para la preservación digital y para la publicación de recursos léxicos actuales. El término refiere a la capacidad de dos o más sistemas independientes para intercambiar información y utilizar esta información intercambiada de manera significativa: “The ability of two or more systems or components to exchange information and to use the information that has been exchanged” (Geraci, 1991, p. 114). Vid. también: <http://www.parthenos-project.eu/>.

diferencias con vistas a una posible reclasificación y codificación, homogénea y armónica, de las abreviaturas que corresponden a las marcas temáticas empleadas.

## 1. LA PROBLEMÁTICA DE LAS MARCAS TEMÁTICAS EN LOS DICCIONARIOS DE LENGUA GENERAL

El sistema de marcación (*labelling system*) representa uno de los temas más importantes y más delicados para los estudios metalexigráficos y para la lexicografía en general, tanto teórica como práctica. Entendemos por marcación el “recurso o procedimiento que se utiliza en el diccionario para señalar la particularidad de uso, de carácter no regular, que distingue a determinados elementos léxicos” (Fajardo, 1996-1997, pp. 31-32).

Los lexicógrafos enfrentan múltiples problemas relacionados tanto con la forma de anotar esta información como con el alcance del posible sistema de marcación y sus límites. Ptaszyński (2010, p. 411), en un artículo en el que investiga las causas del tratamiento teórico insatisfactorio de la información diasistemática en diccionarios, considera que los: “lexicographers have been searching in vain for an exhaustive and precise answer to the questions of which words to label in what kind of dictionaries and how to do it” (2010, p. 411). Según el mismo autor, estos problemas resultan de una “lack of a firm theoretical basis for the application of diasystematic information (i. e. information about restrictions on usage) in dictionaries” (2010, p. 411). En muchos casos, por falta de explicación en las introducciones, es difícil desvendar su valor real y se deduce que los lexicógrafos, en la mayoría de los casos, se limitan a reproducirlas siguiendo una cierta tradición. La cuestión del tratamiento de la información diasistemática en diccionarios de lengua general ya ha sido abordada en literatura anterior (Svensén, 2009; Atkins y Rundell, 2008; Bergenholtz y Tarp, 1995; Hausmann, 1989). Así, para Atkins y Rundell (2008, p. 231), “there’s quite a lot of work involved in putting together a consistent policy on labels in a dictionary”, a la vez que para Sakwa (2011, p. 308) “there is no agreed-on criteria for making usage decisions” y para Fedorova (2004, p. 265) “there is no consistency in the labeling policy”. En esta línea, a propósito de un estudio diacrónico de las marcas temáticas en los diccionarios de la Real Academia Española (RAE), Paz Battaner (1996, p. 104) considera que “la

presencia de marca temática parece aleatoria en la tradición académica, y en todas las que la siguen”.

Por estas razones, discutiremos sobre el concepto de marca y de marca temática antes de pasar al análisis de los diccionarios en observación.

## 2. EL CONCEPTO DE MARCA

El registro de informaciones relativas a las restricciones y limitaciones de empleo de una unidad léxica –especializada o no– es una práctica lexicográfica recurrente y muy antigua. Estas informaciones están presentes en prácticamente todos los diccionarios, usualmente por medio de abreviaturas (*ant.*, *pop.*, *vulg.*), o bajo la forma de una palabra simple que funciona como denominación de un dominio (*Arte*), y suelen aparecer en diferentes posiciones en la estructura del diccionario, generalmente antecediendo al texto definitorio propuesto para cada acepción (ver figura 1). En este caso, tenemos la marca temática o diatécnica “*Cul.*”, abreviatura para el dominio *culinária*, la marca diatópica o geográfica “*Region.*” para “regionalismo” y la marca de nivel o registro “*Fam.*” para “familiar”.

**canja** [kɛʒv]. *s. f.* (Do malaiala *kañji* 'arroz com água').  
**1. Cul.** Caldo de galinha com arroz. *A canja era a sua sopa preferida. «E então o coração das galinhas, que a Estefânia deixava ir na canja, juntamente com a moela?»* (L. DA COSTA, *Vovó Ana*, p. 211). **2.** Arroz cozido até ficar um caldo grosso ou uma papa. **3. Region. Gr.** Estado de embriaguez. ≈ BEBEDEIRA. **ser canja, Fam.**, ser coisa fácil de fazer ou obter; não oferecer dificuldades, nem problemas. *Para ele, resolver esse problema é canja. «Segui à risca os conselhos do meu Ali Babá. E, tal como previra, os palermas caíram na esparrela. Dali em diante foi canja.»* (A. VIEIRA, *Eleito*, p. 157).

Figura 1. Entrada “canja” (“sopa de pollo con arroz”) del DLPC (ACL)

En la figura 2, tenemos la marca temática “*Bot.*” para señalar el dominio *botânica*, ejemplificando un caso en el que la marca abarca todo el lema, es decir, las acepciones 1 y 2 de la unidad terminológica “*ulmáceas*” en entrada:

**ulmáceas** [ulmásje.f]. *s. f. pl.* (Do lat. *ulmus* 'olmeiro' + suf. *-áceas*). *Bot.* **1.** Família de plantas dicotiledóneas que compreende árvores ou arbustos de folhas simples, alternas, frequentemente assimétricas na base e providas de estípulas, flores hermafroditas ou unissexuais, dispostas em fascículos axilares ou solitárias, e à qual pertence o ulmo ou ulmeiro, existente em Portugal. **2.** *s. f.* Planta dessa família.

Figura 2. Entrada “ulmáceas” (“ulmáceas”) del DLPC (ACL)

A través de una mirada más atenta sobre los sistemas de marcación en obras lexicográficas –generalmente ubicadas en las páginas iniciales de las ediciones impresas o en una sección propia en páginas web– podemos observar la existencia de diversos tipos de marcas: marcas diacrónicas o temporales, marcas diatópicas o geográficas, marcas diatécnicas o temáticas, marcas de nivel o registro, marcas connotativas y marcas de frecuencia<sup>2</sup>. Coincidimos con Svensén (2009, p. 315) en que las marcas son fundamentales en tanto informan a los usuarios que “a certain lexical item deviates in a certain respect from the main bulk of items described in a dictionary and that its use is subject to some kind of restriction”. En la misma línea, Fajardo (1996-1997, p. 32) refiere que las marcas son “informaciones concretas sobre los muy diversos tipos de particularidades que restringen o condicionan el uso de las unidades léxicas”. De este modo, las marcas representan particularidades específicas en el contexto de uso de las unidades lexicales donde ‘desvío’ (por ejemplo, cuando el registro de lengua es familiar) y ‘restricción’ (si una determinada unidad pertenece a un dominio de especialidad) son conceptos clave para entender el concepto de marca.

En lo que concierne a la metalexigrafía o a la lexicografía teórica, se verifica una gran variación en torno a la designación del concepto de marca. Veamos algunos de los términos usados por diferentes investigadores: Atkins y Rundell (2008) las denominan *linguistic labels*, Svensén (2009) y Haussmann (1989) prefieren usar la

---

<sup>2</sup> Para cada una de las diferentes marcas, presentamos algunos ejemplos de entradas extraídos del DLPC: marcas diacrónicas o temporales (“beque”, *ant.*, “antiguado”), marcas diatópicas o geográficas (“parabenizar”, *Bras.*, “Brasil”), marcas diatécnicas o temáticas (“linfoma”, *Med.*, “Medicina”), marcas de nivel o registro (“paleio”, *fam.*, “familiar”), marcas connotativas (“maralha”, *deprec.*, “depreciativo”), marcas de frecuencia (“per”, *desus.*, “desusado”).

designación *diasystematic marking* como sinónimo de *diasystematic information*, Monson (1973) las designa como *restrictive labels* y Landau (2001) opta por *usage information*.

### 3. REVISITA DE LAS CLASIFICACIONES DE LA INFORMACIÓN DIASISTEMÁTICA

Antes de analizar las marcas temáticas que figuran en nuestro corpus lexicográfico, es importante visitar el marco teórico en torno a las principales clasificaciones de la información diasistemática en obras lexicográficas.

Recorriendo la literatura existente, encontramos diferentes clasificaciones de la información diasistemática propuestas por diferentes investigadores. La clasificación más desarrollada es la de Hausmann (1989, p. 651), que identifica once tipos de marcas. Esta organización la han adoptado otros autores como Bergenholtz y Tarp (1995, pp. 131-134) y Svensén (2009, pp. 326-332). Atkins y Rundell (2008, pp. 182-186), en cambio, distinguen nueve tipos de marcas lingüísticas (*linguistic labels*, en su terminología). Landau (2001, pp. 217-272) presenta ocho tipos distintos de información. Jackson (2002, pp. 109-115), a su vez, enumera siete tipos de marcas de uso (*usage labels*, en su terminología). Un resumen de las diferentes propuestas de clasificación puede consultarse en la tabla que se presenta a continuación:

Hausmann (1989)	Landau (2001)	Jackson (2002)	Atkins y Rundell (2008)	Ejemplos de marcas
diachronic	currency or temporality	History	time	<i>antiguado, arcaico</i>
diatopic	regional or geographic variation	dialect <sup>3</sup> [geographical restriction]	region dialect	<i>Brasil, América</i>
diintegrative	–	–	–	–
diamedial	style, functional variety, or register	–	–	–
diastratic	restricted or taboo sexual scatological usage and slang	Status	slang and jargon offensive terms	<i>calão, vulgar</i>
diaphasic	style, functional variety, or register	Formality	register	<i>formal, informal</i>

<sup>3</sup> A pesar de hablar de “*dialect*”, afirma que “dialect labels refer to geographical restrictions, and we can take this to include both national varieties and regional dialects within a national variety” (Jackson, 2002, p. 110).

diatextual	style, functional variety, or register	–	style	<i>literário, poético</i>
<b>diatechnical</b>	<b>technical or specialized terminology</b>	<b>topic or field</b>	<b>domains</b>	<b><i>anatomía, desporto, medicina</i></b>
diafrequent	–	–	–	<i>pouco usado, raro</i>
diaevaluative	insult <sup>4</sup> style, functional variety, or register	Effect	attitude	<i>eufemismo, irónico, ofensivo</i>
dianormative	status or cultural level	disputed usage	–	–

**Tabla 1.** Clasificaciones de la información diasistemática propuestas por diferentes investigadores

A pesar de todo este esfuerzo teórico de clasificación, ninguno de los autores presenta reglas o plantea la representación de información diasistemática en diccionarios, lo que sería muy útil para los lexicógrafos para el tratamiento cuantitativo y cualitativo de los datos.

Por último, conviene también poner de relieve que hay investigadores que utilizan un concepto más amplio cuando hacen referencia a marcas. En la metalexigrafía española, el lexicógrafo Porto Dapena (2002, p. 250) considera, por ejemplo, las categorías gramaticales como marcas lexicográficas, no ciñéndose solo a la idea de desvío y de trazos restrictivos: “nosotros preferimos partir de un concepto más amplio que incluya no solo rasgos restrictivos, sino de cualquier otro tipo, como por ejemplo la pertenencia a una determinada categoría y subcategoría gramatical o semántica”. Fajardo (1996-1997, p. 388), en cambio, al asumir que las marcas son informaciones que solo aparecen cuando se verifican particularidades que restringen o condicionan el uso de la unidad léxica, no considera las indicaciones de la categoría gramatical tras cada lema como una marca, una vez que “fuera del concepto de marcación todo lo que es regular y constante en cada uno de los artículos del diccionario” (1996-1997, p. 388).

Estas tensiones nos obligan a asumir, en el ámbito de este artículo, la distinción de los términos ‘marca’, que será usado para hacer referencia a las abreviaturas (*Anat.*,

<sup>4</sup> Landau (2001) divide la “*diaevaluative information*” en dos grupos: aquellos que denotan una connotación positiva (*humorous*) pertenecientes a “style, functional variety, or register” y los de connotación negativa (*offensive*) pertenecientes a esta división y designadas por “insult”.

*Filos., Med.*), y ‘dominio’ para mencionar las designaciones por extenso de cada una de las abreviaturas (*Anatomía, Filosofía, Medicina*).

#### 4. EL CONCEPTO DE MARCA TEMÁTICA

En lo que atañe a la información diasistémica presente en diccionarios de lengua general, podemos especificar el conjunto de informaciones –las marcas temáticas– que indican si una unidad es terminológica, o sea, si procede de un área específica de actividad o saber humanos. También para *marca temática* existe variación en su designación: Atkins y Rundell (2008, p. 182), refiriéndose a *linguistic labels*, clasifican el vocabulario especializado como *domains; field labels*, según Janssen *et al.* (2003, p. 7); *marcas técnicas* para Fajardo (1996-1997) o *marcas temáticas* en Estopà (1998). En el marco de nuestra investigación preferimos el término *marcas temáticas* porque nos parece una designación transparente y reconocible no solo para los lexicógrafos, sino también para los terminólogos.

Estas marcas sirven “para señalar el léxico temáticamente especializado, en contraposición al léxico común”, y aparecen expresadas generalmente mediante abreviaturas (recordando la necesidad de economía en formato papel) y marcas diatécnicas (Hausmann, 1989, Svensén, 2009).

Pero, en rigor, ¿para qué sirve la marca temática y qué pretende marcar? Ya Atkins y Rundell (2008, p. 227) se hicieron esa pregunta: “What does a label label?”. La respuesta es múltiple. La marca puede remitir a informaciones de naturaleza distinta (las ya referidas marcas diatécnicas, diacrónicas, diatópicas, entre otras), pero también es usada por los lexicógrafos para marcar la pertenencia a un dominio, reduciendo de inmediato las posibilidades de interpretación, haciendo posible que el usuario localice un sentido especializado. En la era digital, además, “domain labels have an important role to play in lexical databases [...] where the domain label is useful in word sense disambiguation.” (Atkins y Rundell, 2008, p. 227). Por otra parte, “in publisher’s databases, these labels offer a way of automating lists of specialized vocabular which can be exploited in a number of ways” (Atkins y Rundell, 2008, p. 227). Pero, si las

marcas pueden ayudar a un usuario a buscar una unidad léxica, también funcionan para el lexicógrafo como mecanismo de control de la terminología.

Como ya referimos, si este es un tema delicado cuando observamos las marcaciones dentro de un determinado recurso lexicográfico, la dificultad aumenta cuando comparamos diferentes recursos. No todos usan las mismas marcas y su número, a veces, es bastante dispar.

Recordando y reforzando nuestra premisa inicial: “there’s quite a lot of work involved in putting together a consistent policy on labels in a dictionary.” (Atkins y Rundell, 2008, p. 231).

## 5. LAS MARCAS TEMÁTICAS EN LOS DICCIONARIOS ACADÉMICOS IBÉRICOS

El surgimiento de varias Academias de la Lengua está condicionado por una premisa idéntica: redactar un diccionario con la intención de conservar y perfeccionar la lengua, reglamentar el uso, el vocabulario y la gramática de los idiomas<sup>5</sup>. Un diccionario, en tanto “objeto cultural”, es una forma de afirmación de una nación y constituye una parte significativa de la “memoria cultural” (Ahumada, 2002, p. 20; Rey, 2008, p. 120).

Un diccionario de lengua puede considerarse una obra de referencia en la medida que pretenda representar lo más fielmente posible la norma de la comunidad lingüística a la cual se destina. En palabras de Rey (1983, p. 558), “la fonction du dictionnaire est de fournir à ses usagers une référence sur la norme”, por lo que los diccionarios académicos representan un síntoma de la fijación de esa norma.

El corpus lexicográfico aquí en observación está constituido por dos diccionarios elaborados por diferentes instituciones académicas: RAE (*DLE*) y ACL (*DLPC*) cuyo contenido diccionarístico está en las lenguas de origen de cada una de las instituciones. Son diccionarios de lengua general, contemporáneos, con ediciones impresas de orientación sincrónica, de naturaleza descriptiva, con una preocupación normalizadora y dirigidos a un público amplio cuya lengua materna es el español y el portugués,

---

<sup>5</sup> Aunque muchas de las decisiones académicas no tengan, en realidad, fuerza de ley y que, en ciertos casos, las autoridades no consideren sus decisiones.

respectivamente. Ambos diccionarios se encuentran *online*<sup>6</sup> y están siendo objeto de actualizaciones.

El *DLE* es el diccionario normativo más amplio del español editado y elaborado por la RAE. Su primera edición data de 1780, siendo la más reciente, 24.<sup>a</sup> edición, de 2014<sup>7</sup>, en coincidencia con la conmemoración del tricentenario de su fundación. Desde el 2005, puede ser consultado *online*<sup>8</sup>. Según se lee en su “Preámbulo” (*DLE*, p. XI) de la 23.<sup>a</sup> edición, la nomenclatura del *DLE* está constituida por “93 111 entradas con un total de 195 439 acepciones”<sup>9</sup>.

En Portugal, a pesar de sucesivas tentativas académicas, no fue hasta el siglo XXI, (en el año 2001, más concretamente) cuando la ACL publicó un diccionario completo, el DLPC, en una versión en papel en dos volúmenes. Un equipo de la Universidad de Minho<sup>10</sup> está trabajando en una actualización en versión digital de este diccionario con la participación del NOVA CLUNL<sup>11</sup>. La nomenclatura del DLPC está constituida por un total de 69 426 entradas, 167 556 acepciones.

En este trabajo, no es nuestra intención cuestionar la exactitud de la información de los datos, sino analizar la forma en que estos son presentados.

Así, pasamos ahora a describir detalladamente nuestro planteamiento metodológico:

---

<sup>6</sup> El DLPC solo está disponible para usuarios registrados, pero la primera autora de esta investigación es coordinadora del nuevo diccionario de la ACL que ya se encuentra en base de datos.

<sup>7</sup> Para este trabajo la primera autora ha podido consultar las abreviaturas registradas en la base de datos del diccionario gracias a una beca concedida por ELEXIS en noviembre de 2018 (12/11 – 30/11/2018). Vid. <https://elex.is/ana-de-castro-salgado/>.

<sup>8</sup> <https://dle.rae.es>

<sup>9</sup> Para la presentación de los datos estadísticos del *DLE*, agradecemos al Ilex (Instituto de Lexicografía de la RAE). En noviembre de 2018, la base del *DLE* contaba con 95 410 entradas, del total de 198 176 acepciones.

<sup>10</sup> El equipo para el Procesamiento del Lenguaje Natural de la Universidad de Minho ha desarrollado el soporte digital para la nueva edición digital del *Diccionario académico de portugués*, con la participación de Alberto Simões del IPCA (Instituto Politécnico do Cávado e do Ave), responsable del soporte tecnológico, con la colaboración de José João Almeida, y la asesoría de Álvaro Iriarte Sanromán, ambos investigadores de la Universidad de Minho

<sup>11</sup> El nuevo diccionario digital se rige por las TEI *Guidelines* (<https://tei-c.org/guidelines>). La participación del NOVA CLUNL se relaciona con la transición del formato para TEI Lex-0, una propuesta adaptada específicamente para la codificación de recursos lexicográficos.

### 5.1. PÁGINAS INTRODUCTORIAS

El primer paso en nuestra metodología fue la lectura de las páginas de introducción de las dos ediciones impresas para detectar tanto referencias al tratamiento del léxico especializado como a criterios o, incluso, justificaciones sobre su registro.

Tras la lectura del “Prólogo” de la edición impresa del *DLE*, verificamos que no hay ninguna mención al léxico especializado, lo que nos lleva a creer que el léxico de lengua especializada sufrió el mismo tratamiento que el léxico general. Tan solo en la sección “Advertencias” (pp. LI-LIII) se encuentra una breve alusión a las marcas temáticas con la intención de informar acerca de las decisiones tomadas por los lexicógrafos en la ordenación de las acepciones en los artículos lexicográficos, en los que, lógicamente, se da prioridad a los registros de lengua:

De marcación: las acepciones no marcadas tienden a anteponerse a las marcadas. Dentro de estas, van primero las acepciones que tienen marcas correspondientes a los niveles de lengua o registros de habla, después las que llevan marcas técnicas, después las que tienen marcas geográficas (y dentro de ellas, primero las de España y luego las de América y Filipinas) y finalmente las que llevan una marca de vigencia. (*DLE*, p. LII)

A su vez, en la “Introdução” del *DLPC* (pp. XIII-XXIII), tampoco encontramos ninguna referencia a posibles criterios orientadores para la selección de léxico especializado por parte de los equipos de lexicógrafos responsables, ni tampoco ninguna justificación relativa a la inclusión de áreas temáticas ni al empleo de las marcas respectivas, sino solo una descripción, muy general, del tipo de unidades especializadas que son objeto de registro:

No Dicionário registam-se ainda: tecnicismos generalizados na linguagem usual; tecnicismos que, embora de uso não generalizado, correspondem a noções ou classificações e a aparelhos fundamentais em cada ciência ou técnica; tecnicismos que ocorrem em manuais escolares de natureza científica e técnica. (*DLPC*, p. XIV).

### 5.2. LISTAS DE ABREVIATURAS

El rastreo exhaustivo de las marcas temáticas permitió determinar la cantidad de dominios representados en los dos diccionarios, tanto los exclusivos como los compartidos y sus frecuencias. Hemos analizado si el uso de las marcas era sistemático, o no, y si había dominios recientes y relevantes que no figurasen. Como no se

encontraron criterios relativos al empleo de las marcas, nos vimos forzadas a hacer suposiciones.

A través del análisis llevado a cabo, se constató que los criterios para selección y tratamiento del léxico especializado difieren entre los diccionarios: en el *DLE*, las marcas temáticas se emplean cuando la acepción no se considera que pertenezca al léxico común; en el *DLPC*, la marcación parece limitarse a precisar el dominio de la acepción. Consúltese, por ejemplo, las entradas “corazón” (*DLE*) y “coração” (*DLPC*). En el *DLPC*, la marca temática para el dominio *Anatomia* está presente, pero, en el *DLE* la entrada no presenta ninguna marcación, quizás porque los lexicógrafos la pueden haber considerado como perteneciente al léxico general<sup>12</sup>. A su vez, observamos otros recursos lexicográficos que apuntan a contextos especializados como, por ejemplo, la introducción de expresiones restrictivas en el texto definicional.

A pesar de existir poca o ninguna información concerniente a criterios para la selección, inclusión y marcación de léxico especializado, las secciones “Abreviaturas y signos empleados” (*DLE*, pp. LV-LVIII) y “Classificação do vocabulário quanto à repartição por ciências, técnicas e formas de actividade” (*DPLC*, pp. XXXV-XXXVI) demuestran claramente que estos diccionarios de lengua general hacen una cobertura del léxico especializado. En el caso del *DLPC*, su título nos obliga a preguntarnos cuál sería la distinción que los redactores de la lista del *DLPC* harían entre “ciencias”, “técnicas” y “formas de actividad”. Como encontramos los dominios *Alveitaria*, *Alvenaria* o *Cutelaria* creemos que esa pueda ser la razón para el empleo de “formas de actividad”, por el hecho de que los lexicógrafos redactores del *DLPC* tal vez no consideren que estos dominios puedan ser realmente dominios de especialidad. La ausencia de fundamentación de las opciones de los lexicógrafos redactores no nos permite ir más allá de la deducción.

Por otra parte, aunque ambas lenguas apunten a un mismo sistema de marcación, las marcas empleadas para el efecto no siempre coinciden, como veremos más adelante. Si el *DLE* enumera todas las marcas usadas en una única lista exhaustiva de

---

<sup>12</sup> Pese al interés de este tema, en este artículo no abordaremos su estudio por alejarse del objetivo que nos hemos propuesto.



	Abreviaturas	Tipografía	Justificación
DLE	74	cursiva, minúscula	no
DLPC	184	cursiva, mayúscula	no

Figura 4. Cuadro comparativo del número de marcas temáticas, tipografía y justificación de criterios

Así también, los dominios encontrados en el *DLE* pueden observarse en la figura 5:



Figura 5. Marcas temáticas en el *DLE* (74)

En comparación, las marcas temáticas del *DLPC* registran un número mucho más elevado. Estando, actualmente, la información disponible en base de datos (ver nota 6), ha sido posible verificar que hay marcas que, aunque están listadas en el *DLPC*, no presentan ninguna ocurrencia en el diccionario, como es el caso de los dominios *Bromatología*, *Cibernética*, *Economía Política*, *Escolástica*, *Espiritualismo*, *Futurología*, *Policia*, *Química Biológica*, *Química Orgánica*, *Telefonía Sem Fio* y *Velocipedia*. Por ello, al número total de 184 dominios habría que restarles estos 11 dominios ausentes en el *DLPC*, por lo que el número pasa a 173. Por otro lado, también se detectaron varias marcas que no se encuentran en la lista de abreviaturas y que se encuentran en la base de datos,

como *Bioquímica, Etnología, Metrología, Agronomía, Marítima, Psicanálise, Ecología, Ginástica, História Política, Pirotecnia* y *Transportes*. De este modo, si a los 173 dominios que reconocíamos, les sumamos estos ausentes en la lista inicial, el resultado vuelve a ser de 184 dominios. Creemos que estas discrepancias pudieron ser lapsus cometidos en la publicación del *DLPC*.

La totalidad de los dominios encontrados en el *DLPC* pueden observarse en la figura 6:

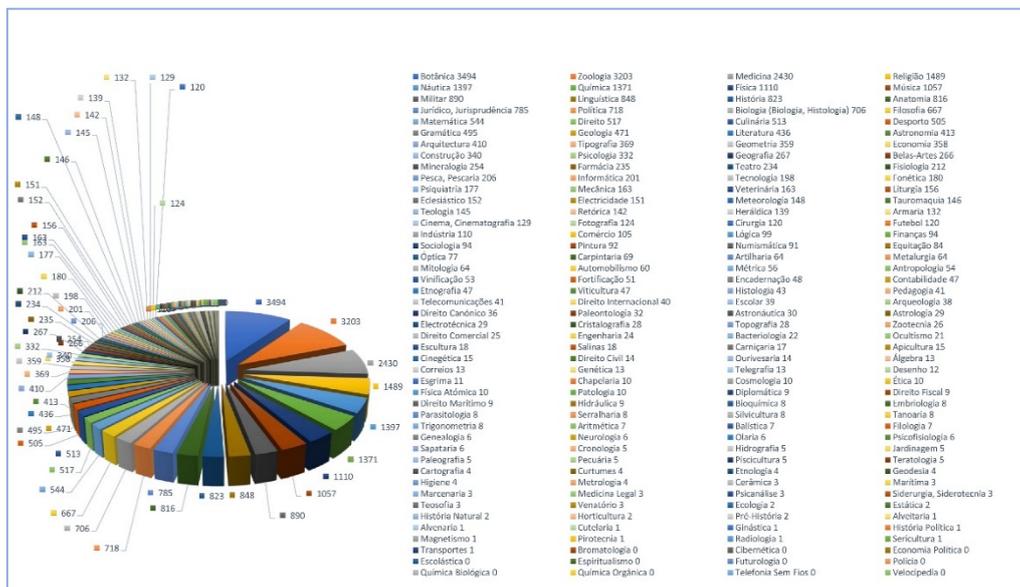


Figura 6. Marcas temáticas en el *DLPC* (184)

Se observa la coexistencia de dominios genéricos y subdominios, como, por ejemplo, es el caso de *Direito* “Direito Canónico”, “Direito Civil”, “Direito Comercial”, “Direito Fiscal”, “Direito Internacional”, “Direito Marítimo” o *Química*: “Química Biológica” y “Química Orgânica” o *Matemática* y sus subdominios “Geometria”, “Álgebra”, “Aritmética” y “Trigonometria”. La proliferación de subdominios de un dominio general hace que haya una mayor presencia de marcas en el *DLPC* y se constate una diferencia de 110 dominios respecto al *DLE*.

Como vemos, el desequilibrio del número total de dominios entre los dos diccionarios es notorio. Ese desequilibrio puede explicarse en que mientras el *DLE* presenta solo dominios genéricos, el *DLPC* especifica algunos de los dominios sin entenderse su razón. La ausencia de explicación de criterios para el establecimiento de

límites no nos permite entender esta opción portuguesa, pudiéndose solo deducir la ausencia de una previa organización del conocimiento especializado, deliberada o no. No detectamos una sistematicidad en la inclusión de los dominios que constan en la lista de abreviaturas del *DLPC*. Una explicación posible para los criterios de selección y uso de estas marcas especializadas nos hablaría de una toma en consideración de las necesidades del usuario o de las preferencias del equipo de lexicógrafos que han trabajado en los diccionarios.

No pensamos que la organización entre dominios y subdominios patente en el *DLPC* presente ventajas porque según nuestro criterio aumenta la posibilidad de hallarse errores que pueden afectar a la coherencia interna de la obra, además de dificultar el trabajo del lexicógrafo ante la ausencia de criterios de marcación. Una clasificación más general salvaguardará el trabajo del lexicógrafo y contribuirá a la coherencia de la descripción de las unidades lexicales y terminológicas.

Si seguimos comparando ambas obras lexicográficas, nos encontraremos con las abreviaturas utilizadas para representar las áreas temáticas.

Las abreviaturas escogidas por los lexicógrafos de estos diccionarios para representar una misma área no siempre coinciden (por ejemplo, “Mús.” para el dominio *Música*; pero, para el dominio *Acústica*, son usadas las abreviaturas “Acús.” en el *DLE* y “Acúst.” en el *DLPC*). Aunque somos completamente conscientes de que llevamos a cabo una comparación entre dos recursos lexicográficos de lenguas diferentes, la proximidad entre las dos lenguas hace deseable proponer la relevancia de una convención homogénea de ciertas marcas temáticas. Como vemos en la figura 7, hay 44 marcas y denominaciones comunes (70 %) y 19 marcas y denominaciones diferentes (30 %) entre las dos obras que sería conveniente revisar.

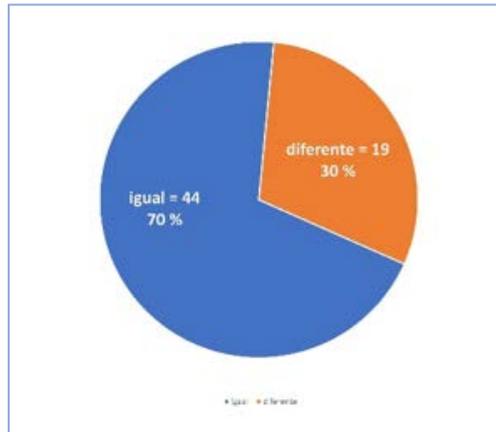


Figura 7. Correspondencia entre las marcas temáticas

Al mismo tiempo, más allá de las marcas usadas, contabilizamos 63 dominios compartidos (ver figura 8).

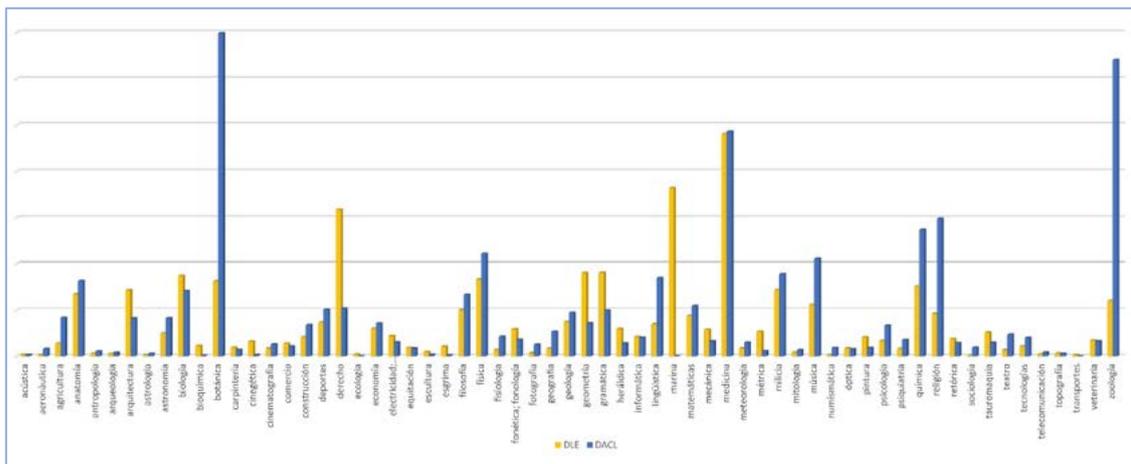


Figura 8. Dominios comunes entre los diccionarios

Se da el caso, también, de la presencia de dominios compartidos pero sin que coincidan sus marcas. Véase, por ejemplo, el caso del dominio *Acústica*, que es marcado “Acús.” en el *DLPC* y “Acúst.” en el *DLE*, “Ret.” y “Retór.”, respectivamente, en el caso del dominio *Retórica*. Una lista de los dominios compartidos con marcas diferentes puede ser observada en la figura 9:

N.º	DLE		DLPC	
1	Acús.	acústica	Acúst.	Acústica
2	Aer.	aeronáutica	Aeron.	Aeronáutica
3	Antrop.	antropología	Antr.	Antropologia
4	Arq.	arquitectura	Arquit.	Arquitectura
5	Com.	comercio	Comérc.	Comércio
6	Dep.	deportes	Desp.	Desporto
7	Der.	derecho	Dir.	Direito
8	Esc.	escultura	Escult.	Escultura
9	Fil.	filosofía	Filos.	Filosofia
10	Fon.	fonética; fonología	Fonét.	Fonética
11	Fotogr.	fotografía	Fot.	Fotografia
12	Geogr.	geografía	Geog.	Geografia
13	Mec.	mecánica	Mecân.	Mecânica
14	Mit.	mitología	Mitol.	Mitologia
15	Psiquiatr.	psiquiatría	Psiq.	Psiquiatria
16	Ret.	retórica	Retór.	Retórica
17	Teatro	teatro	Teat.	Teatro
18	Telec.	telecomunicación	Telecom.	Telecomunicações
19	Topogr.	topografía	Topog.	Topografia

Figura 9. Marcas diferentes y denominaciones comunes

Las siete áreas de conocimiento con mayor y menor número de entradas marcadas diatómicamente se presentan en los gráficos siguientes (figuras 10 y 11):

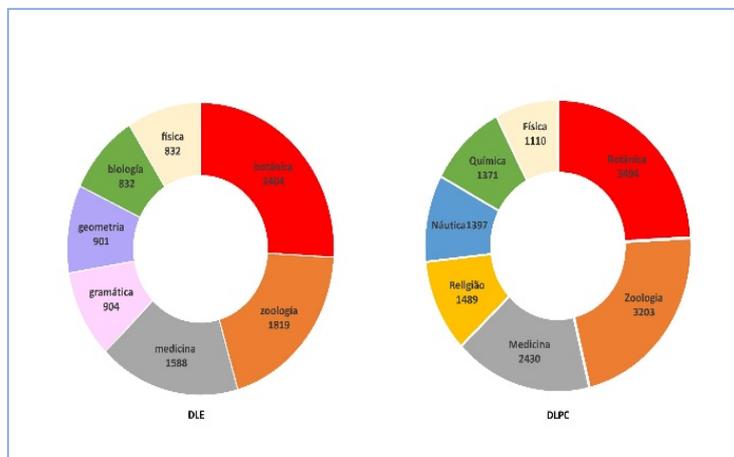


Figura 10. Dominios más frecuentes entre los dos diccionarios

Los dominios clásicos, como *Botánica/Botânica* (2404 entradas *DLE* vs. 3494 *DLPC*), *Medicina* (2404 entradas *DLE* vs. 2430 *DLPC*) y *Zoología/Zoologia* (1819 entradas *DLE* vs. 3203 entradas *DLPC*) son los más frecuentes en estos diccionarios, previsible en tanto continúan con una tradición lexicográfica.

Por otra parte, es necesario cuestionar la presencia de dominios con menor representación, como aquellos que aparecen 1 o 2 veces, como es el caso de todos los

dominios portugueses registrados en el gráfico de la figura 11 y el caso de *Ortografía* del *DLE*, respectivamente.

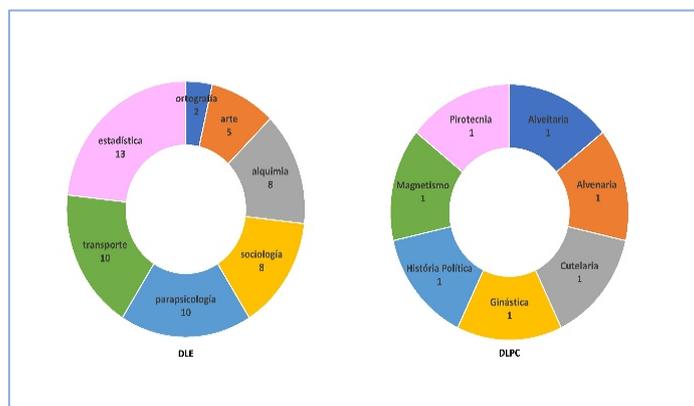


Figura 11. Dominios menos frecuentes entre los dos diccionarios

Si la lista de abreviaturas suele ordenarse alfabéticamente de manera convencional –un recurso eficaz usado para facilitar la localización de una marca dada–, defendemos, sin embargo, una organización previa conceptual de las marcas respectivas y descodificación de su valor respectivo. Como afirma Costa (2013), “Specialised communication, whether monolingual or multilingual, is not solely a matter of language, it is also a matter of knowledge”. El carácter bidimensional de la Terminología, al vincular un componente conceptual –fortalecido por la organización del conocimiento– a un componente lingüístico, permitirá una mejor descripción de artículos lexicográficos.

Por otra parte, ante la ausencia de explicación en las páginas iniciales de los diccionarios impresos sobre sus marcas temáticas, decidimos buscar información en otro tipo de recursos que usan marcas temáticas, fundamentales para sus objetivos. Así, contrastamos de qué manera el EuroVoc<sup>13</sup> y el Tesouro de la UNESCO<sup>14</sup> organizan sus descriptores para poder establecer analogías.

El EuroVoc es un tesouro multilingüe y multidisciplinar de la UE, dividido en 21 dominios y 127 subdominios. Sin embargo, la organización del conocimiento especializado, según el Tesouro de la UNESCO, se estructura en 7 temas principales:

<sup>13</sup> <https://eur-lex.europa.eu/browse/eurovoc.html?locale=pt>

<sup>14</sup> <http://vocabularies.unesco.org/browser/thesaurus/es/>

Educación, Ciencia, Cultura, Ciencias Sociales y Humanas, Información y Comunicación, Política, Derecho y Economía.

En el EuroVoc, por ejemplo, se destacan algunos dominios que no se registran en los diccionarios, por ejemplo: *Medioambiente*, *Tecnología*, *Energía*, *Industria*, o *Educación*. Observando esta breve lista de dominios, se hace evidente el hecho de que ninguno de estos dominios está presente en los diccionarios en estudio. ¿Por qué razón no se encuentra un dominio como el del *Medioambiente*? ¿Será por la intersección entre la terminología de esa área con otros dominios, como la *Ecología* o la *Biología*, por ejemplo? Otro caso: la *Tecnología*. En este caso, nos atrevemos a avanzar con la hipótesis de que este dominio no figura en los diccionarios estudiados por considerarse que sus unidades terminológicas son ya corrientes en el léxico general, ya que hoy en día es normal que un hablante de portugués o español integre vocablos como GPS o *wifi* en su discurso cotidiano. En otras palabras, esas mismas unidades terminológicas estarán, seguramente, registradas en los diccionarios, pero ya no llevan marca temática. Desde nuestra perspectiva creemos que para dar una respuesta homogénea a estas y otras cuestiones es necesario un trabajo previo de organización de una lista de dominios para proporcionar respuestas científicamente rigurosas.

Esta falta de criterio previo haría cuestionable que en los diccionarios de lengua general queden fuera dominios recientes como los arriba mencionados de gran relieve en la sociedad actual, mientras permanecen dominios perfectamente obsoletos.

Como sostienen Atkins y Rundell (2008, p. 184), en vez de concebir “a totally ‘flat’ (non-hierarchical list of domains)”, “it is more practicable to try build a domain list with a certain hierarchical structure, so that instead of ‘physics’, ‘chemistry’, etc., you have ‘science: physics’, ‘science: chemistry’, and so on”. Encontramos, así, ventajas, en la aplicación de una estructura previamente organizada tanto en la fase de composición de un recurso lexicográfico como durante la fase de edición ya que facilita al lexicógrafo el control de la terminología. Se podrá, de este modo, garantizar que no hay “glaring omissions” y “mark vocabular items more accurately” (ibidem). “Those which are common to several domains can receive the ‘higher level’ domain marker, so that items

like test tube and laboratory may be labelled ‘science’ rather than ‘physics, chemistry, biology’, and son on” (pp. 184-185).

Podemos ejemplificar nuestras ideas con el análisis de algunos artículos lexicográficos de especialidad relacionados con el dominio Fútbol.

#### 5.4. ANÁLISIS DE ARTÍCULOS LEXICOGRÁFICOS – EL DOMINIO DEL FÚTBOL

Como el estudio exhaustivo del léxico de especialidad y su tratamiento son tareas demasiado extensas para los límites de este trabajo, optamos por elegir un dominio para análisis de artículos lexicográficos: el Fútbol.

Del corpus lexicográfico constituido, seleccionamos aleatoriamente algunas unidades terminológicas relacionadas con el Fútbol. En el *DLE*, no existe marca temática para “Fútbol”, pero sí “Deportes”, donde las unidades terminológicas pertenecientes al fútbol están tratadas. De las 1915 entradas marcadas con “*Dep.*”, seleccionamos 147 entradas en las que el término “fútbol” aparece en la información de los artículos lexicográficos. En el *DLPC*, el criterio es otro ya que la marca temática “*Fut.*” existe como un subdominio del “Desporto” lo que permite identificar 120 entradas lexicográficas con esa marcación.

En primer lugar, observamos la posición de las marcas “*Dep.*” (*DLE*) y “*Fut.*” (*DLPC*) dentro del artículo lexicográfico.

Consultamos, ahora, dos entradas lexicográficas del *DLE*, a saber “extremo” y “lateral”:

- (1) **extremo**, **ma** m. y f. 14. En el fútbol y otros deportes, miembro de la delantera que, en la alineación del equipo, se sitúa más próximo a las bandas derecha o izquierda del campo
- (2) **lateral** adj. 3. Dicho de un futbolista o de un jugador de otros deportes: Que actúa junto a las bandas del terreno de juego con funciones generalmente defensivas. U. t. c. s.

Estas unidades, que podrían ser vistas como unidades terminológicas, no se encuentran marcadas diatécnicamente, optando los lexicógrafos por otro mecanismo (Porto Dapena, 2002, p. 308): la introducción de expresiones restrictivas en el texto definicional (por ejemplo, en (1) “En el fútbol y otros deportes”, subrayado en nuestro ejemplo). Las unidades “extremo” y “lateral” no son terminológicas, sino unidades lexicales de uso corriente no especializado. En este sentido, del corpus analizado se

concluye que la opción del *DLE* es marcar temáticamente las unidades solo cuando la acepción es específicamente pertenencia de un dominio, teniendo otro empleo fuera del contexto especializado. De esto son ejemplo, *ariete*, *autogol*, *estirada* que se señalan con el dominio *Deporte*.

En el *DLPC*, la presencia de marcas temáticas es significativa. Se verifican tres situaciones distintas:

i) la marca temática figura después de la entrada, así, todo el contenido está abarcado por esa marcación (“águia”, “benfiquista”);

(3) **águia**<sup>4</sup> *s. m. e f. Gír. Fut.* 1. Clube desportivo Sport Lisboa e Benfica, que tem a águia como insígnia. 2. O m. que benfiquista<sup>2</sup>.

ii) la marca temática figura después de los números de acepciones por lo que esa marcación solo se refiere a una acepción particular de manera que también sirve para diferenciar explícitamente, por ejemplo, casos de polisemia (“chapéu”);

(4) **chapéu** *s. m.* 13. *Fut.* Remate que descreve um arco por cima da cabeça do guarda-redes.

iii) también la marca temática puede situarse antes de las unidades léxicas complejas, como colocaciones (“grande penalidade”, “pontapé de baliza”).

(5) **grande**<sup>1</sup> *adj. m. e f. grande penalidade, Fut.*, castigo aplicado por falta grave, cometida na grande área e que se traduz por um remate directo à baliza, a partir de onze metros de distância, sendo a bola apenas defendida pelo guarda-redes.

De lo expuesto, se puede concluir, por tanto, que la posición de la marca no es aleatoria. Esta puede abarcar la totalidad de una entrada lexical (3) o solo sentidos particulares (4, 5).

Sin embargo, se constata que la marcación no siempre es regular. Encontramos artículos donde las marcas son especificadas y otros donde no se indica ningún uso temático. De este modo, dentro de la microestructura del diccionario, no siempre se hace un uso sistemático de una determinada marca. Véase, por ejemplo, el tratamiento de las mismas unidades anteriormente analizadas en el *DLE*, pero ahora en el caso portugués:

- (6) **extremo**:<sup>2</sup> *n. m.* 9. *Desp.* Jogador de futebol, basquetebol... que actua junto à linha lateral.  
 (7) **lateral**:<sup>2</sup> *n. m. e f.* 2. *Fut.* Jogador que actua junto da linha lateral do campo. + direito, esquerdo.

En este caso, para el ejemplo (6), el lexicógrafo no usa la marca “*Fut.*” probablemente porque en la definición detalla los diferentes tipos de deporte que hacen uso del término<sup>15</sup>.

Parece evidente que el *DLPC* se distancia del *DLE* al usar la marca temática para diferenciar acepciones o contextualizarlas, limitándose a precisar el dominio de la acepción. No nos cabe hacer consideraciones sobre los diferentes criterios. En realidad, cualquier criterio puede ser validado siempre y cuando se aplique uniformemente.

Continuando con el examen de las entradas relacionadas con el dominio *Fútbol*, analizaremos y compararemos ahora el comportamiento de algunas unidades que pertenecen al campo conceptual de ‘adepto’:

<i>DLE</i>	<i>DLPC</i>
(8) <b>colchonero, ra</b> 2. <i>adj. coloq.</i> Perteneciente o relativo al Atlético de Madrid, sociedad deportiva española. 3. <i>adj. coloq.</i> Jugador o seguidor del Atlético de Madrid. U. t. c. s.	(10) <b>gilista</b> : <sup>1</sup> <i>adj. m. e f.</i> 1. Que pertence ou diz respeito ao clube desportivo Gil Vicente Futebol Clube. « <i>A equipa profissional sénior gilista voltou, ontem de manhã bem cedo, ao trabalho</i> » ( <i>A Bola</i> , 29.3.1997) 2. Que é jogador, adepto ou simpatizante desse clube.
(9) <b>merengue</b> 5. <i>adj. coloq.</i> Perteneciente o relativo al Real Madrid Club de Fútbol. 6. <i>adj. coloq.</i> Jugador o seguidor del Real Madrid Club de Fútbol. U. t. c. s.	(11) <b>leão</b> : <sup>4</sup> <i>n. m. Gír. Fut.</i> O m. q. sportinguista:2 (12) <b>portista</b> : <sup>1</sup> <i>adj. m. e f. Fut.</i> 1. Que diz respeito ou pertence ao clube desportivo Futebol Clube do Porto. = azul-e-branco = dragão « <i>A equipa portista teve, pelo meio, alguns jogos</i> » ( <i>A Bola</i> , 16.4.1997) « <i>a justiça da vitória portista</i> » ( <i>A Bola</i> , 5.5.1997) 2. Que é jogador, adepto ou simpatizante desse clube.

<sup>15</sup> Este ejemplo también nos lleva a considerar que los lexicógrafos responsables podrían haber optado por usar tan solo la marca genérica “*Desp.*”, cuestión ya abordada en la sección 6.

Como se puede observar, las entradas se caracterizan por la ausencia o presencia de la marca temática. Sin embargo, como vimos, las marcas temáticas son útiles para el usuario y para el lexicógrafo y, por tanto, sería importante uniformizar este tratamiento. Este tipo de armonización será cada vez más importante a medida que avancemos hacia la vinculación de conjuntos de datos léxicos estructurados que cumplan con los estándares con el objetivo de crear recursos lexicográficos accesibles e interoperables.

De acuerdo con nuestro análisis, se puede verificar que no solo las marcas temáticas apuntan a un léxico especializado, sino que ambos diccionarios usan “fórmulas lingüísticas”, usadas en la *definición*, que desempeñan las mismas funciones que las marcas temáticas. Por ejemplo, fórmulas como “no jogo do futebol”, o la introducción de expresiones restrictivas en el texto de la definición a través de colocaciones como “en el fútbol” o del tipo “aplicado a... se aplica a...”. También hay casos en los que se usan más de un recurso simultáneamente.

Para el usuario, la presencia de fórmulas lingüísticas en la definición es una estrategia interesante; pero, para el lexicógrafo, puede suponer una dificultad para el tratamiento de los datos, ya que podría verse afectada la coherencia del recurso lexicográfico. En principio, si en un determinado diccionario el criterio es marcar los dominios con una marca, entendemos que no se debe elegir otro criterio como, por ejemplo, la inclusión de ese dominio en la definición sin marcarlo temáticamente en una etiqueta apropiada. Por otro lado, el uso de herramientas computacionales exige alguna coherencia para que el lexicógrafo pueda controlar adecuadamente este tipo de información —la posibilidad de filtrar el diccionario por un dominio de conocimiento y exportar todas las entradas relacionadas—. Así, una posibilidad sería mantener esas fórmulas lingüísticas pero marcar esas acepciones con el dominio, incluso con una marca que puede ser invisible al usuario.

## 6. CONSIDERACIONES FINALES Y TRABAJO FUTURO

La importancia de la marcación temática en Lexicografía es indiscutible; sin embargo, en los diccionarios analizados en este trabajo se observan inconsistencias que, de forma general, pueden ser atribuidas a la ausencia de una metodología explícita. Sería

conveniente para superar esta situación que los criterios seguidos por los lexicógrafos para tomar decisiones sobre el léxico especializado fueran incluidos en futuras ediciones de las obras.

En las dos introducciones analizadas, la referencia a la inclusión y tratamiento de léxico especializado es prácticamente nula o demasiado generalista. Las decisiones asumidas por los lexicógrafos responsables no están justificadas y solo parecen sostenerse en la presentación de una lista de abreviaturas (marcas temáticas). A su vez, ambos diccionarios no razonan el empleo o el valor de las marcas temáticas mientras difieren en la representación tipográfica. Al mismo tiempo, el número de marcas seleccionadas por los lexicógrafos de estos diccionarios es dispar. Existe también un desequilibrio en cuanto al alcance de las marcas, presentando el *DLPC* muchos ejemplos de subdominios que el *DLE* ignora. La propuesta de armonización de carácter internacional parece, entonces, lejana.

Se impone, de este modo, la necesidad de revisar los sistemas de marcación, eliminando marcas innecesarias o repetitivas como también aquellas distinciones que por ser demasiado finas a veces pueden parecer arbitrarias tanto desde el punto de vista de un lexicógrafo como de un usuario habitual del diccionario. También se observan inconsistencias en las marcas del léxico especializado, ya que algunas veces se utiliza el recurso de la abreviatura, y otras veces se opta por otros mecanismos como el uso de fórmulas presentes en la *definición*, llegando a emplearse incluso más de un mecanismo simultáneamente.

Con vistas a una implementación de buenas prácticas, sería deseable que los lexicógrafos aunaran esfuerzos para colaborar en la propuesta de armonización de las marcas temáticas y así mejorar el proceso de marcación temática en los diccionarios académicos ibéricos. Esta armonización es tanto más útil cuanto más se avanza hasta bases de datos lexicales estructuradas, basadas en normas que permiten la construcción de recursos lexicográficos accesibles adaptados a la necesaria interoperabilidad.

## AGRADECIMIENTOS

Investigación sufragada por financiación nacional portuguesa a través de la FCT – Fundación para la Ciencia y la Tecnología – como parte del proyecto Centro de

Lingüística (UID / LIN / 03213/2019) de la Universidad NOVA de Lisboa (UID / LIN / 03213/2019) y por el Programa de Investigación e Innovación Horizon 2020 de la Unión Europea según acuerdo de subvención nro. 731015 (ELEXIS).  
Las autoras desean agradecer al ILEX (Instituto de Lexicografía de la RAE).

## BIBLIOGRAFÍA

---

### Diccionarios

DLPC = *Dicionário da Língua Portuguesa Contemporânea*, 2001, João Malaca Casteleiro (Coord.), 2 vols. Lisboa: Academia das Ciências de Lisboa & Editorial Verbo.

DLE = *Diccionario de la lengua española* (22.<sup>a</sup> ed.). Real Academia Española [En línea], 2001-2018, [www.rae.es/rae.html](http://www.rae.es/rae.html) [Fecha de consulta: abril de 2019].

### Otra literatura

Ahumada, I. (Ed.) (2002). *Diccionarios y lenguas de especialidad*. Jaén: Universidad de Jaén.

Atkins, B. T. S. y M. Rundell (2008). *Planning the Dictionary. The Oxford Guide to Practical Lexicography*. New York: Oxford University Press.

Battaner, M. P. (1996). Terminología y diccionarios. *Actes de la Jornada Panllatina de Terminologia*. Barcelona: Institut Universitari de Lingüística Aplicada. pp. 93-117.

Bergenholtz, H. y S. Tarp (1995). *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. Amsterdam: John Benjamins Publishing Company.

Boulanger, J. C. y M. C. L'Homme (1991). Les technolèctes dans la pratique dictionnaire générale: quelques fragments d'une culture, *Meta*, 36(1), pp. 23-40.

Costa, R. (2013). Terminology and specialised lexicography: two complementary domains, *Lexicographica*, 29, pp. 29-42.

Estopà, R. B. (1998). El léxico especializado en los diccionarios de lengua general: las marcas temáticas. *Revista de la Sociedad Española de Lingüística*, 28(2), pp. 359-387.

Fajardo, A. (1996-1997). Las marcas lexicográficas: concepto y aplicación práctica en la lexicografía española, *Revista de Lexicografía*, III, pp. 31-57.

Fedorova, I. V. (2004). Style and Usage Labels in Learner's Dictionaries: Ways of Optimization. En G. Williams and S. Vessier (Eds.), *Proceedings of the 11th Euralex International Congress*. Lorient, France: Université de Bretagne-Sud, Faculté des lettres et des sciences humaines. pp. 265-272

Geraci, A. (1991). *IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries*. Piscataway: IEEE Press.

- Hausmann, F. J. (1989). Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht. En F. J. Hausmann, O. Reichmann, H. E. Wiegand y L. Zgusta (Eds.), *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Berlin/New York: Walter de Gruyter. pp. 649-657.
- Jackson, H. (2002). *Lexicography: An Introduction*. London and New York: Routledge.
- Janssen, M., Jansen, F. y H. Verkuyl (2003). The Codification of Usage by Labels. En Piet G. J. van Sterkenburg (Ed.), *A Practical Guide to Lexicography*. Amsterdam, Philadelphia: John Benjamins. pp. 297-311.
- Landau, S. I. 2001. *Dictionaries. The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Monson, S. C. (1973). Restrictive Labels – Descriptive or Prescriptive? En R. I. McDavid, Jr and A. R. Duckert (Eds.), *Lexicography in English*. (*Annals of the New York Academy of Sciences* 211). New York: New York Academy of Sciences. pp. 208-212.
- Porto Dapena, J. A. (2002). *Manual de técnica lexicográfica*. Madrid: Arco Libros.
- Ptaszyński, M. O. (2010). Theoretical Considerations for the Improvement of Usage Labelling in Dictionaries: A Combined Formal-functional Approach, *International Journal of Lexicography*, 23(4), pp. 411-442.
- Rey, A. (1983). Norme et dictionnaire (domaine du français). En E. Bédard y J. Maurais (Eds.), *La norme linguistique*. Québec, Paris: Le Robert, coll. “L’ordre des mots”.
- Rey, A. (2008). *De l’artisanat des dictionnaires à une science du mot: images et modèles*. Paris: Armand Colin.
- Sakwa, L. N. (2011). Problems of Usage Labelling in English Lexicography, *Lexicos*, 21, pp. 305-315.
- Svensén, B. (2009). *A Handbook of Lexicography: The Theory and Practice of Dictionary Making*. Cambridge: Cambridge University Press.
- Wiegand, H. E. (1984). On the Structure and Contents of a General Theory of Lexicography. En R.R.K. Hartmann (Ed.), *LEXeter’83*. Tübingen: Max Niemeyer. pp. 13-30.



## NIVELACIÓN DEL ARTÍCULO LEXICOGRÁFICO EN UN DICCIONARIO MONOLINGÜE DE APRENDIZAJE: EL CASO DE LAS LOCUCIONES ADVERBIALES

### LEVELING OF THE LEXICOGRAPHIC ENTRY IN A MONOLINGUAL LEARNERS' DICTIONARY: THE CASE OF ADVERBIAL IDIOMS

M.<sup>a</sup> Pilar Valero Fernández  
*Universidad de Castilla-La Mancha*  
[profesor.mpvalero@uclm.es](mailto:profesor.mpvalero@uclm.es)  
Ivana Lončar  
*Universidad de Zadar*  
[iloncar@unizd.hr](mailto:iloncar@unizd.hr)

#### RESUMEN

En este texto se persigue mostrar una nivelación plausible de los elementos que configuran la microestructura del artículo lexicográfico en el *DicLELE*, un diccionario monolingüe especializado en locuciones y en el aprendizaje del español como lengua extranjera (ELE). Para ello, a partir de los tres niveles de dominio lingüístico establecidos en el *Marco Común Europeo de Referencia para las lenguas (MCER)*, básico (A1-A2), intermedio (B1-B2) y avanzado (C1-C2), se muestran las características lexicográficas de la microestructura que, en particular, se han seguido para las locuciones adverbiales.

**Palabras clave:** *DicLELE*, microestructura, locuciones adverbiales, niveles A1-A2, B1-B2 y C1-C2

#### ABSTRACT

This text aims to show a plausible leveling of the elements that compose the microstructure of the lexicographical entry in the *DicLELE*, a monolingual dictionary specialized in the idioms and in the learning of Spanish as a foreign language (ELE). To this end, the lexicographic characteristics of the microstructure, obeyed in particular in reference to the adverbial idioms, are shown from the three levels of linguistic domain established by the *Common European Framework of Reference for Languages (CEFR)*, basic (A1-A2), intermediate (B1-B2) and advanced (C1-C2).

**Keywords:** *DicLELE*, microstructure, adverbial idioms, levels A1-A2, B1-B2 and C1-C2

## 1. PRESENTACIÓN GENERAL DEL PROYECTO LEXICOGRÁFICO

El siguiente estudio se centra en la presentación de la microestructura del *Diccionario de locuciones para ELE* (en adelante *DicLELE*)<sup>1</sup>. Esta obra lexicográfica es fruto del proyecto doctoral (2015-921) que se viene desarrollando desde 2016 hasta el momento actual en la Universidad de Zadar (Croacia). En especial, el *DicLELE* sobresale por tratarse de un diccionario de aprendizaje monolingüe digital, dirigido al estudiante de ELE de los niveles A1-A2, B1-B2 y C1-C2, correspondientes respectivamente a las tres búsquedas posibles (búsqueda básica, intermedia y avanzada) para cada lema, y centrado en las locuciones. Como resultado, nos encontramos ante un diccionario especializado (Haensch y Omeñaca, 2004, p. 65) y, como consecuencia, fraccionario, que se compone exclusivamente de un tipo de unidades fraseológicas (en adelante UF), las locuciones, que dan lugar a un vocabulario parcial de la lengua española en su variedad diatópica peninsular.

Para su redacción diccionarística se han seguido fehacientemente los tres fundamentos lexicográficos enunciados por Penadés Martínez (2015, p. 292) en la elaboración de sus cuatro diccionarios, también fraseológicos de corte locucional<sup>2</sup>. Los mismos se sintetizan, primeramente, en la conciencia clara a lo largo de todo el proceso de creación de su propia finalidad y de su destinatario principal (o el *usuario potencial* en términos de la Teoría de la Lexicografía Funcional); por lo que nuestras variables a tener en cuenta fueron la didáctica de ELE y el estudiante de ELE en sus diferentes niveles lingüísticos. Seguidamente, fue imprescindible recabar un leuario representativo acorde con lo primero y, por último, se hizo fundamental analizar

---

<sup>1</sup> Agradecemos la colaboración en este proyecto al ingeniero informático, Palomares Fernández. Su labor ha sido fundamental para la puesta en marcha del *DicLELE* en soporte digital.

<sup>2</sup> Penadés Martínez es autora de cuatro obras lexicográficas fraseológicas locucionales orientadas a ELE, las cuales han sido de vital importancia en el diseño del *DicLELE*: *DICLOCVER* (2003), *DICLOCADV* (2005), *DICLOCNADP* (2008) y *DiLEA* (2019). Asimismo, cabe destacar que, si bien las obras lexicográficas proponen el nivel de lengua de dicha unidad, esto no implica que sus obras lexicográficas se adapten a un nivel de lengua concreto. Reside aquí una divergencia metodológica importante entre el trabajo desarrollado por Penadés Martínez y el *DicLELE*.

detalladamente todos los factores influyentes en el tipo de entradas seleccionadas, las cuales en este artículo se equiparan a las locuciones adverbiales<sup>3</sup>.

Por otro lado, valiéndonos de la definición de *diccionario de aprendizaje*, como aquel que persigue el fin de dar respuesta a todas las necesidades lexicográficas del hablante de una lengua extranjera en referencia a un conjunto de situaciones ‘extra-lexicográficas’ (Tarp, 2006, p. 300), se pretende que el *DicLELE* pueda usarse tanto para el desarrollo de las funciones cognitivas, las cuales permiten una proliferación del conocimiento de la lengua española entre los estudiantes extranjeros, como de las funciones comunicativas, centradas en la resolución de las posibles interferencias propias de la comunicación (Tarp, 2006, p. 303).

Por ello, se parte de la idea de que el diccionario, al igual que otros materiales de aprendizaje como los manuales o las gramáticas, no puede resultar *immune* a la evolución del aprendizaje de una lengua, sino que, por el contrario, el artículo lexicográfico (en adelante AL) responde a un proceso escalar ascendente, en cuanto a la información presentada, desde una descripción sencilla del lema orientada a la comprensión de la locución (nivel A1-A2), continuando con la producción de la locución (nivel B1-B2) hasta alcanzar la comprensión, la producción y el manejo de la unidad fraseológica en múltiples contextos y/o ámbitos (nivel C1-C2).

La defensa de este planteamiento supuso el punto de partida de la elaboración del *DicLELE*, ya que en el presente se desconoce la existencia de diccionarios de aprendizaje locucionales orientados a la enseñanza de ELE<sup>4</sup> en los que el AL albergue dicha modulación según la capacidad lingüística de sus usuarios, siendo este junto al dominio de la lengua materna los dos criterios más importantes en la confección de un diccionario de aprendizaje (Tarp, 2006, p. 310).

---

<sup>3</sup> Se entiende por locución adverbial como aquella unidad fraseológica que se singulariza por equivaler a un adverbio y por desempeñar, por tanto, sus funciones prototípicas (Penadés Martínez, 2012, pp. 167-168).

<sup>4</sup> Esta realidad mejora sustancialmente si la mirada se dirige a los diccionarios didácticos confeccionados según la etapa vital del usufructuario nativo de español. A modo ilustrativo, se puede destacar el trabajo que viene desarrollando la editorial SM, bajo la dirección de Maldonado González, y los diccionarios: *Diccionario básico* (2012), *Diccionario avanzado* (2012) y *Diccionario secundaria y bachillerato* (2012). Desde la Universidad de Jaén, por otra parte, cabe señalar *EnREDate*, *Diccionario temático infantil* (disponible en línea <https://www.enredate.es/>).

En su momento actual, el leuario o la nomenclatura del *DicLELE* está compuesto por 1252<sup>5</sup> entradas, entre las que conviven locuciones de diferente índole: locuciones nominales (52), pronominales (11), adjetivas (77), adverbiales (493), verbales (435), conjuntivas (28) y preposicionales (64). A estas se suman, un conjunto de locuciones con doble naturaleza: locuciones adjetivas/adverbiales (90), locuciones adverbiales/nominales (1) y locuciones adverbiales/preposicionales (1). Sin embargo, hasta la fecha, los datos introducidos en su página web coinciden con los de las locuciones únicamente adverbiales y los de las locuciones con naturaleza adjetiva y adverbial. Esto es, en total, ya hay 583 fichas lexicográficas desarrolladas. Así pues, una vez finalizado al completo el *DicLELE*, este alcanzará su acceso de manera digital u *online* por medio de su propia página web.

Finalmente, en las líneas siguientes nos ocupamos de presentar el AL para cada uno de los tres niveles lingüísticos a partir de una serie de ejemplos de locuciones adverbiales, que atestiguan el trabajo ejecutado. En efecto, la arquitectura del propio AL, máxime si se busca que el diccionario dé respuesta a las necesidades de los discentes de ELE en sus diferentes etapas de aprendizaje, ha de ser el reflejo directo de nuestro objetivo.

## 2. LA MICROESTRUCTURA EN EL NIVEL A1-A2

La microestructura de un diccionario se entiende, *grosso modo*, como el armazón sobre el que se estructura toda la información aportada sobre un lema. De hecho, la microestructura se define como “la ordenación de todos los elementos que componen un artículo lexicográfico” (Haensch y Omeñaca, 2004, p. 47). A su vez, el AL, unidad básica de organización de un diccionario, se compone de dos secciones (Porto Dapena, 2002, p. 183): la enunciativa o entrada (aquí equivalente a una locución adverbial) y la

---

<sup>5</sup> Su macroestructura actual se halla limitada por motivos espacio-temporales sujetos al desarrollo de la tesis doctoral, aunque se aspira a completar su leuario en investigaciones futuras tras la revisión de otras fuentes. Ahora bien, el leuario es el resultado directo del análisis exhaustivo de treinta manuales generales de ELE de todos los niveles del MCER (10 manuales de nivel A1-A2, 13 de nivel B1-B2 y 7 de nivel C1-C2) además de los inventarios de las *Nociones generales*, *Nociones específicas* y *Tácticas y estrategias pragmáticas* del Plan Curricular del Instituto Cervantes (PCIC, 2006).

informativa o los datos ofrecidos, los cuales son variables cualitativa y cuantitativamente según se seleccione la pestaña del nivel A1-A2, B1-B2 o C1-C2.

A continuación, se espera dar respuesta a la forma optada para el diseño tanto de la parte enunciativa como informativa de la microestructura del *DicLELE*. Para este propósito, puesto que hay ciertos elementos que son constantes en los tres niveles, se enuncia solo lo novedoso para cada uno de ellos.

Ya en el interior del nivel A1-A2, la información lexicográfica que se puede hallar ocasionalmente es: lema, definición, dos ejemplos lexicográficos de competencia, dos imágenes, palabras y/o locuciones sinonímicas y antonímicas y, en última instancia, una pestaña despegable con apuntes gramaticales. Ahora bien, antes de iniciar el recorrido por la ficha lexicográfica (en adelante FL), obsérvense las piezas microestructurales del nivel A1-A2 (véase imagen 1).

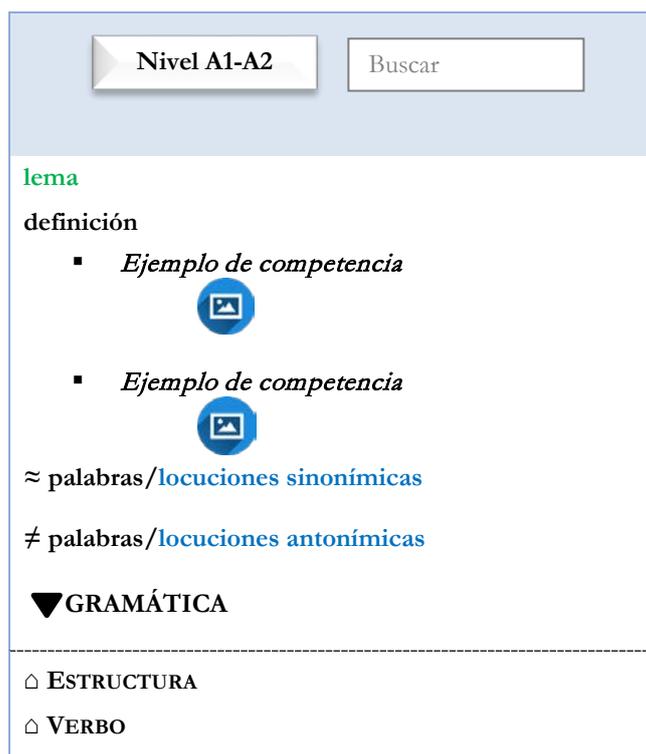


Imagen 1. Ficha lexicográfica. Nivel A1-A2

▪ **Lema o entrada**

Cada lema del *DicLELE*, homólogo a una determinada locución, supone una ficha independiente en el mismo. De esta forma, en función del nivel lingüístico

seleccionado en la pestaña de nivel, la locución aparece desarrollada lexicográficamente en una de las tres formas posibles.

A nivel macroestructural, las entradas se ordenan por orden alfabético, de forma directa, continua y con exclusión de los dígrafos enunciados por la *Ortografía de la lengua española* de la RAE y ASALE (2010). Aun así, al dar por sentado que el usuario no ha de conocer todos los constituyentes de la locución, a este le basta con escribir un constituyente en el buscador para poder acceder a una lista de locuciones que alberguen dicha lexía simple. Por consiguiente, en el *DicLELE* se etiquetan cada uno de los elementos obligatorios y facultativos de todas las locuciones del leuario a excepción de los elementos de contorno. Se trata, pues, de hacer más accesibles e inteligibles las búsquedas de las unidades pluriverbales que ocupan nuestro interés.

En vínculo propio con la redacción<sup>6</sup> de la parte enunciativa del AL, cabe indicar que en la lematización de las entradas, identificables en color verde y con un tamaño superior al resto de elementos, se acogen cuatro realidades:

1. Entrada con una forma única. Por ejemplo: *a continuación, a la vez* o *a lo mejor*.
2. Entrada con diferentes variantes léxicas, sinónimas o antónimas, introducidas entre llaves y, a su vez, diferenciadas entre sí por medio de una barra vertical. Por ejemplo, *a voz en {cuello/grito}*, *con la boca {chica/pequeña}*, *para dar y {regalar/tomar}* o *{más/menos} de la cuenta*. En este caso, se debe advertir que las distintas variantes aparecen, al mismo tiempo, ordenadas alfabéticamente en la redacción de la entrada. Sin embargo, dicho orden se omite si se encuentran entre dichas variantes artículos definidos e indefinidos, determinantes demostrativos, números ordinales y cardinales, así como términos antónimos o graduales. Así se da en: *a {mi/tu/su} parecer* o *a {esta/esa/aquella} altura*. Por su parte, se tienen en cuenta, además, las variantes morfológicas<sup>7</sup> flexivas y derivativas, pónganse como ejemplos **ni**

---

<sup>6</sup> A lo largo de todo el artículo se mantiene la ortotipografía seguida en el *DicLELE* para reflejar explícitamente el trabajo hecho en el mismo con la salvedad del tamaño de letra del lema, en el *DicLELE* a tamaño 16 y en este artículo a 12.

<sup>7</sup> La inclusión de la información morfológica referida a las locuciones es un apunte cada vez más frecuente en los diccionarios de nueva planta o en aquellos que se reeditan. Así lo señala Rodríguez Barcia (2016, p. 225) para el *DLE* y su edición de 2014. Ahora bien, otra cosa distinta es la sistematicidad,

{*borracho/borracha*}<sup>8</sup> o *ni* {*muerto/muerta*}. También, hay que advertir que se tomó la decisión de excluir todas las variantes que se consideran raras o poco usadas en los diccionarios tomados como referencia<sup>9</sup> en respuesta a los atributos del *DicLELE*.

3. Entrada con elementos facultativos. En esta ocasión, todos los constituyentes de la entrada aparecen en fuente negrita y, además, el elemento opcional se discierne entre paréntesis. De esta suerte, el internauta es conocedor de dicha singularidad de la locución en cuestión. Unas muestras las hallamos en: *sin (la menor) duda, en los (puros) huesos o por (lo) tanto*.
4. Entrada con elemento de contorno como ocurre en las locuciones adverbiales *en contra (de algo o de alguien)* o *a favor (de algo o de alguien)*. Por su parte, el elemento de contorno difiere de la propia locución por diferenciarse entre paréntesis y con letra redonda. Se trata, por lo tanto, del uso de una distinción ortotipográfica con el elemento opcional incorporado en fuente negrita.

#### ▪ **Definición lexicográfica y acepciones**

En el nivel A1-A2, la definición lexicográfica se caracteriza por ser sucinta. Se evitó, por ende, el uso de definiciones enciclopédicas, los complementos que acompañan a los verbos definidores y la incorporación o remisión de otra locución en la propia definición (Bosque, 1982, p. 108). Así, se apostó por la propia nivelación lingüística de la definición, tal y como lo expuso Rey Castillo (2017, p. 5) en su artículo dedicado a los criterios en la creación de la macroestructura y la microestructura en los diccionarios monolingües de aprendizaje.

En lo relativo a la propia tipología de definiciones (Porto Dapena, 2002) y a las locuciones adverbiales, se divisan las perifrásticas sustanciales (véanse el lema **en casa**

---

para nosotras muy discutible, con la que se tratan las variantes morfológicas de una locución, si se ahonda en el rastreo del propio *DLE*, por ejemplo.

<sup>8</sup> Repárese que, si las locuciones admiten variación flexiva de género y de número, estas se redactan en su formulación morfológica completa. El fin de esta decisión es aportar información útil referida a estos fenómenos flexivos.

<sup>9</sup> *DLE* (2014), *DFDEA* (2018), *Clave* (2012), *DICLOCVER* (2003), *DICLOCADV* (2005) y *DICLOCNAP* (2008).

y su definición ‘*en la casa propia*’), las definiciones funcionales (verifíquense el lema {*al/por el/todo lo*} *contrario* y su definición ‘*introduce una información opuesta*’), definiciones conceptuales antónimas (obsérvense el lema *de repente* y la definición ‘*de forma no esperada*’), definiciones conceptuales sinonímicas<sup>10</sup> (mírense *a menudo* y la definición ‘*frecuentemente, normalmente*’) o definiciones híbridas (por ejemplo, el lema *a tiempo* y la definición ‘*en el momento correcto, cuando no es tarde*’).

En la definición, además, se tuvo en cuenta el fenómeno del contorno referido a información semántica extensional e intensional (Mellado Blanco, 2018, p. 73). De manera análoga al lema, dicho elemento se incluyó entre paréntesis y en fuente redonda, diferenciándolo debidamente del propio contenido definicional, el cual sobresale por albergar:

1. El elemento referido al que se aplica una determinada locución. Por ejemplo, la entrada *de rechupete* y su contorno definicional de ‘*(comida) con muy buen sabor*’; o el lema *de un trabajo* y su contorno definicional de ‘*(modo de beber) de una vez*’.
2. El complemento indirecto: *a voz en {cuello/grito}* y su contorno en la definición de ‘*hablando muy fuerte (a alguien)*’.
3. El complemento regido: *a la defensiva* y su contorno en la definición de ‘*con carácter de defensa (de alguien)*.’ o *a favor* y su contorno tras la definición ‘*de acuerdo (con algo o con alguien)*’.

#### ▪ Ejemplos lexicográficos

El ejemplo es el único elemento de la microestructura que aparece de manera perpetua y reiterada en cada una de las FL. En un diccionario de aprendizaje como es el *DicLELE*, se considera, pues, la piedra angular para el uso correcto de las entradas. Además, cabe puntualizar que para cada lema se incorporan, como mínimo, dos ejemplos de uso creados *ad hoc*. Como consecuencia los ejemplos manifiestan la naturaleza de ejemplos de competencia (Jacinto García, 2015). En otras palabras, todos los ejemplos que se hallan en las fichas del nivel A1-A2 provienen de la pluma de las

---

<sup>10</sup> Entendidas en el sentido de aquellas definiciones en las que se incluyen varios términos equivalentes y no en el de elegir otra locución para definir la entrada (otra locución) (Penadés Martínez, 2015, p. 199).

lexicógrafas, ya que se desea que la ejemplificación se ajuste, en extremo, desde el punto léxico y gramatical a la capacidad lingüística en la que se halla el usuario. Para conseguir este objetivo se han revisado las sugerencias léxicas y gramaticales formuladas por el PCIC en los inventarios de *Gramática* (capítulo 2), *Nociones generales* (capítulo 8) y *Nociones específicas* (capítulo 9).

Asimismo, en la selección y confección de los ejemplos, se busca que estos reflejen las distintas realidades:

1. Variación morfológica. Por ejemplo, si tomamos el ejemplo de la locución de {*pie/pies*}, se incluyen ejemplos tanto del singular como del plural (■ *Quiero sentarme. Me duelen los pies y no quiero estar de pie;* ■ – *¿Te quieres tumbar en la cama o prefieres estar de pies? – Prefiero estar de pie, ¡gracias!*).
2. Variantes sinónimas como en *ni* {*lo pienses/pensarlo*} y sus ejemplos (■ – *¿Vienes a la fiesta de Juan?; – No, ni pensarlo. Juan es muy tonto;* ■ – *Quiero vivir en Madrid. – Ni lo pienses, es muy caro. No tienes dinero. Es mejor vivir en una ciudad pequeña*).
3. Verbos frecuentes de combinación. Se aprecia en la entrada *por las* {*buenas/malas*}, la nota gramatical de ‘normalmente aparece con el verbo *hacer*’ y su respectivo ejemplo (■ *María hace las cosas por las malas. Nunca tiene buen carácter*).
4. Elementos facultativos. Se tiene en el lema *hoy (en) día* y sus ejemplos (■ *Hoy día es imposible vivir sin móvil. Antes era diferente y vivíamos más tranquilos;* ■ *Me encanta hablar muchas lenguas (inglés, francés, español, etc.). Hoy en día esto es muy importante para viajar o conocer gente nueva de otros países*).
5. Elementos de contorno. Así ocurre en el lema *a favor*, su contorno ‘de algo’ y su respectivo ejemplo: (■ – *Uso el transporte público porque es más barato que el transporte privado. – Yo también estoy a favor del transporte público*).

Desde el punto de vista ortotipográfico, el ejemplo se diferencia del resto de elementos mediante el símbolo de un cuadrado (■), el uso de la fuente redonda y cursiva y el color negro. Además, para remarcar nuestro foco de interés, la locución aparece

subrayada, eso sí, dejando fuera del subrayado tanto sus posibles elementos de contorno como sus plausibles verbos frecuentes de combinación.

#### ▪ **Imágenes**

Otro de los aspectos relevantes del *DicLELE* es la incorporación de la imagen<sup>11</sup> dentro de las fichas lexicográficas, dada la certificación de su beneficencia en los diccionarios orientados a la enseñanza de ELE (Vertikova, 2012; Leal Riol, 2013). En efecto, después del ejemplo, en caso de que se considere factible la anexión de una imagen propia, se decide llevarla a cabo. En cambio, cabe aclarar que la imagen incluida no se basa en la imagen subyacente de los constituyentes de la locución, sino que la imagen viene a representar el significado idiomático de la locución. Es decir, si cogemos como modelo la locución adverbial *de copas* no se plasma el recipiente, sino un grupo de personas en una fiesta.

Ya en el ámbito específico de las locuciones adverbiales asociadas a las distintas *Nociones Generales, Nociones Específicas y Tácticas y Estrategias pragmáticas* recogidas en el *PCIC* se adjuntan imágenes para:

1. Determinadas nociones espaciales de: localización/posición relativa (*a lo lejos*), posición absoluta (*de rodillas, {a/en} hombros, de espaldas*, etc.), medida (*del tiempo*), temporales (*de noche, de sol a sol*, etc.) o cantidad (*a mares, a puñados, como sardinas en lata*, etc.).
2. Determinadas nociones cualitativas de: color (*en blanco y negro, en color*, etc.), limpieza (*manga por hombro, como una patena*, etc.), humedad/sequedad (*a chorros, a borbotones*, etc.) o de tiempo (*en punto, en hora, a tiempo*, etc.).
3. Determinadas nociones específicas referidas al individuo: cuerpo (*a pierna suelta*), características físicas (*en los (puros) huesos*), alimentación (*a rayos*), restaurante (*de copas*) o transporte (*a pie*).

Asimismo, en el caso de incluir imágenes, se dan dos escenarios posibles:

1. Se incluye una única imagen válida para los dos ejemplos.

---

<sup>11</sup> Todas las imágenes anexadas en el *DicLELE* han sido cedidas por Fco. Javier Ruiz de Martín Esteban Fernández y M.ª Amparo Fernández López.

2. Se incluyen dos imágenes, cada una relativa a un ejemplo. Lo vemos en la entrada *en público*. con el ejemplo de ‘ ■ *no me gusta hablar en público. Me pongo muy nerviosa*’ y su respectiva imagen de una persona emitiendo un discurso ante un público, por un lado; y por otro, con el ejemplo de ‘ ■ *cuando hablo en público, todo el mundo me mira. ¡No me gustan las presentaciones en clase!*’ y su imagen de una estudiante hablando en la pizarra ante sus compañeros.

Por el contrario, por su carácter abstracto y poco visual, no se adjunta imagen en las locuciones adverbiales referidas a:

1. Determinadas nociones evaluativas: *a {mi/tu/su} modo de ver, hasta el gorro*, etc.
2. Determinadas nociones temporales, por ejemplo, de desarrollo (*pasado mañana*) o de frecuencia (*de vez en cuando, a veces, de Pascuas a Ramos*, etc.).
3. Determinadas nociones cuantitativas: cantidad numérica (*por lo alto*), proporción (*a escote*), grado (*de espanto*) o medida (*a ojo (de buen cubero)*).
4. Determinadas nociones existenciales de: certeza/incertidumbre (*tal vez*), distancia (*a distancia*), ausencia (*de cuerpo presente*), disponibilidad/indisponibilidad (*a mano*) o realidad/ficción (*de color de rosa*).
5. Nociones de dimensión perceptiva y anímica: carácter (*por las {buenas/malas}, como una fiera*, etc.), sentimientos (*a gusto, entre dientes*, etc.) o estados mentales (*en Babia, en {mis/tus/sus} cabales*, etc.).
6. Nociones específicas referidas al individuo, por ejemplo, la alimentación (*de un trago*) o el trabajo (*a tiempo {completo/parcial}*).
7. Locuciones adverbiales con valor de marcador discursivo (*por una parte, no obstante*, etc.).

■ **Palabras y/o locuciones sinónimas y antónimas**

Tras la definición, los ejemplos y las imágenes, se incluyen las palabras y/o las locuciones sinónimas y antónimas que se creen convenientes para cada lema. Para introducirlas, además, se recurre al uso de dos símbolos. Por una parte, para los

sinónimos, se recurre al símbolo de aproximación ( $\approx$ ) y, por otra, para los antónimos, al de contrario ( $\neq$ ).

A colación de la sinonimia, esta ha de ser interpretada en la línea de ‘unidades de significado relativamente próximo, parecido o semejante en determinadas situaciones’ y, por ende, se es consciente de que la sinonimia no se puede manifestar en términos absolutos. Como reflejo de esta realidad, la locución sinonímica o antonímica de una entrada, eso sí, en su FL correspondiente, incorpora la información precisa para discernir entre los diferentes sinónimos y su aplicación correcta dentro de un contexto.

Por lo tanto, se está de acuerdo con la reflexión de Mellado Blanco (2004) y su apuesta por la existencia de una sinonimia más bien ficticia, e incluso forzada, que real entre las locuciones sinónimas internas, esto es, “FR (fraseologismos) sinónimos de estructura diferente” (Mellado Blanco, 2004, p. 91). Así pues, sin entrar en el debate del fenómeno de la sinonimia, se apuesta por la conveniencia de la inclusión de este tipo de relaciones semánticas (también aplicable a la antonimia) en los diccionarios de aprendizaje y, más todavía, en los destinados a aprendientes de lenguas. Esto ya fue debidamente justificado por la lexicógrafa Penadés Martínez (1999, p. 43) al aseverar que estas relaciones paradigmáticas suponen “una serie de mnemotécnicas virtuales que pueden ser de gran utilidad para memorizar y, al final, aprender unidades lingüísticas”.

La sinonimia se refleja en el nivel A1-A2 en diferentes fichas. Un reflejo de ello lo tenemos en la locución adverbial *en cuanto a*. Para dicha entrada, se muestra un único sinónimo no fraseológico *sobre*. Sin embargo, en los niveles siguientes dicha relación semántica se amplía considerablemente, es decir, el repertorio de unidades ofrecidas es mayor que en el nivel A1-A2. Así, en el nivel B1-B2 se presentan los sinónimos *sobre*, *(con) respecto a* y *en relación con*; y en el nivel C1-C2 se encuentran los sinónimos *sobre*, *(con) respecto a*, *en relación con*, *en lo que respecta a*, *a colación de*, *acerca de*, *referente a*.

En cuanto a la antonimia, relación mucho menos fructífera entre las locuciones (Mellado Blanco, 2004), se introducen antónimos que representan una oposición gradual, direccional, complementaria o inversa. A modo ilustrativo, para el lema *en voz*

*alta* se incluye el antónimo *en voz baja* y viceversa; o para la entrada *de noche* se anexa el antónimo *de día*.

Asimismo, este componente del AL se caracteriza por ser un elemento de referencia. Con este propósito, se dan dos situaciones: la locución aparece en color azul cuando forma parte de la macroestructura del *DicLELE* y se caracteriza por su naturaleza de hipervínculo; pero se registra en color negro aquella locución que se omite, al menos por ahora, en esta obra lexicográfica. Igualmente, todas las lexías simples o sintagmas (no fraseológicos) sinónimos o antónimos referidas al lema a definir se presentan en color negro por ser el *DicLELE* un diccionario únicamente de locuciones.

#### ▪ **Apuntes gramaticales**

El último elemento de la FL se refiere a otros aspectos lingüísticos complementarios de interés para el estudiante de nivel A1-A2, ya que siguiendo las advertencias expuestas por Tarp (2018, p. 378), se comparte la idea de lo contraproducente que puede llegar a ser una sobrecarga informativa, en este caso, gramatical. De este modo, si se desea desplegar la pestaña de ‘GRAMÁTICA’ situada al final de la FL, se incluyen varias puntualizaciones precedidas siempre por su respectivo símbolo (△) y la identificación correspondiente, bien bajo la etiqueta de ESTRUCTURA o bien con la etiqueta de VERBO:

1. △ ESTRUCTURA. La variación morfológica del artículo definido seguido de preposición, esto es, la presentación de los artículos contractos. Así ocurre con locuciones adverbiales como *en presencia* o *a favor* en contacto con su elemento de contorno (*de algo* o *de alguien*). Además, tras dicha información se incluye una muestra.
2. △ VERBO. Los verbos frecuentes con los que se combina la locución definida. De esta forma, si se desea, además el estudiante puede acceder a la conjugación verbal de dichos verbos por medio de un hipervínculo<sup>12</sup>. Ahora bien, puesto

---

<sup>12</sup> De esta forma, se aboga por las postulados recogidos en Fuertes-Olivera y Bergenholtz (2018, p. 18) sobre la conveniencia de presentar los datos, en este punto sobre la conjugación del verbo, de un modo accesible para el usuario sin la inclusión de referencias externas (tales como números o letras) para llegar a las conjugaciones verbales. Sin duda alguna, esto lo facilita el formato digital del *DicLELE*.

que el nivel del discente es A1-A2, según las recomendaciones del *PCIC*, los tiempos verbales que se proporcionan son: presente, pretérito indefinido, pretérito imperfecto y pretérito perfecto compuesto de indicativo, las formas de imperativo y, por último, las formas no personales.

En la imagen 2 se plasma lo expresado para la FL de nivel A1-A2 en relación con la locución *en contra*.

Nivel A1-A2    *contra*

**en contra (de algo)**

en oposición (a algo)

- *Uso el coche todos los días.*
- *Yo no. Yo estoy en contra del transporte privado. El autobús es perfecto para mí.*
- *Nosotros estamos en contra de trabajar por las tardes. Solo queremos trabajar por las mañanas.*

≠ a favor, de acuerdo

▼ GRAMÁTICA

---

△ ESTRUCTURA

en contra + del + nombre masculino singular  
*en contra del horario*

en contra + de los + nombre masculino plural  
*en contra de los exámenes*

en contra + de la + nombre femenino singular  
*en contra de la contaminación*

en contra + de las + nombre femenino plural  
*en contra de las faldas en invierno*

en contra + de + infinitivo  
*en contra de levantarse pronto*

△ VERBO

Normalmente aparece con el verbo **estar**

Imagen 2. Lema “en contra”. Nivel A1-A2

### 3. LA MICROESTRUCTURA EN EL NIVEL B1-B2

En lo que respecta al siguiente nivel, el B1-B2 de español, los resultados expuestos difieren en algunos elementos de la microestructura del nivel anterior. Es, por esta

razón, que ahora nos detenemos en sus aspectos propiamente específicos. No obstante, antes de adentrarnos en la descripción de la propia microestructura, cabe advertir que esta se sustenta sobre una plataforma compuesta por: subcategoría de locución\*<sup>13</sup>, marca de registro\*, definición, dos ejemplos controlados\*, palabras y/o locuciones sinónimas y antónimas y, en último lugar, unas notas gramaticales\*. De un modo visible se atisban dichos elementos en la imagen 3.

Nivel B1-B2    Buscar

**lema** [tipo de locución; marca diafásica]

**definición**

- *Ejemplo controlado*
- *Ejemplo controlado*

≈ palabras/**locuciones sinónimas**  
≠ palabras/**locuciones antónimas**

▼ **GRAMÁTICA**

△ ESTRUCTURA  
△ VERBO  
△ REGISTRO DE SINÓNIMOS Y ANTÓNIMOS  
△ INTENSIFICADORES  
△ PUNTUACIÓN  
△ POSICIÓN  
△ ORTOGRAFÍA

Imagen 3. Ficha modelo. Nivel B1-B2

#### ▪ **Categoría de la locución**

Seguidamente al lema, se muestra una información lingüística relevante para nuestros usuarios del *DicLELE*, el tipo de locución<sup>14</sup> que corresponde a dicha unidad

<sup>13</sup> En aquellos elementos del AL en los que se ha integrado un asterisco (\*), este símbolo indica que resulta ser un constituyente inédito o de distinta índole con respecto al nivel anterior.

<sup>14</sup> Pese a que se escape de los límites de la microestructura, se ha de reparar en el hecho de que, dentro del menú de la página web del *DicLELE*, en *Instrucciones*, se explique brevemente cuáles son las particularidades (pluriverbalidad, idiomática, (in) fijación, principalmente) de una locución y de cada subtipo. En concreto, esta información gramatical se registra en las *Instrucciones* de los niveles B1-B2 y C1-C2. Sin embargo, en el nivel inicial las locuciones se denominan ‘expresiones’, etiqueta que podría ser atacada de ambigua y laxa pero que, desde el punto de vista de didáctica de ELE, nos es válida para una primera aproximación. En cierto modo, el *DicLELE* se halla en la órbita de un material complementario, pero que, a pesar de ello, pretende dar cuenta de la idiosincrasia de las unidades que alberga en su interior. Es, por esta razón, que nos apoyamos en profesores de lenguas extranjeras como Solano Rodríguez (2007) y la repercusión positiva que supone la enseñanza explícita de las características

fraseológica. Así, a partir de este apunte, los interesados pueden incorporar, con mayor seguridad, las locuciones de forma activa en el discurso al ser conocedores de la equivalencia entre una subclase de locución y su respectiva correspondencia gramatical, ahora monoléxica (nombre, adjetivo, adverbio, etc.). En otras palabras, en fuente redonda, en naranja y entre corchetes, se presentan las etiquetas de locución nominal<sup>15</sup>, pronominal, adjetiva, verbal, adverbial<sup>16</sup>, preposicional y conjuntiva.

Asimismo, en el caso de que una misma locución pueda desempeñar, por ejemplo, tanto la función de locución adjetiva como adverbial, se apuesta por la inclusión de la etiqueta simple de *locución adverbial* o de *locución adjetiva* correspondiente a la acepción en particular que desempeña dicha función. Véanse el lema *de veras* y sus acepciones: la primera referida a la locución adjetiva ('1. [*locución adjetiva*] *verdadero, verdadera*') y la segunda a la locución adverbial ('2. [*locución adverbial*] *realmente*'). Igualmente, se evidencia en la entrada *de puta madre* y su primera acepción referente a la función de locución adjetiva ('1. [*locución adjetiva*] *muy bueno*') y la siguiente a la de locución adverbial ('2. [*locución adverbial*] *muy bien*').

Por otra parte, huelga decir que, tras el análisis de nuestros seis diccionarios de referencia (véase nota 9), se ha decidido lexicalizar la locución de la forma más frecuente en estos, pues son muchos y variados los problemas que suponen la lematización de las locuciones y su repercusión directa, por ende, en la consideración del subtipo locucional.

---

circundantes a las locuciones, en su opinión primeramente en la lengua materna y luego en la lengua extranjera.

<sup>15</sup> Un caso especial lo ocupan las locuciones nominales, para las que, además del tipo de locución, se incluyen sus propiedades morfológicas referentes al género y al número. De este modo, una locución como *las tantas* va especificada con la marca gramatical de 'locución nominal femenina plural' o una locución como *números rojos* con la de 'locución nominal masculina plural'. De manera similar (no coincidente) al trabajo realizado por Penadés Martínez en su diccionario (2008), se deja rastro de las propiedades morfológicas de aquellas locuciones nominales sexuadas. Así sucede en {*abogado/abogada*} *del diablo* o en {*hijo/hija*} *de puta*. No obstante, si en el *DICLOCNAP* únicamente se dejaba rastro de dicho fenómeno en la marca gramatical, en el *DiLELE* se hace por partida doble: en el lema y en la marca de 'locución nominal masculina y femenina', además de su verificación en los ejemplos.

<sup>16</sup> Se omite, consiguientemente, el subtipo de locuciones marcadoras (*en primer lugar, dicho de otro modo*, etc.) establecido por fraseólogos como Ruiz Gurillo (2001, 2005, 2010) o Montoro del Arco (2006). En particular, las denominadas locuciones marcadoras se engloban bajo las locuciones adverbiales. La razón de dicha decisión radica por seguir los postulados de la *NGLE* (2009, pp. 2355-2361) y su postura acerca de la consideración del marcador, y por lo tanto de las supuestas locuciones marcadoras, como una categoría pragmática y no gramatical del estatus del nombre, del adjetivo o del adverbio, por ejemplo.

En último lugar, tanto para el tipo de locución como la marca de registro se rechaza el uso de abreviaturas, de forma que el *DicLELE* apuesta por el aprovechamiento de la *cuasi* ilimitación del espacio que ofrece el formato digital.

▪ **Marcas diafásicas**

Después del subtipo de locución, se introduce la marca diafásica que recoge su respectivo registro de uso. En consecuencia, se omite el resto de marcas diacrónicas<sup>17</sup>, diastráticas, diatópicas y diatécnicas, las cuales forman en su conjunto el sistema de marcas diasistemáticas (Porto Dapena, 2002). Así pues, ante la falta de unanimidad por el uso de unas marcas de registro unificadas en los diccionarios (Penadés Martínez, 2013; González Aguiar, 2003), las marcas de registro, hasta el momento requeridas<sup>18</sup>, se ciñen a tres categorías: coloquial, vulgar y formal<sup>19</sup>. De manera que aquella entrada que carece de dicha marca implica su respectiva pertenencia a un ámbito normal o general. Se procede a la explicación de cada una de las marcas así como la inclusión de algunos ejemplos que en el *DicLELE* poseen las diferentes marcas:

1. En primer lugar, se asigna la etiqueta *coloquial*<sup>20</sup> a locuciones propias de una situación informal, distendida o familiar. Estamos ante este contexto en locuciones como: *a chorros*, *a muerte* o *a ojo (de buen cubero)*.
2. En segundo lugar, se adjunta la etiqueta *vulgar* a locuciones con carácter grosero u ordinario y que, por lo tanto, su uso debe estar mucho más cuidado y reservado a situaciones puntuales. Locuciones como *de puta madre* o *de (tres pares de) narices* tienen esta etiqueta. Cabe advertir que, dadas

---

<sup>17</sup> En el *DicLELE* las marcas diatópicas quedan fuera de lugar por considerarse solo la variedad peninsular y, a su vez, omitir todo aquello marcado por los regionalismos.

<sup>18</sup> Repárese, nuevamente, en este punto que la macroestructura del *DicLELE* es el reflejo directo de las locuciones trabajadas didácticamente en un conjunto representativo de manuales desde el nivel A1 hasta el nivel C2 y los diferentes inventarios del *PCIC* (2006). No obstante, no se rechaza la idea de aumentar el número de lemas en proyectos de investigación venideros.

<sup>19</sup> Con esta nómina cerrada de marcas se intenta salvaguardar los tintes de imprecisión que rodean a la tradición lexicográfica y su selección dispar de las marcas diafásicas, a veces incluso entremezcladas con las marcas diastráticas (Garriga Escribano, 2003, p. 117), según el diccionario consultado.

<sup>20</sup> Se ha destacar que, en la asignación de las marcas de registro o diafásicas, las locuciones se ven afectadas más, inequívocamente, por la imagen que desprenden estas en la mente del receptor (Penadés Martínez, 2013) que por el significado recto de los significantes que las integran. Como consecuencia, se atestigua que locuciones remarcadas como coloquiales, como *un rábano*, *un pimiento* o *de bigos a brevas* se ven perturbadas por ello.

nuestras fuentes de formación del corpus, son muy escasas las locuciones que proceden a este grupo.

3. Seguidamente, la marca de *formal* se reserva para las locuciones más frecuentes en la lengua escrita y en las obras literarias, es decir, dichas unidades son más propias de un registro protocolario o elevado. Así sucede con las unidades fraseológicas *de pro* o *a lo sumo*.

#### ▪ **Definición**

Frente a los artículos monosémicos o aquellos compuestos por una única acepción (Rodríguez Barcia, 2016, p. 213), dada la circunstancia de que para un mismo lema se presenten dos o más valores oportunos para el nivel en cuestión, es decir, acepciones, estas dan lugar a artículos lexicográficos múltiples<sup>21</sup>, ordenados por orden categorial (es decir, primero aquellas acepciones referidas a nombres seguidas sucesivamente por las de adjetivos, adverbios, verbos, preposiciones, conjunciones e interjecciones). Además, cada una de las acepciones va precedida por el número cardinal correspondiente en negrita (**1**, **2**, **3**, etc.) y junto con su propia información microestructural (tipo de locución, marca diatécnica, ejemplos lexicográficos, palabras y/o locuciones sinonímicas y antonímicas). Así sucede en el lema **de verdad** y sus tres acepciones: ‘*1. verdadero, verdadera*<sup>22</sup>; *2. realmente*; *3. de una manera seria*’.

Según lo expuesto, dada la circunstancia de que nos enfrentemos ante artículos múltiples, las unidades sinonímicas y antónimas se acompañan del número de la acepción de la entrada a la que se refieren. De esta manera, pongamos por caso que se busca el lema {*al/por el/todo lo*} *contrario*. Ya en el interior de su FL, se aprecia la inclusión de un sinónimo, *en contra*. Ahora bien, como únicamente la segunda acepción de *en contra* es sinónima de {*al/por el/todo lo*} *contrario*, se toma la decisión de incluir la unidad sinonímica seguida del número cardinal correspondiente entre paréntesis: *en contra* (2).

---

<sup>21</sup> Los artículos múltiples son más habituales en los niveles B1-B2 y C1-C2. Una exhibición de este apunte está en la locución conjuntiva *ya que* y sus tres acepciones alcanzadas paulatinamente en las tres búsquedas posibles: **1** expresa causa (niveles A1-A2, B1-B2 y C1-C2); **2** expresa consecuencia (niveles B1-B2 y C1-C2); **3** expresa concesión (nivel C1-C2).

<sup>22</sup> Una vez más se tiene presente la morfología flexiva, en este caso, en lo que respecta al género.

### ▪ Ejemplos lexicográficos

Habitualmente, en los manuales de lexicografía general (Medina Guerra (coord.), 2003; Ahumada, 1989; Martínez de Sousa, 1995; Porto Dapena, 2002, entre otros) se presenta el ejemplo como ‘el elemento esencial’ o ‘parte fundamental’ del AL. En efecto, Olímpio de Oliveira (2007, p. 261), experta fraseógrafa, se detiene en este aspecto certificando que “los ejemplos son muy importantes en el tratamiento lexicográfico de las unidades fraseológicas en cualquier tipo de diccionario y, principalmente, en los que van dirigidos al aprendizaje”. Por ello, en este estadio de aprendizaje del español, se mantiene el cuidado extremo a la ejemplificación. En esta ocasión, frente a los ejemplos de competencia, se seleccionan ejemplos controlados. Particularmente, a partir de la clasificación realizada por Jacinto García (2015, pp. 37-38), los ejemplos controlados se corresponden con aquellos provenientes de corpus<sup>23</sup> (incluido Internet como corpus *per se*) sobre los que recae una manipulación. La adaptación de los ejemplos se apoya en diferentes argumentos: el reflejo de las diferentes variantes léxicas y/o morfológicas, de los elementos facultativos, de los elementos de contorno, la flexión de los verbos de una entrada o la inclusión de los verbos fijados como habituales para un lema.

Por otra parte, a los dos ejemplos adjuntados tras la definición se suman otros ejemplos que vienen a ejemplificar diferentes aspectos<sup>24</sup> incluidos bajo la pestaña desplegable de GRAMÁTICA. Así, en el lema *en {primer/segundo/tercer...} lugar* se evidencian dos apuntes:

1. El primero referido a la puntuación de la locución. Esto es: ∆ PUNTUACIÓN: ‘después de *en {primer/segundo/tercer...} lugar* se necesita una coma<sup>25</sup> (,): *En*

---

<sup>23</sup> Tarp y Fuertes-Olivera (2016, p. 277) establecen una clasificación de corpus: por un lado, los corpus tradicionales, como pueden ser el *CREA*, el *CORDE* o el *CORPES XXI*; por otro, los corpus compuestos de textos recabados de Internet; en última instancia, Internet como un corpus en sí mismo.

<sup>24</sup> Si la evidencia la hallamos en la locución verbal *darse cuenta*, se tiene esta información: (1) verbo pronominal: *Me* doy cuenta; *Se* ha dado cuenta; (2) darse cuenta + DE + NOMBRE/PRONOMBRE: *Le* encanta enseñar. *Se* ha dado cuenta *de* *ello*; (3) darse cuenta + DE + QUE + INDICATIVO: *Me* he dado cuenta *de* *que* *he* perdido las llaves. Nótese que se ha seguido el consejo de Tarp (2018, p. 379) y su propuesta de presentación de la información sintáctica.

<sup>25</sup> En el nivel siguiente se incluyen otras posiciones posibles, como es la de inciso.

*primer lugar, lava los platos. En segundo lugar, recoge la mesa. En tercer lugar, limpia la ventana*’.

2. El segundo en relación a la posición de la locución en el discurso. Esto es:  $\Delta$  POSICIÓN: ‘siempre aparece(n) en posición inicial de frase: *En primer lugar, escribiré un mensaje a los invitados*’.

A diferencia de los ejemplos presentados en primera posición, en la sección de apuntes gramaticales, los mismos se caracterizan por ser de competencia.

#### ▪ **Notas gramaticales**

Para las locuciones adverbiales, entre las anotaciones gramaticales, se aglutinan datos lingüísticos de naturaleza dispar junto con las ejemplificaciones oportunas. Los diferentes elementos se reconocen por el uso de las distintas etiquetas que los encabezan. En atención a ello, se enumeran, a partir del orden de aparición en el AL, los siguientes datos:

1.  $\Delta$  ESTRUCTURA. Combinatoria sintagmática de la locución, así como selección del modo verbal que realiza la locución. En este caso, siempre se produce una ampliación de las distintas posibilidades combinatorias con respecto al nivel anterior.
2.  $\Delta$  VERBO. Verbo frecuente de combinación (integrado ya en la ficha de A1-A2) junto con la posibilidad de obtener la conjugación verbal completa mediante el uso del hipervínculo situado en el propio verbo. Asimismo, en este mismo punto se incluye, en caso de situar el cursor sobre el verbo, la naturaleza del verbo incluido en la propia locución o del verbo señalado como usual: verbo regular *vs.* irregular, verbo pronominal. De este modo, en la entrada *de morros* por ejemplo, se hace referencia a *estar* como verbo asiduo en la locución pronominal. En particular, se introduce: ‘aparece frecuentemente con el verbo *estar*> verbo irregular’.
3.  $\Delta$  REGISTRO DE SINÓNIMOS Y ANTÓNIMOS. Referencia a la marca diafásica de las unidades sinonímicas y/o antonímicas incluidas previamente. Verbigracia, en la entrada *ni torta* se inscribe, en su lugar correspondiente de la FL, un sinónimo, *ni papa*). Posteriormente, en el apartado de las notas

gramaticales, se inscribe dicha unidad seguida de su correspondiente valor ‘forma coloquial’. Igual sucede en el lema *de vez en cuando* y su sinónimo *de unas a peras* el cual se recupera en la parte gramatical agregado con la etiqueta de ‘forma coloquial’. De esta forma, el usuario no solo conoce el registro de uso del lema, sino que también puede tener constancia del valor que aportan sus respectivos sinónimos o antónimos.

4. △ INTENSIFICADORES. Posibilidad de inclusión de intensificadores así como su debida posición. En la locución *de acuerdo* se adjunta la nota de ‘puede aparecer con el adverbio *muy* para destacar el acuerdo con una opinión o pensamiento: *Estoy muy de acuerdo con tu plan*’.
5. △ PUNTUACIÓN. Apuntes sobre la puntuación, como ocurre frecuentemente en las locuciones adverbiales con valor de conector discursivo. De esta forma, *por una parte* va acompañado de ‘después de *por una parte* se necesita una coma (,): *Por una parte, el piso está cerca de la Universidad. Por otra parte, el alquiler es barato*’.
6. △ POSICIÓN. Posición de la locución en el discurso. Por su parte, el lema *en resumidas cuentas* registra la nota de ‘aparece al inicio de frase y de forma externa a la oración<sup>26</sup>: *En resumidas cuentas, tenemos mucho trabajo para esta semana y no podremos ir al congreso de Valencia*’.
7. △ ORTOGRAFÍA. Apuntes ortográficos. El lema *a tocateja*, en su caso, recoge la puntualización de ‘se puede escribir también como *a toca teja*’. Otro ejemplo, también, es el de *en Babia* y la nota de ‘se puede escribir también con minúscula *en babia*’.

En la imagen 4 se recoge lo indicado para la FL de nivel B1-B2 a colación de la locución *en contra*.

---

<sup>26</sup> Según se expone en el *Manual* de la NGLLE (2010, p. 578), dicha locución adverbial, por su incidencia sintáctica, se corresponde a la categoría adverbial de los adjuntos, más concretamente, a los periféricos con valor de conector discursivo.

Nivel B1-B2	contra
-------------	--------

**en contra** (de algo o de alguien) [locución adverbial]

**1 en oposición** (a algo o a alguien)

- *Día de protestas en contra de la educación privada y el aumento de los impuestos.*
- *La mayoría de los ciudadanos votó en contra del partido conservador.*

≠ **a favor, de acuerdo**

**2 contrariamente** (a algo)

- *En contra de lo que suele pasar en Madrid, vivo en piso de 200 m<sup>2</sup>.*
- *Tus padres, en contra de lo que ocurre en otras familias, te han dado total libertad siempre.*

≈ {**al/por el/todo lo**} **contrario**

▼ **GRAMÁTICA**

△ **ESTRUCTURA**

en contra + de + artículo definido/indefinido + nombre  
*en contra del gobierno corrupto / en contra de un gobierno corrupto*

en contra + de + infinitivo  
*en contra de trabajar por un salario tan bajo*

en contra + de + lo + adjetivo  
*en contra de lo clásico*

en contra + de + lo que + verbo en indicativo  
*en contra de lo que pasa*

en contra + de + de + pronombre posesivo  
*en contra de mí*

△ **VERBO**

Normalmente aparece con los verbos **estar, ponerse**

Imagen 4. Lema “en contra”. Nivel B1-B2

#### 4. LA MICROESTRUCTURA EN EL NIVEL C1-C2

El último escalón en la descripción de las microestructuras del *DiLELE* lo ocupa el nivel C1-C2. Para este nivel, se cree oportuno que la FL presente los siguientes elementos: subcategoría de locución, marca de registro, marca de ámbito<sup>27\*</sup>, definición, dos ejemplos reales\*, palabras y/o locuciones sinonímicas y antonímicas, y, en última instancia, una notas gramaticales\* y pragmáticas\*. De un modo visible se contemplan dichos elementos en la imagen 5.

<sup>27</sup> En aquellos elementos del AL en los que se ha integrado un asterisco (\*), dicho símbolo indica que resulta ser un constituyente inédito o de distinta índole con respecto al nivel anterior.

The image shows a digital dictionary interface. At the top, there is a search bar with the text 'Nivel C1/C2' and a 'Buscar' button. Below the search bar, the entry for 'lema' is displayed, including its type '[tipo de locución; marca diafásica]' and a definition '[marca diatécnica]definición'. Two example sentences are listed: 'Ejemplo real (corpus CORDE/CORPES XXI)'. Below the examples, there are two lines of text: '≈ palabras/locuciones sinónimas' and '≠ palabras/locuciones antónimas'. The entry is divided into three sections: 'GRAMÁTICA', 'PRAGMÁTICA', and 'FUERZA ILOCUTIVA'. Each section contains a list of sub-entries, each preceded by a triangle symbol (upward for 'GRAMÁTICA' and 'FUERZA ILOCUTIVA', downward for 'PRAGMÁTICA').

**Nivel C1/C2**    Buscar

**lema** [tipo de locución; marca diafásica]  
[marca diatécnica]**definición**

- *Ejemplo real (corpus CORDE/CORPES XXI)*
- *Ejemplo real (corpus CORDE/CORPES XXI)*

≈ palabras/locuciones sinónimas  
≠ palabras/locuciones antónimas

▼ **GRAMÁTICA**

- △ ESTRUCTURA
- △ VERBO
- △ REGISTRO DE SINÓNIMOS Y ANTÓNIMOS
- △ INTENSIFICADORES
- △ PUNTUACIÓN
- △ POSICIÓN
- △ ORTOGRAFÍA
- △ MODALIDAD ENUNCIATIVA
- △ NORMA

▼ **PRAGMÁTICA**

- △ FUERZA ILOCUTIVA
- △ FUNCIÓN DISCURSIVA

Imagen 5. Ficha modelo. Nivel C1-C2

### ▪ **Marcas diatécnicas**

Las marcas diatécnicas o de ámbito de especialidad se definen como aquellas que “informan sobre las áreas de conocimiento y ámbitos profesionales en los que esas voces se utilizan” (Rodríguez Barcia, 2016, p. 226). Concretamente, en el *DicLELE* las marcas de ámbito no se insertan junto al lema sino que aparecen encorchetadas, en fuente redonda y en color naranja entre el número de acepción y la definición propiamente dicha. La razón de ello es que, en nuestro corpus actual, un lema puede presentar diferentes acepciones y solo una de ellas hace mención a dicho campo de especialidad. Como consecuencia, en la ordenación de las acepciones, se apuesta por incluir primeramente la no marcada y después la marcada.

Entre las marcas que hasta el momento han sido necesarias para nuestro leuario, destacan: agricultura, deportes, economía, milicia, fútbol, biología, tauromaquia, televisión, teatro, entre otras.

De este modo, en el lema *en cierne* se incluyen dos acepciones: ‘1. *en una fase inicial, no acabada*; 2. *[agricultura] en flor*’. Es, justamente, en la segunda acepción donde se aprecia la marca diatécnica. Por su lado, en la entrada *entre bambalinas*, se acogen también dos acepciones: ‘1. *De manera reservada*; 2. *[teatro] fuera del escenario de representación de una obra*’. Las segundas acepciones de los lemas *en cierne* y *entre bambalinas* son omitidas tanto en el nivel A1-A2 como en el nivel B1-B2.

#### ▪ Ejemplos lexicográficos

En este punto del *DicLELE*, los ejemplos se distinguen por su condición de ejemplificaciones reales extraídas de dos corpus de referencia del español: *CORPES XXI* y *CORDE*. El alegato a favor de sus respectivos usos se debe a una cuestión de recoger ejemplos veraces de la lengua española (en su variedad peninsular) y, además, con el filtro impuesto por las lexicógrafas de textos publicados en las últimas décadas. No obstante, a pesar de ser ejemplos reales, se modifica todo lo que se califica como: erratas, normas de acentuación anteriores a la *Ortografía de la lengua española* (2010) o la inclusión de grafemas suprimidos por su uso en un registro coloquial.

Asimismo, los dos ejemplos ofrecidos tras la definición de un lema van identificados con el nombre del corpus de procedencia entre corchetes. Siguiendo con la misma tónica que en los niveles precedentes, los ejemplos vienen a demostrar diferentes aspectos relevantes en torno a cada una de las locuciones. Estos son: variantes léxicas, variantes morfológicas, elementos facultativos, verbos habituales de combinación sintagmática y elementos de contorno. Por ejemplo, en el lema *{al/por el/todo lo} contrario* se reflejan dos de sus variantes léxicas: ‘■ [...]’<sup>28</sup> observó que a Tristana no le causaba disgusto. Al contrario, en ocasiones parecía ser de la misma opinión, [...] [*CORDE*]; ■ Vivir solo tiene sus ventajas: no hay que explicarle nada a nadie y pude,

---

<sup>28</sup> Esto refleja que el ejemplo ha sido acertado por la escasa representatividad que suponía el mantenimiento de más elementos.

por el contrario, decirme a mí mismo en voz alta: “Esa mujer te necesita, Gálvez”. [CORPES XXI].

Ahora bien, en este nivel tampoco se olvida que “los ejemplos lexicográficos pueden tener varias finalidades, entre ellas la de ejemplificar los datos sintácticos explícitos” (Tarp, 2018, p. 374). Es, por esta razón, que se realiza un minucioso rastreo del ejemplo más conveniente en cada caso, pues como ya se anunció, los ejemplos no aparecen exclusivamente con el fin de plasmar el significado de una locución, sino que también se registran para ilustrar los datos sintácticos recogidos en la parte de los apuntes gramaticales. De ahí que en esta sección del AL se incluyan ejemplos ora controlados, ora de competencia, ora reales. De esta manera acontece en la entrada *de cabo a rabo* y la siguiente información sobre la naturaleza aspectual<sup>29</sup> de dicha locución adverbial: ‘VERBO ‘Se combina con verbos que expresan un proceso delimitado o que tienen un límite propio. Esto ocurre con verbos como *leer, recorrer, examinar, revisar*, etc.: *Recorrí la ciudad de cabo a rabo*’.

#### ▪ **Notas gramaticales**

Igualmente, las notas gramaticales amparan nuevos elementos informativos. Como resultado, aunque *grosso modo*, la información más relevante se aporte en el nivel B1-B2, a lo expuesto en este nivel, se incorpora lo siguiente:

1. △ ESTRUCTURA. A la combinatoria sintagmática ya incluida en el nivel B1-B2, se adjuntan puntualizaciones semánticas aportadas por el modo verbal seleccionado. A modo de ejemplo, en la locución adverbial **tal vez** se ofrece ‘en posición pospuesta al verbo se combina con indicativo: *Voy tal vez*; en posición antepuesta al verbo se admite tanto INDICATIVO como SUBJUNTIVO: *Tal vez voy/vaya*. El uso de subjuntivo expresa menor probabilidad’.
2. △ VERBO. Información sobre las conjugaciones verbales y carácter del verbo de la locución o de los verbos frecuentes así como el carácter de variación flexiva de determinadas locuciones, póngase por caso la locución

---

<sup>29</sup> A partir del *Manual* de la NGLÉ (2010, pp. 587-589), se asevera que los adverbios aspectuales enfatizan un suceso, expresan su reiteración o indican la fase en la que se encuentra el propio suceso.

adverbial *que se las pela* y su única alternativa *que se las pelan*. En particular, para dicha entrada se incluye la nota de ‘*que se las pela* solo es posible en 3.ª persona del singular o del plural: *los jugadores corren que se las pelan*’.

3. △ REGISTRO DE SINÓNIMOS Y ANTÓNIMOS.
4. △ INTENSIFICADORES.
5. △ PUNTUACIÓN.
6. △ POSICIÓN.
7. △ ORTOGRAFÍA.
8. △ MODALIDAD ENUNCIATIVA. Especialmente, este apartado se destina a las locuciones de polaridad negativa. Un ejemplo es la locución adverbial *ni {loco/loca}* y la referencia de ‘aparece con oraciones de modalidad negativa: *No haré esto ni loco*’. Otra huella de la modalidad enunciativa negativa se atisba en *ni en broma*.
9. △ NORMA. En el último apartado de las notas gramaticales se acoge información de carácter normativo. Se observa claramente en el lema *en contra* y en el dato de: ‘En la estructura *en contra* + *pronombre posesivo* es incorrecto el uso de la forma masculina *\*en contra mío*. En (2) un sinónimo incorrecto es *\*por contra*’.

#### ▪ Pragmática

La carga pragmática de las locuciones supone para el estudiante de ELE y para las obras lexicográficas didácticas (Santamaría Pérez, 2011, p. 182) una orientación fundamental si se persigue el fin de usarlas con firmeza en el discurso y, con cierta cautela, nos atreveríamos a decir que es uno de los caminos idóneos para evitar los máximos malentendidos entre los hablantes.

Por ello, aunque por primera vez en la FL se incluyen unas observaciones específicas de carácter pragmático, se ha de tener presente que, tal y como apuntó Penadés Martínez (2018), la información pragmática se viene extendiendo a lo largo de todo el artículo (elementos de contorno, marcas diafásicas, carácter animado o inanimado del sujeto, entre otros), eso sí, de manera más indirecta o absorbida por otros objetivos lexicográficos.

En función de las recomendaciones aportadas por Penadés Martínez (2015, p. 271), se presenta seguidamente la información obtenida tras la pestaña despegable de PRAGMÁTICA:

1. △ FUERZA ILOCUTIVA. En el caso de la locución **para colmo** cuenta con el apunte pragmático de ‘enfatisa la molestia que supone un acción’. Por su parte, la locución **largo y tendido** posee la información de ‘enfatisa el tiempo de desarrollo de una acción’.
2. △ FUNCIÓN PRAGMÁTICA. Por ejemplo, en la locución adverbial **a saber** tiene el apunte de ‘función pragmática de conector discursivo explicativo. Se usa para explicar, aclarar o esclarecer una idea anterior’. Otra muestra está en la locución **por añadidura** y su nota de ‘función pragmática de conector discursivo de precisión o particularización. Se usa para añadir, incorporar una nueva idea relacionada con la anterior’.

En la imagen 6 se retoma lo dicho para la FL de nivel C1-C2 en vínculo con la locución *en contra*, al igual que se hizo para los dos niveles precedentes.

Nivel C1/C2	contra
-------------	--------

**en contra** (de algo o de alguien) [locución adverbial]

**1 en oposición** (a algo o a alguien)

- *Estados Unidos fue uno de los países que estuvo en contra del Tribunal Interamericano de Justicia, ya que los militares estadounidenses se opusieron a que su personal quedase bajo jurisdicción internacional.* [CORPES XXI]
- *Para estos atormentados, los argumentos más razonables en contra de matarse, parecen triviales, absurdos.* [CORPES XXI]

≈ **frente a**  
≠ **a favor, de acuerdo, en pro**

**2 contrariamente** (a algo)

- *Pero como en aquella casa todo era raro, y pasaban las cosas en contra de lo usual y corriente, se guardó su sorpresa y no dijo nada.* [CORDE]
- *Sin embargo, y en contra de lo que se ha dicho, no es cierto que nos importe demasiado que dejen la tapa del váter levantada.* [CORPES XXI]

≈ **{al/por el/todo lo} contrario, en cambio**

▼ **GRAMÁTICA**

△ **ESTRUCTURA**  
en contra + de + artículo definido/indefinido + nombre  
*en contra del gobierno corrupto/ en contra de un gobierno corrupto*

<p>en contra + de + infinitivo <i>en contra <u>de</u> trabajar por un salario tan bajo</i></p> <p>en contra+ de + lo + adjetivo <i>en contra <u>de</u> lo clásico</i></p> <p>en contra + de + lo que + verbo en indicativo <i>en contra <u>de</u> lo que pasa</i></p> <p>en contra + de + pronombre posesivo <i>en contra <u>de</u> mí</i></p> <p>en contra + pronombre posesivo <i>en contra <u>mía</u></i></p> <p>en +determinante posesivo + contra <i>en <u>mi</u> contra</i></p> <p>△ <b>VERBO</b> Normalmente aparece con los verbos <b>estar, ponerse, situarse, posicionarse, ubicarse, declararse</b></p> <p>△ <b>NORMA</b> En la estructura <i>en contra + pronombre posesivo</i> es incorrecto el uso de la forma masculino *<i>en contra mío</i>. En (2) un sinónimo incorrecto *<i>por contra</i></p> <p>▼ <b>PRAGMÁTICA</b></p> <p>△ <b>FUNCIÓN DISCURSIVA</b> En (2) se tiene la función pragmática de conector discursivo contraargumentivo</p>
---

Imagen 6. Lema “en contra”. Nivel C1-C2

## 5. CONCLUSIONES

A lo largo de esta investigación se ha podido constatar cómo se ha elaborado la microestructura del *DicLELE* en función del nivel de lengua del usuario. En particular, se ha querido reflejar de qué forma la FL ha sufrido su característico desarrollo desde el nivel A1-A2, avanzando por el nivel B1-B2, hasta llegar el nivel C1-C2 por medio de diversos lemas de naturaleza adverbial. En definitiva, se ha puesto en relieve el peso que supone la nivelación lingüística e informativa aportada en la FL del *DicLELE*, ya que “desde el punto de vista de la lexicografía de aprendizaje [...] lo que le interesa son los cambios cualitativos que cambian el carácter de las soluciones lexicográficas que corresponden a las necesidades lexicográficamente relevantes de los usuarios” (Tarp, 2006, p. 313).

Como ya apuntaron Maldonado González (2017) y Penadés Martínez (2015), en un diccionario didáctico como el *DicLELE*, la representatividad del leuario se cree que es esencial. Así, se ha recurrido a un número significativo de manuales de ELE de todos los niveles y del *PCIC*, obra de referencia para todos los agentes implicados (profesores, alumnos, editoriales, lexicógrafos) en la enseñanza de ELE. Asimismo, el

*DicLELE*, puesto que está *en ciernes*, no se cierra a posibles ampliaciones futuras, aunque siempre acordes a las directrices fraseológicas aquí defendidas.

En suma, el *DicLELE*, ante todo, pretende ser considerada como una herramienta lexicográfica digital representada por una interfaz intuitiva, fácil y manejable por sus usuarios potenciales, ya que “de momento no se conoce ningún diccionario digital para aprendices extranjeros del español que aproveche plenamente las tecnologías disponibles para ofrecer un servicio dinámico a sus usuarios con soluciones tanto monolingües como bilingües” (Tarp, 2018, p. 371).

## BIBLIOGRAFÍA

- Ahumada Lara, I. (1989). *Aspectos de la lexicografía teórica. Aplicaciones al Diccionario de la Real Academia Española*. Granada: Universidad de Granada.
- Bosque Muñoz, I. (1982). Sobre la teoría de la definición lexicográfica. *Verba*, 9, pp. 105-123.
- Castillo Carballo, M.ª A. (2003). La macroestructura del diccionario. En Medina Guerra (Coord.), *Lexicografía española*. Barcelona: Ariel, pp. 79-101.
- Consejo de Europa Consejo de Europa. (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid: MECO y Anaya.
- Fuertes-Olivera, P. A., y H. Bergenholtz (2018). Diccionarios del español para la producción de textos. *RILEX*, 1(I), pp. 5-28.
- Garriga Escribano, C. (2003). La microestructura del diccionario: las informaciones lexicográficas. En Medina Guerra (coord.), *Lexicografía española*. Barcelona: Ariel, pp. 103-126.
- González Aguiar, M. I. (2003). Fraseología y lexicografía: análisis y propuestas. *Revista de Lexicografía*. La Coruña: Servicio de Publicaciones de la Universidad de La Coruña, 9, pp. 29-55.
- Haensch, G. y C. Omeñaca (2004). *Los diccionarios del español en el siglo XXI*. Universidad de Salamanca: Ediciones Universidad Salamanca.
- Instituto Cervantes. (2006). *Plan Curricular del Instituto Cervantes. Niveles de Referencia*. Madrid: Instituto Cervantes, Biblioteca Nueva.
- Jacinto García, E. (2015). *Forma y función del diccionario. Hacia una teoría general del ejemplo lexicográfico*. Universidad de Jaén. Servicio de Publicaciones de la Universidad de Jaén.
- Leal Riol, M. J. (2013). Estrategias para la enseñanza y aprendizaje de la fraseología en español como lengua extranjera. *Paremia*, 22, pp. 161-170.
- Maldonado González, C. (2017). El diccionario didáctico monolingüe y bilingüe. En Domínguez Vázquez y Sanmarco Bande (Eds.), *Lexicografía y didáctica: diccionarios y otros recursos lexicográficos en el aula*. Frankfurt: Peter Lang, pp. 161-176.
- Medina Guerra, A. M. (Coord.). (2003). *Lexicografía española*. Barcelona: Ariel Lingüística.
- Mellado Blanco, C. (2004). *Fraseologismos somáticos del alemán*. Frankfurt: Peter Lang.
- Mellado Blanco, C. (2018). El contorno lexicográfico en fraseología: naturaleza y ubicación en la microestructura de los diccionarios generales y fraseológicos. En Mogorrón Huerta y Albaladejo-Martínez (Eds.), *Fraseología, diatopía y traducción*. Ámsterdam: John Benjamins Publishing Company, pp. 72-93.
- Mendoza Puertas, J. D. (2013). Las locuciones en los diccionarios de ELE: las dificultades del usuario. *Tonos digital: Revista electrónica de estudios filológicos*, 24, pp. 1-22.

- Montoro del Arco, E. T. (2006). *Teoría fraseológica de las locuciones particulares. Las locuciones prepositivas, conjuntivas y marcadoras en español*. Frankfurt: Peter Lang.
- Olímpio de Oliveira Silva, M. E. (2007). *Fraseología teórica y práctica*. Frankfurt: Peter Lang.
- Penadés Martínez, I. (1999). *La enseñanza de las unidades fraseológicas*. Madrid: Arco/Libros.
- Penadés Martínez, I. (2012). *Gramática y semántica de las locuciones*. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá.
- Penadés Martínez, I. (2013). La imagen subyacente a las locuciones como criterio de marcación diafásica. En Olza Moreno y Manero Richard (eds.), *Fraseopragmática*. Berlín: Frank & Timme, pp. 23-48.
- Penadés Martínez, I. (2015). *Para un diccionario de locuciones. De la lingüística teórico a la fraseografía práctica*. Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá.
- Penadés Martínez, I. (2018) (en prensa). La información pragmática sobre las locuciones en el diccionario. *I Simposio Internacional de Fraseología. Tendencias Actuales de investigación en fraseopragmática y fraseodidáctica*. Murcia: Universidad de Murcia.
- Porto Dapena, J. Á. (2002). *Manual de técnica lexicográfica*, Madrid: Arco Libros.
- Real Academia Española y Asociación de Academias de la Lengua Española. (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Real Academia Española y Asociación de Academias de la Lengua Española. (2010). *Manual de la Nueva gramática de la lengua española*. Madrid: Espasa.
- Real Academia Española y Asociación de Academias de la Lengua Española. (2010). *Ortografía de la lengua española*. Madrid: Espasa.
- Rey Castillo, M. (2017). Criterios macroestructurales y microestructurales en los diccionarios monolingües para el aprendizaje del léxico en español como lengua extranjera: entre la teoría y la práctica. *MarwoELE*, 25, pp. 1-34.
- Rodríguez Barcia, S. (2016). *Introducción a la lexicografía*. Madrid: Editorial Síntesis.
- Ruiz Gurillo, L. (2001). *Las locuciones en español actual*. Madrid: Arco/Libros.
- Ruiz Gurillo, L. (2005). Las locuciones marcadoras del español: análisis y aplicaciones. En Trives, Wotjak y Almela (Eds.), *Fraseología contrastiva: Con ejemplos tomados del alemán, español, francés e italiano*. Murcia: Ediciones de la Universidad de Murcia, Editum, pp. 242-257.
- Ruiz Gurillo, L. (2010). El tratamiento de la fraseología en el Diccionario de Partículas Discursivas del español. En Mellado *et al.* (eds.): *Nuevas perspectivas de la fraseología del siglo XXI*. Berlín: Frank & Timme, pp. 211-230.
- Santamaría Pérez, I. (2011). Entre la Pragmática y la Lexicografía: la marca humorístico en los diccionarios monolingües de aprendizaje del español. *Revista de Lexicografía*, 17, 217-240.
- Solano Rodríguez, M. A. (2007). El papel de la conciencia fraseológica en la enseñanza y aprendizaje de una lengua extranjera. *Les expressions figées en didactique des langues étrangères*, pp. 201-211.
- Tarp, S., y Fuertes-Olivera, P. A. (2016). Advantages and Disadvantages in the Use of Internet as a Corpus: The Case of the Online Dictionaries of Spanish Valladolid-Uva. *Lexikos*, 26 (1), pp. 273-295.
- Tarp, S. (2006). Lexicografía de aprendizaje. *Cadernos de tradução*, 2(18), pp. 295-317.
- Tarp, S. (2018). El desafío de los datos sintácticos en los diccionarios para aprendices extranjeros del español. *Revista de Filología*, 36, pp. 367-391.
- Vertikova, M. (2012). *El uso de ilustraciones en los diccionarios de ELE. Las ventajas*. Islandia: Universidad de Islandia.

### Diccionarios

- Maldonado González, C. (Coord.) (2012). *Diccionario avanzado*. Madrid: SM.
- Maldonado González, C. (Coord.) (2012). *Diccionario básico*. Madrid: SM.
- Maldonado González, C. (Coord.) (2012). *Diccionario secundaria y bachillerato*. Madrid: SM.

SECCIÓN: LEXICOGRAFÍA  
NIVELACIÓN DEL ARTÍCULO LEXICOGRÁFICO EN UN DICCIONARIO MONOLINGÜE DE APRENDIZAJE:  
EL CASO DE LAS LOCUCIONES ADVERBIALES  
M.ª Pilar Valero Fernández y Ivana Lončar

- Maldonado González, C. (Dir.) (2012). *Clave. Diccionario de uso del español actual*. Madrid: SM. [Clave]
- Moreno Moreno, M.ª A., Contreras Izquierdo, N. M., Torres Martínez, M., Camacho Niño, J., e I. Ruiz Sánchez (2018). *Diccionario temático infantil*. Disponible [en línea] [www.enredate.es](http://www.enredate.es) [EnREDate]
- Penadés Martínez, I. (2002). *Diccionario de locuciones verbales para la enseñanza del español*. Madrid: Arco/Libros. [DICLOCVER]
- Penadés Martínez, I. (2005). *Diccionario de locuciones adverbiales para la enseñanza del español*. Madrid: Arco/Libros. [DICLOCADV]
- Penadés Martínez, I. (2008). *Diccionario de locuciones nominales, adjetivas y pronominales para la enseñanza del español*. Madrid: Arco/Libros. [DICLOCNAP]
- Penadés Martínez, I. (2019). *Diccionario de locuciones idiomáticas del español actual*. Disponible [en línea] <http://www.diccionariodilea.es/> [DiLEA]
- Real Academia Española. (2014). *Diccionario de la lengua española* (23ª ed.). Madrid: Espasa. [DLE]
- Seco, M., Andrés, O., y Ramos, G. (2018). *Diccionario fraseológico documentado del español actual: Locuciones y modismos españoles*. Madrid: Aguilar. [DFDEA]

### Corpus

- REAL ACADEMIA ESPAÑOLA: Banco de datos (CORPES XXI). *Corpus del Español del siglo XXI*. Disponible [en línea] <http://web.frl.es/CORPES/view/inicioExterno.view>
- REAL ACADEMIA ESPAÑOLA: Banco de datos (CORDE). *Corpus diacrónico del español*. Disponible [en línea] <http://www.rae.es>
- REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA). *Corpus de referencia del español actual*. Disponible [en línea] <http://www.rae.es>



## TRATAMIENTO DE LAS DEFINICIONES DE LOS CONECTORES DISCURSIVOS EN LA LEXICOGRAFÍA DIDÁCTICA CUBANA. EL CASO DEL *DICCIONARIO BÁSICO ESCOLAR*

### PROCESSING OF THE DEFINITIONS OF DISCURSIVE CONNECTORS IN CUBAN DIDACTIC LEXICOGRAPHY. THE CASE OF *THE DICCIONARIO BÁSICO ESCOLAR*

Kelly Linares

*Instituto de Literatura y Lingüística "José Antonio Portuando Valdor"*  
[klinares89@gmail.com](mailto:klinares89@gmail.com)

#### RESUMEN

El interés que despiertan los *marcadores del discurso* (MD) dentro del proceso de enseñanza-aprendizaje de lenguas; el escaso análisis de su tratamiento en los diccionarios escolares hispanoamericanos y la carencia de investigaciones lexicográficas sobre MD en Cuba han motivado la realización de la presente investigación que tiene como objetivo principal examinar los tipos de definición empleados para un grupo de MD, los *conectores*, en un diccionario escolar cubano. En él se repasan las características semántico-pragmáticas, señaladas por autores de referencia como Portolés (1999, 2001), Martín Zorraquino (1999, 2004), Fuentes (1996, 2003, 2009), Montolío (2001), de cada conector registrado en el diccionario con la finalidad de valorar si se incluyen o no en las definiciones objeto de análisis. Además, se evalúa en qué medida se ajusta el tratamiento de los conectores al desiderátum teórico de que como unidades léxicas con significado de procesamiento sus definiciones han de ser funcionales o explicativas.

**Palabras clave:** marcadores del discurso, conectores, diccionarios escolares, definición, lexicografía, pragmática

#### ABSTRACT

The interest aroused by discourse markers (MD) within the process of teaching--learning languages; the scant analysis of its treatment in Spanish-American school dictionaries and the lack of lexicographical research on MD in Cuba have motivated the realization of the present investigation whose main objective is to examine the definitional types used for a group of MD, the connectors, in a Cuban school dictionary. It reviews the semantic-pragmatic characteristics, identified by reference authors such as Portolés (1999, 2001), Martín Zorraquino (1999, 2004), Fuentes (1996, 2003, 2009), Montolío (2001), of each connector registered in the dictionary in order to assess whether or not they are included in the definitions analysis. In addition, it is evaluated to what extent the treatment of the connectors is adjusted to the theoretical desideratum that as lexical units with processing meaning their definitions must be functional or explanatory.

**Keywords:** discourse markers, connectors, school dictionaries, definition, lexicography, pragmatics

## 0. INTRODUCCIÓN

A partir de los años sesenta y, fundamentalmente, en la década del 70 se desarrollan disciplinas lingüísticas como la lingüística del texto, la pragmática y el análisis del discurso, que no tienen la oración como nivel último analizable, sino que extienden su campo de análisis a niveles supraoracionales como el texto, el discurso, la conversación. Para López Serena y Borreguero (2010), a partir de la aparición de estas disciplinas se produce un salto desde el estudio inmanentista de la lengua hasta el pragmático-discursivo: “una necesaria inflexión desde la abstracción excesiva del sistema y la insuficiencia de la oración como nivel superior de análisis a la variación y organización propia del uso comunicativo” (2010, p. 382).

Por consiguiente, unidades lingüísticas como los *marcadores del discurso* (MD), que eran desterradas de los estudios gramaticales de la oración, fueron analizadas a partir de los diversos aparatos teóricos y metodológicos propuestos por las disciplinas antes mencionadas.

Los diversos estudios sobre los MD se han ocupado de su definición, clasificación, descripción y tratamiento lexicográfico. Esta última labor se ha dificultado por varios motivos, entre ellos: a) la pertenencia de estas unidades lingüísticas a diferentes categorías gramaticales; b) su heterogeneidad terminológica ligada a sus diversos abordajes teóricos; c) la complejidad de su clasificación funcional<sup>1</sup>; d) la naturaleza instruccional o de procesamiento de su significado<sup>2</sup>; e) su configuración prosódica y movilidad posicional (Martín Zorraquino, 2004, pp. 53-55; Vázquez, 2000). Sin

---

<sup>1</sup> Su clasificación funcional ha sido tratada, por una parte, desde una perspectiva onomasiológica (a partir de la existencia de diversos actos de habla, los MD se adscriben a funciones textuales vinculadas a estos, por ejemplo, advertencia, refutación, énfasis, etc.) y, por otra, semasiológica (a partir de la caracterización morfosintáctica, semántica y pragmática de los MD se crean grupos según sus equivalencias funcionales en el discurso, lo cuales permiten “determinar, con mejores garantías, las regularidades del comportamiento lingüístico de las unidades acotadas” (Martín Zorraquino, 2004, p. 54).

<sup>2</sup> Consiste en “una serie de instrucciones semánticas que guían las inferencias que se han de efectuar de los distintos miembros del discurso en los que aparecen estas unidades” (Martín Zorraquino y Portolés, 1999, p. 4071). Si bien la división realizada por Blakemore (1987) entre *significado conceptual* y *procedimental* ha recibido críticas pues las palabras con *significado conceptual* pueden, también, orientar el procesamiento de los enunciados y algunos MD, aun cuando presentan *significado procedimental*, conservan, en parte, un *significado conceptual* (Portolés [1998] 2001).

embargo, a pesar de estas dificultades, se pondera la importancia de la aparición de los MD en la nomenclatura de los diccionarios y la necesidad de su adecuado tratamiento.

El auge de investigaciones lexicográficas de los MD en el ámbito hispánico, fundamentalmente en España, no se produce hasta la década del 90 del pasado siglo. Los acercamientos teóricos, descriptivos y lexicográficos al tema facilitaron y permitieron la elaboración de diccionarios específicos que sistematizan la información disponible sobre MD en artículos lexicográficos. Entre estos se encuentran el *Diccionario de partículas* de Santos Río (2003), el *Diccionario de partículas discursivas del español (DPDE)*, coordinado por Briz, Portolés y Pons (2008), el *Diccionario de conectores y operadores del español* de Fuentes (2009) y, más recientemente, el *Diccionario de Marcadores Discursivos para Estudiantes de Español como Lengua Extranjera (DIMAELE)* de Holgado (2018). Los niveles de especialización de estos repertorios léxicos (que se traslucen en el mayor o menor empleo de metalenguaje, en los contenidos de los artículos lexicográficos y su grado de formalización y condensación) difieren en correspondencia con su función, determinada por el público meta y los propósitos perseguidos por los diccionaristas (finalidad), por ejemplo: realizar un diccionario para la codificación de mensajes o para su interpretación o descodificación.

A la proliferación de estudios, que se ocupan del tratamiento lexicográfico de los MD en diccionarios tipológicamente diferentes, y a la creación de diccionarios especializados se suma el interés por la enseñanza de este tipo de unidades lingüísticas en las clases de lenguas, tanto de primeras como de segundas. En relación con su legitimidad en las clases de español como lengua extranjera (E/LE), destaca Martín Zorraquino que los MD cumplen diversas funciones discursivas; forman parte de la competencia comunicativa de los hablantes; son palabras frecuentes con acepciones diversas, por lo tanto, necesarias para la construcción del discurso y la expresión de la subjetividad del hablante, y los errores que los aprendices de L2 cometen en su uso demuestran la impericia en el dominio de la lengua que se desea manejar como hablante nativo (Martín Zorraquino, 2004, pp. 56-57). Debemos añadir que esta impericia en su uso la demuestran no solo los aprendices de E/LE sino también los estudiantes de

español como lengua materna (E/LM), que hacen una pobre utilización de ellos (Vázquez, 2008).

Dentro del proceso de enseñanza-aprendizaje de lenguas los diccionarios se erigen como una herramienta fundamental para el estudio de los MD. Por lo tanto, es importante que estas unidades lingüísticas aparezcan y sean tratadas de forma adecuada no solo en los diccionarios especializados, sino también en los generales de la lengua y en los didácticos, aun cuando los fines de estos repertorios lexicográficos son diferentes debido sus distintas naturalezas.

Los acercamientos a los MD desde una perspectiva metalexigráfica son numerosos en la parcela de la enseñanza de E/LE, según la bibliografía consultada. Sin embargo, el panorama resulta más desolador en la de E/LM, pues se ha examinado escasamente el tratamiento que los diccionarios escolares hispanoamericanos dan a los MD.

Por otra parte, en Cuba son escasos los trabajos sobre MD. Estos derivan, fundamentalmente, de proyectos de investigación como el “Proyecto de estudio de la norma lingüística culta hispánica Juan M. Lope Blanch”, y “Habana, Español Coloquial” (Hab.Es.Co, 2012). Han estado dirigidos, fundamentalmente, hacia la descripción del comportamiento de MD en el habla culta o coloquial de la variedad cubana del español. Destacan los acercamientos de Perdomo (2013), González Mafud y Perdomo (2014), Martínez Abreu (2015) y García Roche (2018).

El panorama presentado (el interés que despiertan los MD dentro del proceso de enseñanza-aprendizaje de lenguas; el escaso análisis de su tratamiento en los diccionarios escolares hispanoamericanos y la carencia de investigaciones lexicográficas sobre MD en Cuba) ha motivado la realización de esta investigación que se centrará en el examen de los tipos definición empleados para un grupo de MD, los conectores, en un diccionario escolar cubano: *Diccionario básico escolar (DBE)* del Centro de Lingüística Aplicada de Santiago de Cuba (2014), dirigido por Miyares Bermúdez.

Se ha optado por el estudio de la definición por ser un elemento medular en el artículo lexicográfico y uno de los que comporta mayor dificultad en su elaboración, sobre todo en los diccionarios escolares, si se tiene en cuenta que sus destinatarios son

hablantes nativos que se encuentran en un momento de aprendizaje y afianzamiento de la lengua materna.

La presente investigación se inserta dentro de los estudios sobre MD desde la perspectiva de la lexicografía aplicada a la didáctica. Se aproxima a estas unidades lingüísticas, específicamente a los conectores, desde un enfoque semántico-pragmático a partir de los estudios teóricos desarrollados, fundamentalmente, por los hispanistas Portolés (1999, 2001) y Martín Zorraquino (1999, 2004), aunque no se desatienden los aportes al tema de otros estudiosos como Fuentes (1996, 2003, 2009) o Montolío (2001). Para el análisis lexicográfico se siguen, principalmente, los postulados teóricos de Porto Dapena (2002, 2014).

## 1. SOBRE LOS MARCADORES DEL DISCURSO: LOS CONECTORES

El término *marcador del discurso*, popularizado por la autora norteamericana Schifffrin (1987) en su libro *Discourse markers*, es de los más empleados en los estudios lingüísticos hispánicos. En los diferentes abordajes investigativos del tema se constata que este término coexiste con otros, por ejemplo, con el término *conector*, en una relación sinonímica o es tomado como un hiperónimo “referido a casi cualquier clase de palabras sin significado proposicional” (Pons, 1998, s/p). En esta investigación, siguiendo a Martín Zorraquino y Portolés (1999), se adopta el término *marcador del discurso* como hiperónimo, pues se coincide con estos autores cuando declaran:

Sólo una parte de los marcadores conectan, no lo hacen, por ejemplo, los operadores argumentativos y mucho menos los ‘marcadores conversacionales’. Ciertamente, los marcadores pueden relacionar una oración con otras unidades externas a ella, pero también es frecuente que relacionen unidades intraoracionales o unidades que pertenecen a categorías sintagmáticas distintas de la oración. Los marcadores, en fin, cohesionan el discurso, pero esta es sólo una de sus posibles propiedades (1999, p. 4057, nota al pie 6).

Por lo tanto, términos como *conector* y *operador* se refieren en este trabajo a tipos de MD<sup>3</sup>. Asimismo se adopta que:

[Los] marcadores del discurso son unidades lingüísticas invariables, no ejercen una función sintáctica en el marco de la predicación oracional —son, pues, elementos

---

<sup>3</sup> Para consideraciones diferentes a las adoptadas en este trabajo sobre los términos *marcador del discurso*, *conector* y *operador discursivo*, véanse Pons (1998) y Fuentes (2003).

marginales— y poseen un cometido coincidente en el discurso: el de guiar, de acuerdo con sus distintas propiedades morfosintácticas, semánticas y pragmáticas, las inferencias que se realizan en la comunicación (Martín Zorraquino y Portolés, 1999, p. 4057).

Los autores antes citados, a partir de una propuesta de clasificación que se basa en buscar para el marcador un significado unitario que dé cuenta de todos sus usos, distinguen cinco tipos de MD:

1. estructuradores de la información: *pues, en primer lugar/ en segundo lugar, por cierto, a propósito*, entre otros.
2. conectores: *además, encima, en consecuencia, entonces, sin embargo*, entre otros.
3. reformuladores: *es decir, de todos modos, en conclusión, en fin*, entre otros.
4. operadores argumentativos: *en el fondo, de hecho, por ejemplo, en particular*, entre otros.
5. marcadores conversacionales: *bueno, claro, hombre, mira, eh*, entre otros.

### 1.1. LOS CONECTORES

Es de interés para este trabajo el grupo de los *conectores*. Si bien autores como Gili y Gaya (1943), Pons (1998), Portolés ([1998] 2001), Briz (1993) reconocen que las conjunciones coordinantes (*y, pero, sino*) y la subordinante *aunque* son los conectores prototípicos —siguiendo las palabras de Pons (1998)—, en este estudio no se incluyen dentro del inventario de los MD porque, aunque vinculan semánticamente dos miembros del discurso y orientan el proceso inferencial, no cumplen con todas las características morfosintácticas que se han adoptado para distinguir los MD, entre ellas:

1. Son unidades lingüísticas invariables (no conservan la capacidad de flexión y de combinación de su miembros. Por tal razón *por consiguiente* es un MD y no así *por este motivo*):
  - (1) a. Juan está muy ocupado y, *por consiguiente*, no te podrá acompañar al teatro.
  - b. Juan está muy ocupado y, *por este motivo*, no te podrá acompañar al teatro.
  - (2) a. \*Juan está muy ocupado. Tiene que entregar un informe, leer los trabajos de los alumnos y evaluarlos. *Por consiguientes*, no te podrá acompañar al teatro.
  - b. Juan está muy ocupado. Tiene que entregar un informe, leer los trabajos de los alumnos y evaluarlos. *Por estos motivos*, no te podrá acompañar al teatro.
2. Se deben a un proceso de gramaticalización, aunque no todos presentan el mismo grado de gramaticalización. Señalan Martín Zorraquino y Portolés

(1999, p. 4060) que “los marcadores del discurso proceden de la evolución de una serie de sintagmas que, de una parte, van perdiendo sus posibilidades de flexión y combinación, y, de otra, van abandonado su significado conceptual y se especializan en otro de procesamiento”.

3. Presentan movilidad sintáctica<sup>4</sup>.

- (3) a. Está muy callada. *Por cierto*, ¿le comentaste sobre aquel asunto?  
b. Está muy callada. ¿Le comentaste, *por cierto*, sobre aquel asunto?  
c. Está muy callada. ¿Le comentaste sobre aquel asunto?, *por cierto*.

4. No modifican ningún elemento dentro de la sintaxis oracional.

- (4) Lo puso todo *encima* de la mesa.

*Encima* funciona como adverbio de lugar que modifica al verbo.

- (5) Llegas tarde a clases y, *encima*, no haces la tarea.

*Encima* funciona como marcador del discurso (conector).

5. No reciben especificadores ni adyacentes complementarios<sup>5</sup>.

- (6) a. *En todo caso*, vive lejos.  
b. \**En todo caso* de esos, vive lejos.  
c. \*Muy *en todo caso*, vive lejos (Martín Zorraquino y Portolés, 1999, p. 4066)

6. No pueden ser negados.

- (7) a. No vino, *pues*, Juan a la fiesta, sino Antonio.  
b. No vino, *pues*, Juan a la fiesta, sino que se quedó en casa.  
c. No vino, *pues*, Juan a la fiesta, sino que ya estaba.  
d. No vino, *pues*, Juan a la fiesta, sino a la cena de después.  
e. \*No vino, pues, Juan a la fiesta, sino por tanto (Martín Zorraquino y Portolés, 1999, p. 4067).

7. No se coordinan entre sí (por ejemplo, \**además y sin embargo*.)

Martín Zorraquino y Portolés (1999) para definir los *conectores* tienen en cuenta las instrucciones argumentativas de su significado. Son los MD que “vinculan semántica y pragmáticamente un miembro del discurso con otro anterior, de tal forma que el

---

<sup>4</sup> Esta característica diferencia a los MD de las conjunciones de la gramática tradicional porque, aun cuando la noción de conexión les es intrínseca (Pons, 1998), estas siempre anteceden al sintagma que introducen ( Fuentes, 1996; Martín Zorraquino y Portolés, 1999). Dicha razón hace que conjunciones como *pero*, *y*, o *aunque* queden fuera de nuestro estudio.

<sup>5</sup> Excepto los MD *además*, *encima* y *aparte*, que pueden recibir complementos con *de* (Martín Zorraquino y Portolés, 1999, p. 4066).

marcador guía las inferencias que se han de efectuar del conjunto de los dos miembros discursivos conectados” (1999, p. 4080). La progresión discursiva no depende de la capacidad de los conectores para unir informaciones sino de su capacidad para orientar y restringir el proceso argumentativo del discurso.

Es por ello que solo incluyen dentro del grupo de los conectores los MD que tienen instrucciones de conexión y argumentativas en su significado. Por lo tanto, excluyen los temporales y ordenadores discursivos (Fuentes, 1996), los organizadores de la información (Montolío, 2001) y los reformulativos (Fuentes, 1996); los cuales no realizan una operación argumentativa y se limitan a actuar en el marco informativo, aun cuando tienen carácter fórico.

En esta investigación se asume que los *conectores* son MD que cohesionan dos miembros del discurso y actúan, debido a las instrucciones argumentativas de su significado, sobre las inferencias que han de efectuarse de los miembros que relacionan, al condicionar la aparición de unas inferencias y la negación de otras. En un ejemplo como:

(8) Luis es muy inteligente; *sin embargo*, suspendió el examen.

Los enunciados están antiorientados argumentativamente: el primero (*Luis es muy inteligente*) orienta hacia unas conclusiones contrarias a la que aparece en el segundo (*suspendió el examen*). Estos enunciados antiorientados se relacionan a través del conector *sin embargo*, el cual por su significado indica que contraria a la conclusión que se puede esperar del enunciado *es muy inteligente*, Luis *suspendió el examen*.

Se adopta en este trabajo la clasificación de los conectores de Martín Zorraquino y Portolés (1999), quienes los dividen en tres grupos de acuerdo con su significado:

1. *Conectores aditivos*: unen a un miembro discursivo anterior otro con su misma orientación argumentativa. De esta forma, posibilitan la inferencia de conclusiones que serían difíciles de obtener si los dos miembros permaneciesen independientes (1999, pp. 4093-4099).

(9) Tiene unas condiciones físicas increíbles y, *además*, es muy alta. Puede jugar en el equipo.

2. *Conectores consecutivos*: Presentan el miembro discursivo en el que se encuentran como una consecuencia del miembro anterior. (1999, pp. 4099-4109).

(10) Este sector económico empieza a entrar en crisis; *por tanto*, es el momento de diversificar nuestras inversiones (Montolío, 2001, p. 100).

3. *Conectores contraargumentativos*: unen dos miembros del discurso, de forma que el segundo suprime o atenúa alguna conclusión que pueda inferirse del primero (1999, pp. 4109-4121).

(11) Los críticos consideran que su libro es excelente; *sin embargo*, sus niveles de venta son bajos.

## 2. SOBRE EL *DICCIONARIO BÁSICO ESCOLAR* Y LA SELECCIÓN DE LA MUESTRA

En Cuba la *lexicografía didáctica* (LD)<sup>6</sup> desde una perspectiva práctica se ha desarrollado, principalmente, a partir de los estudios que se realizan en el Centro de Lingüística Aplicada (CLA) de Santiago de Cuba.

Entre los años 1990 y 1995 se llevó a cabo en el CLA el proyecto “Estudio del vocabulario activo-funcional del escolar cubano”<sup>7</sup>, dirigido por Miyares Bermúdez. Tuvo como finalidad “conocer el léxico de los educandos cubanos, así como el uso que hacen estos de las palabras, especialmente las comprendidas en la esfera conceptual —sustantivos, adjetivos, verbos y adverbios— sin dejar de analizar también las otras clases de palabras y las frases, locuciones y siglas presentes en el vocabulario de los niños” (Miyares Bermúdez *et al.*, 2006, p. 7). La muestra estuvo integrada por textos orales y escritos, realizados por estudiantes de 2.º a 6.º grados y de 10 territorios del país. En total se recogieron 700 000 palabras y 26 251 palabras diferentes que son la base fundamental del *Diccionario básico escolar* (DBE, en sus diversas ediciones).

---

<sup>6</sup> La LD es una disciplina, una rama de la lexicografía, vinculada con la elaboración de diccionarios para estudiantes. Propone Azorín (2000, p. 21) emplear “la denominación *Diccionario didáctico* como etiqueta genérica para designar a los diccionarios especialmente diseñados para los usuarios que se encuentran en el proceso de aprendizaje de una lengua”. Estos se dividen en dos grandes grupos según sus objetivos y usuarios potenciales: *diccionarios didácticos de español como lengua materna*, para usuarios en fase de aprendizaje de la lengua materna, y *diccionarios didácticos de español como lengua no materna*, para usuarios en fase de aprendizaje de una segunda lengua (Hernández Hernández, 1989, 2003, 2008; Azorín, 2000). A los primeros se los denomina *escolares* y dentro de los segundos se distinguen los *de aprendizaje* (diccionarios monolingües para extranjeros), los *bilingües* y los *semibilingües* (diccionarios mixtos que combinan los bilingües con los monolingües).

<sup>7</sup> De este proyecto deriva la obra *Léxico activo-funcional del escolar cubano* (2006).

El *DBE* es la obra más difundida de la lexicografía escolar en Cuba, muestra de ello son sus múltiples ediciones, reimpresiones, y su distribución en diversos formatos (papel, CD, internet, intranet nacional). Sus destinatarios son los escolares con edades comprendidas entre los 10 y los 18 años, es decir, los que se encuentran en segundo ciclo de la enseñanza primaria (4.º y 5.º grados), los de la enseñanza secundaria y los de la preuniversitaria y politécnica. Esta franja temporal resulta excesivamente amplia (Sánchez Muñoz, 2002-2003) y ha sido criticada por varios investigadores que han reseñado las diferentes ediciones de la obra (Forgas y Bargalló, 2005-2006, p. 254; Nomdedeu, 2012, pp. 245-246).

Uno de los aspectos que se echa en falta en la “Introducción” de esta obra es la declaración de las fuentes utilizadas para la conformación de su nomenclatura. En Ruiz Miyares y Tarp (2015) se declara que las 12 000 entradas se seleccionaron del corpus *Léxico activo-funcional del escolar cubano*, de libros de texto de diferentes grados, de periódicos y revistas cubanas dirigidas a los jóvenes, así como de un corpus basado en un estudio sobre el uso de los afijos por parte de estudiantes de 7.º a 9.º grados. Sin embargo, no quedan claros los criterios seguidos para la selección de estos textos y de las unidades léxicas que componen no solo la macroestructura sino también la microestructura (unidades léxicas sublematizadas) del diccionario.

La nomenclatura de este diccionario en su cuarta edición, versión papel, constituye el corpus de este trabajo.

En primer lugar, para la selección de la muestra: los conectores presentes en el *DBE*, se elaboró un listado de referencia a partir de los conectores inventariados en el capítulo 63 de la *Gramática descriptiva de la lengua española* (Martín Zorraquino y Portolés, 1999). A saber:

1. Aditivos: *además, encima, aparte, incluso, inclusive, por añadidura, es más.*
2. Contraargumentativos: *en cambio, por el contrario, al contrario, antes bien, sin embargo, no obstante, con todo, por contra, empero, ahora bien, ahora, eso sí.*
3. Consecutivos: *por (lo) tanto, por consiguiente, consiguientemente, consecuentemente, por ende, en consecuencia, de ahí, entonces, pues, así, así pues.*

Una vez establecido este listado se procedió a su búsqueda en el *DBE*. Se tuvo en cuenta no solo la aparición de la forma (como lema o sublema) sino también que se recogiera con el significado de MD. Los resultados obtenidos fueron los siguientes:

1. Aditivos: *además, encima*.
2. Contraargumentativos: *por el contrario, al contrario, sin embargo, con todo*.
3. Consecutivos: *por (lo) tanto, entonces, pues*.

Estos conectores registrados en el *DBE* constituyen la muestra de nuestra investigación<sup>8</sup>. Se debe aclarar que aparecen lematizadas otras formas como *incluso, inclusive, ahora, así*, pero no con el significado de conector discursivo, por lo tanto, no serán analizadas.

### 3. ASPECTOS GENERALES SOBRE LAS DEFINICIONES DE LOS MARCADORES DEL DISCURSO

Resulta recurrente la afirmación de que la tarea más ardua en el proceso de realización de un artículo lexicográfico de un MD es la elaboración de su definición (Vázquez, 2000; Martín Zorraquino, 2004; Sanmartín, 2010; Holgado, 2014).

Las *definiciones sinonímicas* de estas unidades léxicas han marcado el proceder de muchos diccionarios generales de lengua que no dan cuenta de los sentidos diversos de los MD, sino que en una entrada se limitan a ofrecer como significado otros MD que, o no son sinónimos del definido o no lo son en todos los contextos (Vázquez, 1995-1996). De igual forma, se incurre en el fenómeno de la *circULARIDAD*, pues el uso de este tipo de definición conduce a un recorrido de un MD a otro que actúa como definición y que a su vez es definido con el primer MD. El usuario que realiza la consulta culmina sin saber el significado de los MD buscados. También aparecen *definiciones perifrásticas* poco aclaratorias y deficientes y, en el peor de los casos, se llega a las *pistas perdidas*, pues aparece como definición de un MD otros que no se encuentran lematizados o sublematizados.

---

<sup>8</sup> La comparación con otras dos obras de similar naturaleza: el *Diccionario del estudiante. Secundaria y Bachillerato* (RAE, [2005] 2011) y el *Diccionario Secundaria y Bachillerato. Lengua Española* (2003) de Ediciones SM, dirigido por Maldonado, arroja que el *DBE* ha incluido una cantidad menor de conectores. Por ejemplo, el *Diccionario del estudiante...* de la RAE del listado de referencia recoge los siguientes conectores: *además, encima, incluso, por añadidura, en cambio, por el contrario, antes bien, sin embargo, no obstante, con todo, empero, ahora bien, ahora, por (lo) tanto, por consiguiente, consiguientemente, por ende, pues, así, así pues*.

Por esta razón, lo más recomendable es el empleo, para este tipo de unidades léxicas, de la *definición lexicográfica formal, funcional o explicativa*, es decir, la que se realiza en metalengua de signo y a través de la cual “en lugar del contenido, lo que se pone de manifiesto son características relativas a aspectos formales —funcionales, morfológicos, sintácticos o combinatorios, pragmáticos, etc.— de la palabra utilizada como *definiendum*” (Porto Dapena, 2014, p. 61). Esta debe ser empleada en la descripción de las palabras gramaticales, en las que predomina el significado procedimental y las que, por tanto, no se dejan someter a la *ley de la sinonimia*, que consiste en la “exigencia de que el definiente esté constituido por una forma adecuada a la función sintáctica propia del definido” (Seco, 1987, p. 21).

Acertadamente expresa Portolés ([1998] 2001, p. 153) que “a diferencia de las definiciones del léxico conceptual<sup>9</sup> la de los marcadores han de reflejar no lo que la palabra representa, sino lo que la fuerza en la comprensión del miembro del discurso en el que aparece o al que acompaña”. Aunque, como destaca Martín Zorraquino (2004, pp. 61-62), la dificultad que entraña construir este tipo de definición conduce al uso en muchas obras lexicográficas generales de la antes mencionada definición sinonímica.

Muchas de las deficiencias en la construcción de las definiciones de los MD han sido superadas en las obras lexicográficas especializadas. Por ejemplo, diccionarios de MD como el *DPDE* y el *DIMAELE* han optado por el uso de definiciones funcionales sencillas, sin tanta carga metalingüística, y han creado artículos lexicográficos que difieren de los que se encuentran en los diccionarios generales de lengua, pues contienen una mayor cantidad de apartados que ofrecen información sobre la prosodia, la sintaxis y los registros de uso de los MD.

Portolés ([1998] 2001, pp. 152-155) cuando se refiere al tipo de definición que debe emplearse para los MD presenta el modelo empleado en el *DPDE* y ejemplifica con la

---

<sup>9</sup> A las palabras con significado conceptual (sustantivos, adjetivos, verbos y algunos adverbios) le corresponde la *definición conceptual* que se formula en metalengua de contenido y en la que se expresa “en otras palabras de la misma lengua el contenido significativo o conceptual del definido” (Porto Dapena, 2002, p. 282). Esta puede ser *sinonímica*: el definidor es un sinónimo del definido o *perifrástica*: el definidor está constituido por una frase o sintagma (2002, p. 285).

definición del conector *sin embargo*: “Presenta el miembro del discurso en el que aparece como una conclusión ‘descriptor’. Esta conclusión es contraria a otra que se pudiera inferir a partir de un argumento anterior ‘diferenciador’”. Como vemos consta de un ‘descriptor’ que destaca la función discursiva que permite relacionar al MD con otros con la misma función, y de un ‘diferenciador’ que destaca lo que diferencia al MD de los otros que integran su mismo grupo.

A continuación analizaremos qué tipos de definiciones emplea el *DBE* para los conectores de la muestra. Se hará un breve repaso de las características semántico-pragmáticas de cada conector señaladas por algunas obras de referencia con el objetivo de valorar si se incluyen o no en las definiciones objeto de análisis. Además, se evaluará en qué medida se ajusta el tratamiento de los conectores al desiderátum teórico de que como unidades léxicas con significado de procesamiento sus definiciones han de ser funcionales o explicativas y no conceptuales.

#### 4. SOBRE LAS DEFINICIONES DE LOS CONECTORES EN EL *DICCIONARIO BÁSICO ESCOLAR*

##### 4.1. EL CASO DE *ADEMÁS*

*Además* es el conector de uso más frecuente dentro del grupo de los aditivos, así lo señalan autores como Fuentes (1996, p. 27), Martín Zorraquino y Portolés (1999, p. 4094), y Montolío (2001, p. 143). Apuntan Martín Zorraquino y Portolés (1999, p. 4094) que “vincula dos miembros del discurso con la misma orientación argumentativa<sup>10</sup>. El segundo facilita inferencias que sería difícil lograr únicamente del primero”. Este segundo miembro, puede aumentar la fuerza argumentativa del primero y se presenta, por lo general, como más relevante (Montolío 2001, p. 157).

En el *DBE* este conector se define como sigue:

(12) **además** adv. c. Significa también, aparte de.

La definición es sinonímica múltiple, pues se utilizan dos unidades léxicas como sinónimas del conector definido: *también* y *aparte de*. Por lo tanto, se emplea un tipo de

---

<sup>10</sup> Los miembros que vincula el conector aditivo *además* conducen a la misma conclusión.

definición conceptual para definir una palabra con significado de procesamiento: recurso, como ya expresamos, no recomendable.

También aparece explícito el verbo cópula *significa* como introductor del enunciado que funciona como definición de tipo conceptual; uso que se documenta en diccionarios antiguos como el *Tesoro* de Covarrubias de 1611 (Llopis, 2006, p. 1087 y Porto Dapena, 2014, p. 26), pero no es propio de los diccionarios modernos, los cuales sistematizan la información lexicográfica prescindiendo del verbo cópula, el cual solo se hace explícito en las definiciones funcionales o explicativas introducidas por *indica*, *presenta*, *introduce*, *sirve para*, *se usa para*, etc.

Por un lado, el camino que impone la definición de *además* en el *DBE* nos lleva a la entrada *también* que en su segunda acepción se define de manera sinonímica como *además*. El escolar que se enfrente a la búsqueda del significado del conector de la muestra terminará en el mismo punto donde comenzó su pesquisa. Una vez más, el fenómeno de la circularidad se hace presente en la práctica lexicográfica.

A pesar de que estas unidades léxicas se ofrecen como sinónimas y tienen desde el punto de vista semántico el significado básico de ‘añadir o sumar una información a la ya presentada o conocida con anterioridad’, no son intercambiables en todos los contextos<sup>11</sup>. Al respecto nos dice la *Nueva gramática de la lengua española* (NGLE, 2009, pp. 2362-2363):

En contra de agregar *además* al grupo de los adverbios focales, y a favor de considerarlo entre los conectores aditivos, está el hecho de que se construye con complementos preposicionales, sean expresos o tácitos [...]. El complemento preposicional (ya sea expreso, ya tácito), que caracteriza al adverbio *además* lo diferencia marcadamente de *también*, con el que no puede alternar en muchos contextos, en especial en los de foco nominal.

Por otro lado, la búsqueda de la otra unidad léxica que se pone como sinónimo (*aparte de*) nos lleva a la definición: ‘indica omisión’ que, aun cuando es funcional, resulta poco esclarecedora por su brevedad, y su significado no se corresponde con el del conector aditivo *además*.

---

<sup>11</sup> Véase la comparación que realiza Cuartero (1995) de las unidades léxicas *además* y *también* y las consideraciones de Fuentes (1996, p. 29).

Lo más llamativo es que en este caso se incurre en un error por omisión, porque el adverbio *aparte* sí funciona como conector aditivo (Martín Zorraquino y Portolés, 1999, p. 4096), pero el significado que tiene como tal no se registra en el *DBE*. De igual forma, no resultaría apropiado definir el conector discursivo más frecuente de los aditivos por remisión a uno que es semejante, pero menos frecuente.

Asimismo, la unidad léxica que se consigna como sinónimo es *aparte de* que, por su construcción, es afín con *además de*, combinación a la cual no se hace alusión en el artículo lexicográfico de *además*. Sobre estas combinaciones apuntan Martín Zorraquino y Portolés (1999) que, aun cuando permanecen con el mismo significado y la posición de inciso:

Pierden en estos casos su capacidad de conexión entre dos miembros del discurso, ya que el elemento al que se hacía referencia anafóricamente con el marcador es el que aparece ahora en el complemento: a. Es muy simpático y, *además*, cocina muy bien. b. *Además de ser muy simpático*, cocina muy bien. (p. 4066).

La remisión de una entrada a otra en la que no se encuentra el significado esperado, sino otra acepción de la palabra a la que se remite nos recuerda los errores en los que ha incurrido, otrora con mayor asiduidad que en la actualidad, la lexicografía didáctica hispana, porque la macroestructura de los diccionarios ha derivado, en algunos casos, de la reducción de obras lexicográficas mayores (diccionarios generales), lo que conduce a inconsistencias similares a la aquí presentada.

#### 4.2. EL CASO DE *ENCIMA*

El conector aditivo *encima* es más frecuente en la lengua hablada, en registros informales y en escritos poco formales, según se declara en alguna bibliografía (Montolío, 2001, p. 158; *DPDE*, 2008).

Al igual que *además* vincula dos miembros del discurso con la misma orientación argumentativa, pero a diferencia de este indica que el miembro que lo precede es suficiente para arribar a determinada conclusión. Asimismo, el argumento que introduce aumenta la escala argumentativa del argumento anterior y refuerza “el sentido de exceso, importancia, abuso o sorpresa” (*DPDE*, 2008) que de él se desprende; lo que lo diferencia también de *además*. A ello se suma que la valoración de

la información que se desprende de su uso es, por lo general, de carácter negativo (Fuentes, 1996, p. 27; Montolío, 2001, p. 158; *NGLE*, 2009: p. 2363). Al decir de Montolío (2001, p. 160), por este carácter subjetivizador y su relación con el registro oral espontáneo: “El uso de este conector no resulta frecuente —ni recomendable— en los textos que se presenten como objetivos [...], de hecho, en la lengua escrita, el uso de *encima* se restringe a los textos de opinión, es decir, aquellos en los que el género textual admite que el autor exprese su propia evaluación”.

En el diccionario que nos ocupa se define el conector *encima* como se muestra a continuación:

(13) **encima** adv. 1 [...] 3 Además, por si fuera poco.

En este caso, observamos que la definición es igualmente sinonímica. En primer lugar se remite al conector más frecuente de la serie de los aditivos *además*; lo que lleva, como vimos anteriormente, a la circularidad y a la obtención de un significado erróneo.

En segundo lugar, se da como significado una expresión (*por si fuera poco*) cuyo estatuto como conector aditivo vacila, pues no toda la bibliografía consultada la considera como tal. No se hace alusión a esta expresión en Fuentes (1996), Martín Zorraquino y Portolés (1999) y Portolés ([1998] 2001); Montolío (2001, p. 161) se refiere a ella como una secuencia que, junto a *para colmo*, presenta un valor semejante al del MD *encima*; y la *NGLE* (2009) la presenta como una “fórmula asimilable a las locuciones adverbiales” (2009, p. 2356) y como un “adverbio conector aditivo y de precisión y particularización” (2009, p. 2361) al igual que *además*, *encima*, *asimismo*, *por añadidura*, entre otros. En esta misma obra más adelante se explica que “*encima* y *por si fuera poco* añaden información que se sitúa en un punto superior de la línea argumentativa introducida en el texto que los precede” (2009, p. 2361), y que la interpretación de *por si fuera poco* es más literal que la de *encima*, “donde *poco* se interpreta como ‘de peso argumentativo escaso e insuficiente’” (2009, p. 2361). El *DPDE* lo registra en la entrada de *encima*<sup>1</sup> como partícula semejante.

¿La vacilación puede deberse a que aún no está totalmente gramaticalizado y a que su comportamiento como MD todavía no es estable?, aunque llama la atención que las obras que lo registran como MD son las más actuales.

Se tome *por si fuera poco* como un sinónimo del definido o como una definición parafrástica —la ausencia de esta expresión como sublema en el *DBE* nos inclina hacia esta segunda opción—, las pistas que ofrece sobre el significado de *encima* son escasas.

#### 4.3. EL CASO DE *SIN EMBARGO*

*Sin embargo* es un conector perteneciente al grupo de los contraargumentativos e indica que el argumento que introduce lleva a una conclusión contraria a la que se podía esperar del argumento anterior, que puede estar explícito o ser una suposición contextual (Martín Zorraquino y Portolés, 1999, p. 4115; Portolés [1998] 2001). Señala Montolío que este conector presenta un argumento más fuerte que el que lo precede, es decir, con mayor fuerza argumentativa, por lo que “inclina el signo de la conclusión en su misma dirección argumentativa” (2001, p. 51).

Aparece definido en el *DBE* como sigue:

(14) **embargo** [...] // conj. adversativa. **sin embargo**. No obstante, a pesar de.

Nuevamente nos encontramos ante una definición sinonímica múltiple. El primer sinónimo que aparece como definidor no se sublematiza en el *DBE*. Por tanto, estamos ante una pista perdida, que nos dice poco o nada sobre el significado de *sin embargo*. La impropiedad en el uso del conector contraargumentativo *no obstante* como definición de *sin embargo* se manifiesta en varios aspectos que podemos sintetizar como sigue: *no obstante* no se sublematiza ni se define en el *DBE* (pista perdida); es un conector contraargumentativo que introduce un argumento fuerte, por lo tanto, es afín con *sin embargo*, pero su uso es menos frecuente y su significado no es totalmente idéntico (Martín Zorraquino y Portolés, 1999, p. 4116; Montolío, 2001, p. 63).

El otro sinónimo que se consigna en la definición es la locución preposicional *a pesar de* que aparece sublematizada en la obra<sup>12</sup>. Lo incorrecto de su uso radica, por una parte, en el empleo de una definición de tipo conceptual sinonímico para el conector *sin embargo* y, por otra, en que no constituye un sinónimo del conector aludido. Montolío clasifica esta locución preposicional como un conector contraargumentativo integrado entonativa y sintácticamente en la oración y, además, introductor de un argumento débil, es decir, presenta un impedimento que no será un obstáculo para lo se expresa en la oración principal (Montolío, 2001, pp. 47-62). De igual forma, la definición que se ofrece de *a pesar de* brinda pocas luces sobre el significado de *sin embargo*. Véase a continuación:

(15) **pesar** [...] loc. conj. **a pesar** o **a pesar de**. Contra todos los obstáculos, no obstante.

#### 4.4. EL CASO DE *CON TODO*

Este conector contraargumentativo es semejante a *sin embargo* y *no obstante* porque introduce un argumento que conduce a una conclusión contraria a la que se podía inferir del argumento anterior. La diferencia radica en que el miembro anterior se presenta como un argumento fuerte para arribar a una conclusión contraria a la que introduce el conector (Martín Zorraquino y Portolés, 1999, pp. 4116-4118; Montolío, 2001, pp. 70-74). Entre los refuerzos argumentativos que se emplean para hacer que el primer miembro discursivo sea fuerte se encuentran: “prolija acumulación de datos informativos; acopio de múltiples argumentos coorientados; uso de elementos léxicos, partículas y construcciones sintácticas con carácter intensificador, etc.” (Montolío, 2001, p. 74).

En el *DPDE* se señala su mayor frecuencia en el registro formal de la lengua. Este conector se define en el *DBE* de la manera siguiente:

(16) **todo**, **a** [...] // loc. conj. **con todo**. Sin embargo, no obstante.

En este caso el uso de la definición sinonímica nos conduce, por las dos vías de búsqueda que impone la remisión, hacia la pista perdida *no obstante*. También nos lleva

---

<sup>12</sup> En el *DBE* aparece con la marca gramatical correspondiente a la locución conjuntiva. Sobre la consideración de esta locución como preposicional o conjuntiva véase la *NGLE* (2009, pp. 2276-2284; 2460).

a *a pesar de* que, como vimos, aun cuando tiene un sentido contraargumentativo, se comporta sintácticamente diferente a estos conectores y desde el punto de vista semántico-pragmático introduce argumentos débiles.

Nos hallamos ante una definición que no clarifica el significado del definido; que acuña como sinónimos conectores que, si bien son semejantes, no son totalmente idénticos, y que envuelve al usuario en una búsqueda que lleva a pistas perdidas.

#### 4.5. EL CASO DE *AL CONTRARIO*, *POR EL CONTRARIO*

Este par de conectores ha sido tratado de diversas maneras en la bibliografía consultada. Algunos diccionarios los registran como sinónimos (*Diccionario de la lengua española, DLE*, 2014), otros, como variantes (*Diccionario del estudiante...* de la RAE). Sin embargo, Portolés ([1998] 2001) afirma que esto no se corresponde con la realidad. Propone para probarlo un ejercicio en el que se debe elegir entre los conectores *al contrario*, *por el contrario* y *en cambio* para rellenar varios espacios vacíos en un diálogo. La solución demuestra que estos tres marcadores no son intercambiables en varios contextos, por lo que no son sinónimos (Portolés [1998] 2001, pp. 83-85). Esta posición es la que asume el *DPDE* que tiene una entrada para cada uno de estos conectores.

El conector *por el contrario* introduce un miembro discursivo que es contrario al miembro que lo antecede. Tiene no solo el significado de contraste, sino también el de contrariedad (Martín Zorraquino y Portolés, 1999, pp. 4110-4111) o contraposición (Montolío, 2001, p. 85), lo que lo diferencia, por ejemplo, del conector *en cambio*. También introduce un miembro discursivo que puede refutar el miembro anterior (*DPDE*, 2008). Señala Montolío que en este último caso el conector introduce un argumento que niega la verdad que se introduce en el primer miembro, y esto sucede porque el conector vincula dos miembros que comentan el mismo tópico y el primero de ellos presenta polaridad negativa (Montolío, 2001, p. 87). También destaca el *DPDE* (2008) que se documenta su uso en posición final del miembro discursivo y aparece con mayor frecuencia en el registro formal de la lengua.

Por su parte, el conector *al contrario* introduce, igualmente, un miembro que contradice el miembro discursivo anterior y en muchas ocasiones lo refuta. En Martín

Zorraquino y Portolés se presentan algunas pruebas que demuestran su no gramaticalización como conector (1999, p. 4112). Sin embargo, se presentan otras en las que su comportamiento se acerca al de los MD y hace que se distinga de gran parte de los adverbios (1999, p. 4113). En esta investigación nos basamos en esos comportamientos que lo acercan a los conectores contraargumentativos y lo consideramos como tal.

Sus propiedades distribucionales, por ejemplo, que puede aparecer como elemento autónomo en un turno de palabra, lo diferencian de *por el contrario*. Es por ello que lo consideraremos como conector semejante a *por el contrario*, pero no una variante.

En el *DBE* se recogen estos conectores como variantes. Véase:

- (17) **contrario, a** adj. [...] // loc. adv. **al contrario, por el contrario, de lo contrario**. De modo distinto, al revés.

La definición que se ofrece es mixta, es decir, perifrástica<sup>13</sup> y sinonímica, y en cualquiera de los dos casos, conceptual. Responde al significado de *al contrario* cuando actúa como locución adverbial y no como conector.

Creemos que los residuos del significado conceptual de la unidad léxica *contrario, a* en *al contrario* y *por el contrario* hace que ambas se definan de forma conceptual, que no se distinga entre el significado de *al contrario* como locución adverbial y como conector discursivo, y que el significado de procesamiento de *por el contrario* no se refleje<sup>14</sup>.

#### 4.6. EL CASO DE *POR (LO) TANTO*

Este conector consecutivo se caracteriza por presentar el miembro que introduce como una consecuencia del miembro que lo antecede, derivada de una reflexión, de un razonamiento (Martín Zorraquino y Portolés, 1999, pp. 4100-4101). Se indica en el *DPDE* que la variante *por tanto* tiene mayor frecuencia de uso en el registro formal de la lengua.

---

<sup>13</sup> En este caso una *definición perifrástica relacional*, muy empleada para los adverbios. Para las características de este tipo de definición, véase Porto Dapena (2014, pp. 92-101).

<sup>14</sup> Un diccionario que recoge de cierta forma esta diferencia es el *Diccionario de uso del español* de María Moliner ([1966-1967] 2008). Véase la entrada **contrario, -a**.

Su definición en el diccionario escolar que nos ocupa es la siguiente:

(18) **tanto, a** [...] // loc. conj. **por tanto** o **por lo tanto**. Por lo que, en consecuencia.

La definición se aleja de ser funcional o explicativa. Nuevamente este diccionario recurre a una definición sinonímica para apuntar el significado de un conector.

En primer lugar, se define con una construcción (*por lo que*) que Montolío (2001) clasifica como conector integrado en la oración. Al decir de esta autora, esta expresión conectiva relaciona de manera estrecha la consecuencia que introduce con la causa que la antecede (2001, pp. 105-108). Aun cuando esta construcción conectiva enlace causa y consecuencia al igual que *por lo tanto*, sus características sintácticas, prosódicas, distribucionales e, incluso, semántico-pragmáticas difieren. Por tal motivo, no constituye un sinónimo de *por (lo) tanto* y su uso como definición no esclarece el significado de este último.

En segundo lugar, el conector *en consecuencia*, si bien pertenece al grupo de los consecutivos, se diferencia de *por (lo) tanto* en que introduce una consecuencia que no deriva de un razonamiento previo. Solo conduce de un estado de cosas, que constituye la causa, al efecto. No así *por (lo) tanto*, que el hecho de estar fundamentado en un razonamiento hace que se puedan invertir, generalmente, los miembros discursivos que vincula (Martín Zorraquino y Portolés, 1999, pp. 4104-4105; Montolío, 2001, pp. 125-129).

De otro lado, hemos de señalar que la búsqueda del conector *en consecuencia* en el *DBE* conduce a una pista perdida, pues este no se encuentra sublematizado.

#### 4.7. EL CASO DE *ENTONCES*

Nos hallamos ante un conector con un débil sentido consecutivo, según señalan Martín Zorraquino y Portolés (1999). Se emplea para dar continuidad al discurso a partir de la presentación de informaciones nuevas que aportan al desarrollo de un tópico general. Estos autores precisan que este sentido laxo de consecuencia se deriva de su originario sentido temporal (Martín Zorraquino y Portolés, 1999, pp. 4107-4109).

El *DBE* define este conector como sigue:

(19) **entonces** [...] 2. conj. En consecuencia, por lo tanto.

Se acude a la definición sinonímica múltiple que nos lleva, por una parte, a la pista perdida que es el conector *en consecuencia* y, por otra, al conector *por lo tanto*, que, como se analizó, conduce igualmente a *en consecuencia* (pista perdida) y a *por lo que* (falso sinónimo de *por (lo) tanto*).

Una vez más este diccionario hace que su usuario potencial se enfrente a una serie de remisiones que lo dejan en el total vacío.

#### 4.8. EL CASO DE *PUES*

El conector consecutivo *pues* se emplea para introducir un miembro discursivo que se presenta como consecuente del miembro discursivo anterior (Martín Zorraquino y Portolés, 1999, p. 4099). Una de las características fundamentales de este conector es que rechaza la posición inicial en el miembro del discurso que introduce y va seguido de pausa<sup>15</sup>. Montolío destaca que es un conector de uso poco frecuente, dado su aspecto formal (2001, p. 118).

Se recoge en el *DBE* como se muestra a continuación:

(20) **pues** [...] 2 conj. consecutiva. En consecuencia, por lo tanto.

Observamos que la definición es sinonímica y coincide con la presentada para el conector consecutivo *entonces*<sup>16</sup>. Por una parte, consideramos que estos conectores aun cuando tienen un valor consecutivo no son sinónimos. Por otra, nos llama la atención la falta de coherencia interna del diccionario analizado, pues establece relaciones sinonímicas que no consigna en cada artículo: no se ha empleado la marca Sin. (sinónimo) para relacionar los conectores *entonces* y *pues*. Esto nos mueve a pensar que no se ha sido consciente de que ambas unidades léxicas están definidas de igual manera.

## 6. CONCLUSIONES

El diccionario se erige como una herramienta fundamental en el proceso de enseñanza –aprendizaje de una lengua, ya sea materna o extranjera. A él acuden los docentes, los

---

<sup>15</sup> Cuando esto no ocurre se confunde con el MD estructurador de la información (comentador) *pues* (Martín Zorraquino y Portolés, 1999, p. 4083; 4100) o con la conjunción causal *pues* (Montolío, 2001, p. 117).

escolares y todo aquel que necesite obtener información de diversa naturaleza sobre la lengua (significado de las palabras, aspectos gramaticales, ortográficos, sinónimos y antónimos, etc.). Su compromiso social y el carácter de ‘portador de la verdad sobre el significado y uso de las palabras’ que se le atribuye conllevan que sea realizado con el mayor rigor científico.

El análisis realizado reveló algunas inconsistencias internas del *DBE* con respecto al tratamiento de las definiciones de los conectores discursivos, así como incongruencias con postulados teóricos actuales. A continuación resumiremos las más relevantes:

- No se emplea en ningún caso la definición funcional o explicativa, que es el desiderátum en cuanto a tipo de definición para las palabras gramaticales.
- En todos los casos se emplean definiciones de tipo sinonímico, aunque realmente lo que encontramos son falsos sinónimos (los definidores no son auténticos sinónimos de los definidos); por lo que las definiciones se pueden clasificar de *parasinonímicas*, siguiendo la terminología de Porto Dapena (2002, 2014).
- El empleo de las definiciones antes mencionadas condujo a problemas infraestructurales en el diccionario como la circularidad y las pistas perdidas.

Lo antes expuesto conduce a pensar en la necesidad de redefinir estas unidades lingüísticas en el *DBE*, a partir del empleo de definiciones sencillas, de tipo funcional, y despojadas de un metalenguaje complejo; en aras de que los escolares, destinatarios potenciales de esta obra lexicográfica, comprendan el significado y la función de este tipo de unidades léxicas, y puedan utilizarlas de manera correcta y efectiva. Cerramos con el siguiente ejemplo, a modo de propuesta:

(21) **contrario**, a. adj. [...] // **al contrario**. loc. adv. De modo distinto. EJEMPLO<sup>17</sup> / nexo con valor adversativo. Se utiliza para indicar contrariedad con la información anterior o refutarla. EJEMPLO. (Puede ir seguido o no de la información que indica cómo debe interpretarse este sentido de contrariedad en relación con lo dicho anteriormente. EJEMPLO. En el diálogo, puede aparecer de manera independiente. EJEMPLO). Af. por el contrario. // **por el contrario**. loc. adv. nexo con valor adversativo. Se utiliza para introducir una información contraria a la información anterior o refutarla. EJEMPLO. Ú. m. en el registro formal de la lengua. Af. al contrario.

---

<sup>17</sup> Se ha mantenido este significado para que quede más clara la posición que adoptamos con respecto a los cambios que creemos han de efectuarse en este artículo lexicográfico.

## BIBLIOGRAFÍA

- Azorín Fernández, D. (2000). Los diccionarios didácticos del español desde la perspectiva de sus destinatarios, *ELUA. Estudios de Lingüística Universidad de Alicante*, 14, pp. 19-44.
- Blakemore, D. (1987). *Semantic constraints on relevance*. Oxford: Blackwell.
- Briz Gómez, A. (1993). Los conectores pragmáticos en español coloquial (I): su papel argumentativo, *Contextos*, XI, 21/22, pp. 145-188.
- Briz, A., Pons, S. y J. Portolés (Coords.) (2008). *Diccionario de partículas discursivas del español* [en línea]  [<www.dpde.es>](http://www.dpde.es)
- Cuartero Sánchez, J. M. (1995). El estatuto categorial de *además* y sus propiedades distribucionales, *Dicenda*, 13, pp. 103-118.
- Forgas, E. y M. Bargalló (2005-2006). Los diccionarios del Centro de Lingüística Aplicada de Santiago de Cuba, *Revista de Lexicografía*, XII, pp. 249-266.
- Fuentes Rodríguez, C. (1996). *La sintaxis de los relacionantes supraoracionales*. Madrid: Arco Libros.
- Fuentes Rodríguez, C. (2003). Operador/conector, un criterio para la sintaxis discursiva, *RILCE*, 19.1, pp. 61-85.
- Fuentes Rodríguez, C. (2009). *Diccionario de conectores y operadores del español*. Madrid: Arco Libros.
- García Roche, A. de la C. (2018). Los marcadores conversacionales enfocadores de la alteridad en una muestra del español coloquial de La Habana, *Universidad de La Habana*, 286, pp. 72-81 [en línea] [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0253-92762018000200005&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0253-92762018000200005&lng=es&tlng=es).
- Gili y Gaya, S. (1966 [1943]). *Curso superior de sintaxis española*. La Habana: Edición Revolucionaria.
- González Mafud, A. M. y M. Perdomo Carmona (2014). Marcadores del discurso de La Habana, *Cuadernos de la ALFAL*, 5, pp. 107-139.
- Hernández Hernández, H. (1989). *Los diccionarios de orientación escolar. Contribución al estudio de la lexicografía monolingüe española*. Tubinga: Max Niemeyer Verlag.
- Hernández Hernández, H. (2003). La lexicografía escolar del español: desde sus orígenes hasta su consolidación, *Revista Internacional de Lingüística Iberoamericana*, I, pp. 7-25.
- Hernández Hernández, H. (2008). Retos de la lexicografía didáctica española. En Azorín Fernández D. et al. (Eds.), *El diccionario como puente entre las lenguas y las culturas del mundo. Actas del II Congreso Internacional de Lexicografía Hispánica*. Alicante: Fundación Biblioteca Virtual Miguel de Cervantes, pp. 22-32.
- Holgado Lage, A. (2014). Pragmática lingüística aplicada al español para extranjeros, con especial atención a los marcadores discursivos: El *Diccionario de Marcadores Discursivos para Estudiantes de Español como Lengua Extranjera*, Tesis doctoral dirigida por Elena Bajo Pérez, Facultad de Filología, Universidad de Salamanca [en línea] <http://hdl.handle.net/10366/132282>.

- Holgado Lage, A. (2018). *Diccionario de Marcadores Discursivos para estudiantes de español como segunda lengua*. Berlín / New York: Peter Lang.
- Llopis Cardona, A. (2006). Aplicación de la teoría de Rey-Debove a las definiciones lexicográficas de los marcadores discursivos. En M. Villayandre Llamazares (Ed.), *Actas del XXXV Simposio Internacional de la Sociedad Española de Lingüística*. León: Universidad de León, Departamento de Filología Hispánica y Clásica. pp. 1086-1105
- López Serena, A. y M. Borreguero Zuloaga (2010). Los marcadores del discurso y la variación lengua hablada vs. lengua escrita. En O. Loureda Lamas y E. Acín-Villa (Coords.), *Los estudios sobre marcadores del discurso en español, hoy*. Madrid: Arco Libros, pp. 325-405.
- Maldonado González, C. (Dir.) (2003). *Diccionario Secundaria y Bachillerato. Lengua Española*, (2.ª ed.). Madrid: Ediciones SM.
- Martín Zorraquino, M.ª A. (2004). El tratamiento lexicográfico de los marcadores del discurso y la enseñanza de E/LE. En M.ª A. Castillo Carballo (Coord.), *Actas del XV Congreso Internacional de ASELE*, pp. 53-67.  
[en línea] [http://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/15/15\\_0051.pdf](http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/15/15_0051.pdf)
- Martín Zorraquino, M.ª A. y J. Portolés Lázaro (1999). Los marcadores del discurso. En I. Bosque y V. Demonte (Dirs.), *Gramática descriptiva de la lengua española*. Madrid: Espasa Calpe, 3 Vols. pp. 4051-4213.
- Martínez Abreu, Y. (2015). Glosar el marcador discursivo *bueno* en el español coloquial de La Habana. En *IV Coloquio Internacional Marcadores del discurso en las lenguas románicas*. Universidad de Heidelberg (Alemania).
- Miyares Bermúdez, E. et al. (2006). *Léxico activo-funcional del escolar cubano*. Santiago de Cuba: Centro de Lingüística Aplicada.
- Miyares Bermúdez, E. (Dir.) (2014). *Diccionario básico escolar* (4.ª ed.). Santiago de Cuba: Editorial Oriente.
- Moliner, M. (1966-1967). *Diccionario de uso del español*, versión electrónica 3.0. Madrid: Gredos.
- Montolío Durán, E. (2001). *Conectores de la lengua escrita*. Barcelona: Ariel.
- Nomdedeu Rull, A. (2012). Miyares Bermúdez, Eloína (Dir.) (2009): *Diccionario básico escolar*. Santiago de Cuba, Centro de Lingüística Aplicada, *Revista de Lexicografía*, XVII, pp. 243-256.
- Perdomo Carmona, M. (2013). Estudio de los marcadores discursivos conversacionales en el habla culta de La Habana, en *Memorias. VIII Conferencia Internacional Lingüística 2013* (CD-ROM). La Habana: Instituto de Literatura y Lingüística.
- Pons Bordería, S. (1998). *Conexión y conectores, Estudio de su relación en el registro informal de la lengua*. Valencia: Universitat de València.
- Porto Dapena, J. Á. (2002). *Manual de técnica lexicográfica*. Madrid: Arco Libros.
- Porto Dapena, J. Á. (2014). *La definición lexicográfica*. Madrid: Arco Libros.
- Portolés Lázaro, J. (2001). *Marcadores del discurso* (2.ª ed.). Barcelona: Ariel.

- Real Academia Española y Asociación de Academias de la Lengua Española. (2009). *Nueva gramática de la lengua española*. Madrid: Espasa.
- Real Academia Española. (2011). *Diccionario del estudiante. Secundaria y Bachillerato* (2.ª ed.). Madrid: Santillana Ediciones Generales.
- Real Academia Española. (2014). *Diccionario de la lengua española* (23.ª ed.). Madrid: Espasa-Calpe.
- Ruiz Miyares, L. y S. Tarp (2015). Los diccionarios escolares cubanos: una experiencia para compartir. *Estudios de Lexicografía*, 4, pp. 185-198.
- Sánchez Muñoz, T. (2002-2003). Los diccionarios escolares a comienzos del siglo XXI, *Revista de Lexicografía*, IX, pp. 175-188.
- Sanmartín Sáez, J. (2010). La dificultad de definir en un diccionario de partículas. En M.ª A. Castillo Carballo y J. M. García Platero (Coord.), *La lexicografía en su dimensión teórica*. Málaga: Universidad de Málaga. pp. 561-585.
- Santos Río, L. (2003). *Diccionario de partículas*. Salamanca: Luso-Española.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Seco, M. (1987). Problemas formales de la definición. En *Estudios de lexicografía Española*. Madrid: Gredos. pp. 15-34.
- Vázquez Veiga, N. (1995-96). Los marcadores discursivos en las obras lexicográficas, *Revista de Lexicografía*, 2, pp. 133-149.
- Vázquez Veiga, N. (2000). Los marcadores discursivos en el *Diccionario de colocaciones y marcadores del español*. En U. Heid *et al.* (Eds.), *Proceedings of the Ninth euralex International Congress*. Stuttgart: Universität Stuttgart. pp. 761-771 [en línea] [http://www.euralex.org/elx\\_proceedings/Euralex2000/087\\_Nancy%20VAZQUEZ%20VEIGA\\_Los%20marcadores%20discursivos%20en%20el%20Diccionario%20de%20colocaciones%20y%20marcadores%20del%20espanol.pdf](http://www.euralex.org/elx_proceedings/Euralex2000/087_Nancy%20VAZQUEZ%20VEIGA_Los%20marcadores%20discursivos%20en%20el%20Diccionario%20de%20colocaciones%20y%20marcadores%20del%20espanol.pdf).
- Vázquez Veiga, N. (2008). Las etiquetas discursivas en un diccionario de marcadores pragmáticos: Un puente entre lenguas. En D. Azorín Fernández *et al.* (Coords.), *El diccionario como puente entre las lenguas y culturas del mundo: actas del II congreso Internacional de Lexicografía Hispánica* [en línea] [http://media.cervantesvirtual.com/s3/BVMC\\_OBRAS/020/6b1/768/2b2/11d/fac/c70/021/85c/e60/64/mimes/0206b176-82b2-11df-acc7-002185ce6064.pdf](http://media.cervantesvirtual.com/s3/BVMC_OBRAS/020/6b1/768/2b2/11d/fac/c70/021/85c/e60/64/mimes/0206b176-82b2-11df-acc7-002185ce6064.pdf).



## APROXIMACIÓN AL INDIGENISMO LÉXICO EN EL ESPAÑOL DE EL SALVADOR (SIGLOS XVII-XVIII)\*

### A LEXICON FROM SPANISH OF EL SALVADOR (17<sup>th</sup>-18<sup>th</sup> CENTURIES): AN APPROXIMATION TO INDIGENISMS

José Armando San Martín Gómez  
*Universidad Autónoma de Querétaro*  
[armandosmg21@gmail.com](mailto:armandosmg21@gmail.com)

#### RESUMEN

La historia del español salvadoreño ha sido muy poco estudiada en general (Ramírez Luengo, 2010), y entre los vacíos que existen está el tema de los indigenismos léxicos que se han incorporado a la lengua. Precisamente por eso, esta investigación pretende abordar la influencia léxica de las lenguas indígenas en el español de El Salvador de los siglos XVII y XVIII; con este propósito, se estudiarán 34 documentos de estos siglos escritos en la región (Ramírez Luengo, 2017), de los que se extraerán las unidades léxicas para analizar las siguientes cuestiones: a) lenguas de donde se toman; b) campos léxicos en que se utilizan; c) primeras dataciones de las mismas. El propósito final es contribuir a la historia léxica del español de Centroamérica y conocer mejor cómo las lenguas indígenas influyen en este.

**Palabras clave:** Historia del español de América, El Salvador, léxico, indigenismo, siglo XVIII

#### ABSTRACT

The history of the Salvadorean Spanish remains understudied until now (Ramírez Luengo, 2010), and the indigenisms are not an exception. This paper aims to study the lexical influence of the indigenous languages in the Spanish spoken in this country during 17<sup>th</sup>. And 18<sup>th</sup>. Centuries (Ramírez Luengo, 2017). For this purpose, a corpus of 34 historical documents will be studied in order to analyze the following issues: a) the etymological origin of indigenisms; b) lexical fields where they are used, c) their first appearance. The final goal is to increase our knowledge of the lexical history of Central American Spanish and also the influence of indigenous languages in it.

**Keywords:** History of Latin American Spanish, El Salvador, lexicon, indigenisms, 18th. Century.

\* Este trabajo constituye uno de los productos resultantes del proyecto de investigación Léxico histórico del español de El Salvador tardocolonial (1650-1803) (FFL-2018-02), financiado por la Universidad Autónoma de Querétaro por medio del Fondo para el Fortalecimiento de la Investigación UAQ-2018.

## 1. ESTADO DE LA CUESTIÓN: LA HISTORIA DEL LÉXICO SALVADOREÑO

Sin lugar a dudas, el español de Centroamérica puede considerarse como una de las áreas menos estudiadas por los historiadores de la lengua, tal y como señala Ramírez Luengo (2017, p. 4) cuando dice que “las variedades del español de los países que componen América central constituyen, dentro del ámbito de la dialectología hispánica, sin duda las más desconocidas, tanto desde el punto de vista sincrónico como muy especialmente diacrónico”: exceptuando el caso de Costa Rica, que cuenta ya con bastantes estudios (entre otros, Quesada Pacheco, 1995 y 2009), el resto de las regiones ha quedado bastante abandonado por parte de los investigadores, por lo que hace falta obtener más información diacrónica que permita conocer mejor su evolución; este aserto es especialmente evidente en el caso del español salvadoreño, para el que solo existen dos trabajos sobre el siglo XVIII (Ramírez Luengo, 2010 y en prensa), ninguno de los cuales aborda la cuestión del vocabulario en general, ni en particular de las voces indígenas que han dejado huella en su léxico, de manera que es necesario llevar a cabo trabajos que, basándose en documentación del pasado, aporten luz sobre esta cuestión.

## 2. ASPECTOS METODOLÓGICOS

Teniendo en cuenta lo descrito hasta ahora, el presente trabajo pretende aportar unos primeros datos sobre los indigenismos léxicos presentes en el español salvadoreño de los siglos XVII y XVIII. Específicamente, los objetivos que se persiguen en el estudio son los siguientes: 1) registrar y clasificar todos los vocablos de origen indígena presentes en el corpus; 2) analizar su origen etimológico y su distribución por campos léxicos; y 3) señalar la aparición de primeras dataciones, sea absoluta o centroamericana. Se pretende, por tanto, llevar a cabo un análisis desde diferentes puntos de vista que contribuya a un aporte para la lexicología histórica del español americano y sirva para ampliar el conocimiento existente hasta el momento sobre el uso de voces indígenas en los siglos XVII y XVIII en El Salvador.

Por lo que se refiere al corpus seleccionado para el estudio, se ha optado por utilizar 34 documentos redactados en esta región que aparecen publicados en Ramírez Luengo (2017). En concreto, se trata de documentación de carácter inquisitorial tomada del Archivo General de la Nación (AGN; Ciudad de México, México), que en el volumen señalado se ha transcrito siguiendo los criterios CHARTA (2015). En total, dicha documentación equivale a un total de 17837 registros repartidos cronológicamente entre 1650 y 1803<sup>1</sup>; por lo que se refiere al factor diatópico, los documentos proceden de todo el país, y más específicamente de San Salvador —actual capital de El Salvador—, las ciudades de San Miguel, San Francisco Gotera, Sonsonate y San Vicente y de los municipios de San Alejo, Cuscatancingo e Izalco. Respecto a su tipología textual, la naturaleza inquisitorial de estos documentos hace que la mayor parte de ellos se enmarquen dentro del polo de la inmediatez comunicativa (Oesterreicher, 2004, p. 752), y en concreto que se encuentren con mayor frecuencia algunos tipos diplomáticos concretos, tales como las denuncias, las autodenuncias y los testimonios, si bien se incorporan así mismo al corpus otros textos diferentes como certificados de bautizo o inventarios de bienes, por citar solo algunos de los que componen la antología<sup>2</sup>.

### 3. LOS INDIGENISMOS EN EL ESPAÑOL SALVADOREÑO TARDOCOLONIAL: ANÁLISIS DE LOS DATOS

Como es más que sabido, el contacto que los españoles tienen con los pueblos indígenas en América es relevante para la incorporación de unidades de origen indígena en su español, ya que, debido a la falta de nombres para nombrar los referentes nuevos que no pertenecen a su realidad, es necesario acudir a las lenguas habladas por estos pueblos para dotar de nombre a tales realidades; en este sentido, es importante recordar que, según Enguita Utrilla (2010, p. 199), “la contribución más importante y segura de las lenguas indígenas del Nuevo Mundo al español está en el vocabulario, ya que la

---

<sup>1</sup> Este corpus parece ser altamente relevante y suficientemente amplio si se considera que se está hablando de un territorio muy pequeño y dialectalmente homogéneo, al menos desde el punto de vista léxico (Ramírez Luengo, 2017, p. 16; Lipski, 1996, p. 279).

<sup>2</sup> Para una descripción más detallada de las características geográficas, cronológicas y tipológicas del corpus, véase Ramírez Luengo (2017, pp. 15-18).

incorporación de préstamos léxicos a las hablas hispanoamericanas tiene lugar incluso en las zonas en las que [...] deben descartarse otros fenómenos de transferencia en los niveles fonético-fonológico y gramatical”.

Por supuesto, en el caso concreto del corpus estudiado también se pone de manifiesto este contacto intercultural, en concreto en forma de voces de origen autóctono, indigenismos, que aparecen empleadas en los textos estudiados: así, es necesario señalar en primer lugar se registra un total de diez elementos de esta naturaleza, en concreto *bajareque* (doc. 27, 1769, Izalco, Sonsonate, El Salvador), *canoá* (doc. 27, 1769, Izalco, Sonsonate, El Salvador), *chiquigüite* (doc. 26, 1769, San Vicente, San Vicente, El Salvador), *chocolate* (doc. 5, 1650, San Miguel, San Miguel, El Salvador), *güipil* (doc. 2, 1650, San Miguel, San Miguel, El Salvador), *maíz* (doc. 32, 1800, Sonsonate, Sonsonate, El Salvador), *mangle* (doc. 32, 1800, Sonsonate, Sonsonate, El Salvador) *naborío* (doc. 4, 1650, San Miguel, San Miguel, El Salvador), *petaca* (doc. 27, 1769, Izalco, Sonsonate, El Salvador) y *zompopo* (doc. 17, 1739, San Miguel, San Miguel, El Salvador).

Por lo que se refiere a su empleo, la aplicación del programa ANTCONC<sup>3</sup> a la documentación analizada ofrece un total de 653 sustantivos (*types*) y 3276 apariciones de ellos (*tokens*)<sup>4</sup>; dado que los indigenismos son diez y aparecen en trece ocasiones, se puede concluir que su uso es muy bajo, al equivaler al 1,53% de estos elementos y apenas un 0,39% de sus ocurrencias. Así pues, quizá la primera de las conclusiones que se puede extraer de este análisis sea la notable escasez de estos elementos, en una situación muy semejante a la que se descubre en otras zonas de Centroamérica en este mismo momento, tales como Guatemala (Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017, p. 114) u Honduras (Ramírez Luengo, 2018, p. 26).

Dentro ya de los indigenismos, cabe señalar que entre estos elementos es detectable la aparición de lo que se pueden considerar *ocasionalismos*, esto es, “palabras que no

---

<sup>3</sup> Como es bien sabido, se trata de una herramienta de libre acceso que permite trabajar con corpus lingüísticos, tanto sincrónicos como diacrónicos; puede descargarse gratuitamente en la siguiente página de internet <<http://https://www.laurenceanthony.net/software/antconc/>>.

<sup>4</sup> Dado que los indigenismos son en su totalidad sustantivos, se tiene en cuenta únicamente la cantidad de estos para extraer la frecuencia de uso.

pertenecen al uso habitual de la lengua receptora, sino que se usan ocasionalmente en ella [...] con plena conciencia de su condición de extranjeras” (Álvarez de Miranda, 2009, p. 144): efectivamente, mientras que los indigenismos integrados —es decir, aquellos que “se sienten como formas léxicas ya absorbidas plenamente por la lengua española” (Enguita Utrilla, 2010, p. 207)— aparecen de forma plenamente incorporada a la narración y carecen de rasgos discursivos que permitan interpretarlos como ajenos a la lengua, los ocasionalismos demuestran una marca de corte lingüístico que adquiere la forma de una explicación, y que pretende, en cierto modo, facilitar la comprensión de un término que hasta el momento es desconocido para el hablante (Enguita Utrilla, 2010, pp. 207-208). Así pues, partiendo de la presencia o no de esta marca lingüística en el texto, es posible clasificar los indigenismos registrados en el corpus en dos grupos diferenciados, tal y como aparecen en la tabla 1:

INCORPORACIÓN	CASOS	VOCES
Voces integradas	9 (90%)	<i>bajareque, canoa, chocolate, güipil, maíz, mangle, naborío, petaca, zompopo</i>
Voces no integradas	1(10%)	<i>chiquigüite</i>
<b>TOTAL</b>	10 (100%)	

**Tabla 1.** Indigenismos integrados y no integrados en el corpus

A partir de estos datos, es evidente que la mayor parte de los indigenismos del corpus —un total de 9 de 10, el 90%— pertenecen al grupo de los integrados, mientras que los ocasionalismos se reducen a un único vocablo náhuatl, probablemente de uso más restringido. En este mismo sentido, resulta también interesante señalar el modo con el que el autor del documento incorpora el vocablo: en concreto, se trata de la estrategia conocida técnicamente por Enguita Utrilla (2010, p. 209) como *sinonimia*, que consiste en la conformación de un binomio léxico formado por el indigenismo y un vocablo patrimonial de significado semejante unidos por las conjunciones *o* e *y* (Buesa Oliver y Enguita Utrilla, 1992, p. 43), tal y como se descubre en el ejemplo siguiente:

- (1) Lo afirmó dicho Santiago Reyes diciendo que, aviéndoseles perdido una nabaja, hisieron la prueba de las tixeras, poniéndolas en un *chiquigüite o sesta* clabadas de una punta y poniéndoles el dedo arriba (doc. 26; 1769, San Vicente, El Salvador)

Más allá del listado de voces expuesto más arriba y de su carácter integrado y no integrado de los indigenismos, es también interesante analizar los orígenes etimológicos de estos, pues esta información ayuda a entender la relevancia de las diferentes lenguas amerindias en la configuración léxica del español salvadoreño de los siglos XVII y XVIII; en lo que concierne a esta cuestión, los principales trabajos lexicográficos que facilitan información etimológica —en concreto, el *DCECH* (1980-1991), así como Santamaría (1959), Morínigo (1998), el *DAMER* (2010) y el *DLE* (2014)— han permitido clasificar de acuerdo con sus orígenes lingüísticos los diez vocablos del corpus, según se descubre en la siguiente tabla (tabla 2)<sup>5</sup>:

LENGUA	CASOS	VOCES
Náhuatl	4 (40%)	<i>chocolate, chiquigüite, güipil, petaca</i>
Lenguas antillanas (arahuaco/taíno)	5 (50%)	<i>bajareque, canoa, maíz, mangle, naborío</i>
Maya	1 (10%)	<i>zompopo</i>
<b>TOTAL</b>	10 (100%)	

**Tabla 2.** Distribución de los indigenismos del corpus por origen etimológico

Como se puede apreciar, son cuatro las lenguas que proporcionan los indigenismos del corpus, si bien no todas ellas muestran la misma importancia desde el punto de vista cuantitativo: en efecto, mientras que cinco de los diez vocablos del corpus, el 50%, son de origen antillano y cuatro más —el 40%— tienen su origen en el náhuatl, el aporte del maya es muy minoritario, y supone tan solo una aparición testimonial, la voz *zompopo*. Por supuesto, estos datos no resultan en modo alguno sorprendentes, pues la situación que se acaba de describir concuerda a grandes trazos con lo que se ha dicho en otros trabajos sobre la influencia léxica de las distintas lenguas autóctonas sobre el

<sup>5</sup> Dado que el objetivo del estudio no es analizar la discusión lexicográfica acerca del origen de estos elementos léxicos, se toma exclusivamente la que aparece en las obras citadas.

español de América (Ramírez Luengo, 2007, pp. 76-79), muy especialmente en lo que se refiere a la extensión de voces antillanas por la totalidad de las variedades del Nuevo Mundo, la importancia del náhuatl en los territorios de la Nueva España y la escasez de elementos de origen maya incluso en aquellas zonas donde las lenguas de esta familia son habladas, cuestión que se había puesto de manifiesto ya para el español yucateco (Ramírez Quintana, 2008, p. 1482; Quirós García y Ramírez Luengo, 2015, p. 200) o guatemalteco (Polo Cano, 2005, p. 192; Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017, p. 5) de diversos momentos de la historia.

Por otro lado, junto al origen etimológico de los indigenismos léxicos resulta también interesante analizar el campo léxico al que estos se incorporan: dado que, en palabras de Enguita Utrilla (2010, pp. 203-204), estos elementos “reflejan aspectos diversos de la cultura material (plantas, animales, vestimenta, utensilios, configuración del terreno y habitáculos) [... pero] también se documentan, no obstante otros términos de significado antropológico”, es posible sostener que la observación de esta cuestión permite analizar las diferentes realidades en las que la incidencia de las lenguas autóctonas sobre el español salvadoreño es mayor. Con este propósito, se presenta una clasificación de los datos del corpus que sigue de cerca la propuesta al respecto por Polo Cano (2005, p. 189) y que ofrece los siguientes resultados (tabla 3):

CAMPO LÉXICO	CASOS	VOCES
Enseres/utensilios	3 (30%)	<i>chiquigüite, canoa<sup>6</sup>, petaca</i>
Fauna	1 (10%)	<i>zompopo,</i>
Flora	2 (20%)	<i>maíz, mangle</i>
Organización social	1 (10%)	<i>naborío</i>
Productos manufacturados	2 (20%)	<i>bajareque, chocolate,</i>
Vestimenta	1 (10%)	<i>güipil</i>
<b>TOTAL</b>	10 (100%)	

**Tabla 3.** Distribución de los indigenismos del corpus por campo semántico

Así pues, a partir de los datos expuestos en la tabla anterior es posible extraer una serie de conclusiones que ayudan a comprender mejor la influencia léxica que presentan las lenguas indígenas sobre el español salvadoreño de los siglos XVII y XVIII: en primer lugar, es relevante mencionar el hecho de que en la totalidad de los casos los campos léxicos a los que se incorporan los indigenismos se relacionan claramente con la realidad material americana —sea natural, como la fauna y la flora, o se refiera a la cultura de los pueblos autóctonos, como la vestimenta, los enseres y utensilios y los productos manufacturados—, pues es precisamente el desconocimiento de esta nueva realidad lo que promueve el apropiamiento y la incorporación de los términos originarios de las lenguas amerindias a las distintas variedades del español americano (Mejías, 1980, p. 19; Ramírez Luengo, 2007, pp. 79-80); por otro lado, es también destacable el hecho de que los escasos vocablos registrados en el corpus se distribuyan por numerosos campos léxicos, pues dicha circunstancia muestra la evidente trascendencia que el indigenismo adquiere en la conformación de las distintas

<sup>6</sup> Téngase en cuenta que en el ejemplo del corpus (“Item, un *canoas* que sirve de guardar maíz”; doc. 27, Izalco, Sonsonate, 1769) esta voz no corresponde al valor de ‘bote’, sino que remite al de “especie de artesana o cajón de forma oblonga que sirve para dar de comer a los animales y para otros usos” (DLE, 2014: s.v. *canoas*), motivo por el cual se incorpora al campo léxico de los enseres.

realidades que componen el español americano en general, y en concreto el hablado en El Salvador.

Por último, es importante estudiar también las primeras dataciones, tanto absolutas como centroamericanas<sup>7</sup> —entendidas como es decir, registradas bien en América en general o bien en los países que componen América Central—, que ofrecen estos vocablos, pues esta cuestión permite obtener una imagen más precisa del momento en el que se produce la incorporación del indigenismo en el español americano y especialmente salvadoreño. A partir, por tanto, de la búsqueda de los elementos descubiertos en los textos en los tres grandes repositorios documentales americanos (*CORDE*, *CORDIAM* y *LEXHISP*) y en Quesada Pacheco (1995) se han obtenido los datos siguientes (tabla 4):

VOCABLO	DAT. CORPUS	DAT. ABSOLUTA	DAT. CENTROAMERICANA
Bajareque	1769	1543, Panamá ( <i>LEXHISP</i> )	1543, Panamá ( <i>LEXHISP</i> )
Canoa	1769	1525, México ( <i>CORDIAM</i> )	1669, Costa Rica (Quesada Pacheco, 1995)
Chiquigüite	1769	1531, México ( <i>LEXHISP</i> )	1581, Guatemala ( <i>LEXHISP</i> )
Chocolate	1650	1575, México ( <i>LEXHISP</i> )	1582, Guatemala ( <i>LEXHISP</i> )
Güipil	1650	1543, México ( <i>LEXHISP</i> )	1581, Guatemala ( <i>LEXHISP</i> )
Maíz	1769	1513, Tierra Firme ( <i>LEXHISP</i> )	1526, Nicaragua ( <i>LEXHISP</i> )
Mangle	1800	1519, Puerto Rico ( <i>LEXHISP</i> )	1533, Panamá ( <i>LEXHISP</i> )
Naborío	1650	1514, Puerto Rico ( <i>LEXHISP</i> )	1527, Nicaragua ( <i>LEXHISP</i> )
Petaca	1793	1531, México ( <i>LEXHISP</i> )	1550, Panamá ( <i>LEXHISP</i> )
Zompopo	1739	1943, El Salvador ( <i>LEXHISP</i> )	1943, El Salvador ( <i>LEXHISP</i> )

**Tabla 4.** Primera datación de los indigenismos del corpus

<sup>7</sup> Como su propio nombre indica, se entiende por primera datación centroamericana aquella aparición que se registra por primera vez en la documentación de los países que componen esta región; para la primera datación absoluta se ha tenido en cuenta su aparición más antigua, independientemente de la zona geográfica en que se documenta.

Una vez más, el análisis ofrece unos datos que concuerdan con lo que en principio se podía esperar al respecto: en efecto, la revisión de los materiales documentales publicados permite constatar que la mayoría de estas voces presentes en el corpus salvadoreño son empleadas a partir del siglo XVI en diversas zonas americanas, muy especialmente México (*canoa, chiquigüite, chocolate, güipil, petaca*), pero también otras como Tierra Firme y Puerto Rico (*maíz, mangle, naborío*). Ahora bien, frente a esta primera situación, es especialmente destacable el caso de *zompopo*: según el *DLE* (2014: s.v. *zompopo*) se trata de un vocablo diatópicamente circunscrito al español centroamericano que, con el significado de ‘hormiga de color café o rojizo’, se emplea actualmente en El Salvador, Guatemala, Honduras y Nicaragua, y cuya primera datación hasta el momento se daba a mediados del siglo XX, en concreto en El Salvador en 1943 (*LEXHISP*); sin embargo, la documentación estudiada permite aportar más datos sobre la historia de este indigenismo, adelantando su incorporación a esta variedad del español más de dos siglos, en concreto hasta 1739, cuando se emplea en la documentación al parecer con el mismo significado que presenta en la actualidad. Véase ejemplo 2:

- (2) respondió ella que la enseñaron Petrona Calero y Manela, su hija, quienes avían aprendido de los indios de Yayantique, las cuales Petrona y Manuela viben en el mismo barrio, y que así ella como ellas savían otras vevidas para asonsar a los hombres, y le espresó una del gusano, del caracol del monte, con unos güesitos de sapo y unas cavesas de *sompapos* (doc. 17, 1739, San Miguel, San Miguel, El Salvador)

De este modo, si hasta el momento la historia de este mayismo era del todo desconocida, el ejemplo que se acaba de mencionar confirma su empleo en el español salvadoreño del siglo XVIII, cuestión que a su vez permite situar su incorporación a la variedad lingüística de la zona en algún momento anterior a esta fecha, en una muestra evidente de la antigüedad de este vocablo de empleo actual en el español de América Central (*DLE*, 2014: s.v. *zompopo*).

#### 4. UNAS PRIMERAS CONCLUSIONES

De este modo, el análisis desarrollado a lo largo de estas páginas permite extraer ya una serie de conclusiones que aportan datos de interés acerca del uso del indigenismo en

el español salvadoreño de los siglos XVII y XVIII y, en consecuencia, sobre la historia del léxico de esta variedad americana de la lengua.

En primer lugar, es importante mencionar que la cantidad de indigenismos que aparece en la documentación analizada es notablemente escasa, dado que se reduce a diez únicos elementos (*bajareque, canoa, chiquigüite, chocolate, güipil, maíz, mangle, naborío, petaca, zompopo*) y trece apariciones de los mismos, empleos que equivalen al 1.53% de los distintos sustantivos y apenas un 0.39% de estos en los documentos. A partir de aquí, parece posible sostener la idea de que, en contraste con su importancia cualitativa, desde el punto de vista cuantitativo el empleo de los préstamos indígenas es bastante escaso, rasgo que el español salvadoreño del siglo XVIII comparte con otras variedades centroamericanas de esta misma centuria como Honduras y Guatemala (Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017, p. 114) cuyas causas habrá que analizar con calma en posteriores trabajos.

Junto a lo anterior, los datos demuestran que la mayoría de los indigenismos —en concreto, nueve de los diez elementos, el 90% del total de estos vocablos— aparece integrado en el español salvadoreño tardocolonial, pues tan solo se registra un único elemento no integrado, el nahuatlismo *chiquigüite*, cuya naturaleza no integrada se descubre por el hecho de contar con una marca de corte lingüístico en forma de glosa explicativa (Enguita Utrilla, 2010, pp. 207-208), en concreto un doblete sinonímico, como se sabe una de las estrategias discursivas habituales para la introducción de los indigenismos no integrados en los textos americanos (Buesa Oliver y Enguita Utrilla, 1992, p. 43).

Por otro lado, resulta también interesante el análisis del origen etimológico de las voces indígenas, y a este respecto cabe señalar que la mayoría de ellas son de origen antillano y náhuatl —50% y 40% respectivamente—, mientras que los aportes del maya se reducen a un único elemento, *zompopo*, que equivale al 10% restante. Por supuesto, estos datos no resultan en modo alguno sorprendentes, dada la extensión de los préstamos tomados del arahuaco y del taíno por todo el continente americano (Ramírez Luengo, 2007, pp. 76-79) y la importante presencia de nahuatlismos en el español de los territorios que componen la Nueva España, de los que El Salvador

actual forma parte; finalmente, la escasez de voces mayas también entra dentro de lo esperable, pues la baja presencia de préstamos de esta lengua ya se ha señalado en la bibliografía que analiza el léxico histórico de diferentes zonas donde es empleada, tales como la Península de Yucatán (Ramírez Quintana, 2008, p. 1482; Quirós García y Ramírez Luengo, 2015, p. 200) y Guatemala (Ramírez Luengo, Aquino Melchor y Ramírez Vázquez, 2017, p. 5; Polo Cano, 2005, p. 192).

Por lo que se refiere a los campos léxicos en que aparecen los vocablos de origen indígena, estos se extienden, a pesar de su escasez, por un amplio número de ellos – tales como la vestimenta, la fauna, la flora, la organización social y muy especialmente los enseres y utensilios y los productos manufacturados, con cinco de los diez términos localizados– que se caracterizan por referirse en todos los casos a realidades materiales y por presentar una naturaleza muy variada; ambas ideas son de gran relevancia, pues ponen en evidencia que el indigenismo no se circunscribe a un único campo léxico o a una realidad concreta, sino que se extiende por muchas de ellas y, por tanto, adquiere una gran trascendencia en los procesos de americanización léxica y en la conformación de las distintas variedades que componen el español americano.

Finalmente respecto a las primeras dataciones absolutas y centroamericanas, se ha aportado nueva información sobre una de estas voces, *zompopo*, cuya primera datación en los grandes repositorios históricos se encontraba establecida a mediados del siglo XX; sin embargo, gracias al presente estudio es posible adelantar esta datación poco más de dos siglos, en concreto hasta 1739, lo que muestra su incorporación ya en el siglo XVIII y, por tanto, la profundidad histórica que presenta este término en el español de El Salvador.

En definitiva, se puede decir que el análisis de los indigenismos de la documentación salvadoreña tardocolonial estudiado concuerda en sus líneas principales con las ideas fundamentales que, por el momento, se han indicado en la bibliografía sobre la incorporación de elementos de este origen en la historia del léxico del español americano (Ramírez Luengo, 2007, pp. 76-79); ahora bien, más allá de estos datos concretos acerca de los indigenismos salvadoreños expuestos en estas páginas, el acercamiento que se ha llevado a cabo en estas páginas demuestra que queda aún

mucho por estudiar acerca de la historia del español salvadoreño, muy especialmente en lo que se refiere a su configuración léxica, pues un mayor acercamiento tanto a los indigenismos como a los otros elementos configuradores de su vocabulario sin duda ayudará a comprender mejor la historia de una variedad lingüística como es la de este país, el cual, como se ha dicho previamente, constituye sin duda una de las áreas más descuidada en la bibliografía del español del Nuevo Mundo.

## BIBLIOGRAFÍA

- Álvarez de Miranda, P. (2009). Neología y pérdida léxica. En E. De Miguel (Ed.), *Panorama de lexicología*. Barcelona: Ariel. pp. 133-158.
- Buesa Oliver, T. y J. M. Enguita Utrilla (1992). *Léxico del español de América. Su elemento patrimonial e indígena*. Madrid: Mapfre.
- CORDE. Real Academia Española (2019). *Banco de datos. Corpus diacrónico del español* [en línea] <<http://www.rae.es>> [28/04/2019].
- CORDLAM. Academia Mexicana de la Lengua / Academia Nacional de Letras del Uruguay (2019). *Corpus diatópico y diacrónico del español de América* [en línea] <<http://www.cordiam.org>> [22/04/2019].
- DAMER. Asociación de Academias de la Lengua Española (2010). *Diccionario de Americanismos*. Madrid: Santillana.
- Enguita Utrilla, J. M. (2010). Indoamericanismos léxicos y estructuras discursivas en la relación de Cristóbal de Molina. En P. Jiménez, P. Cuenca y E. López (Eds.), *Relación de las fábulas y ritos de los incas*. Frankfurt/Madrid: Vervuert/ Iberoamérica. pp. 199-216.
- LEXHISP. Boyd-Bowman, Peter. (2003). *Léxico hispanoamericano. 1493-1993*. Nueva York: Hispanic Seminary of Medieval Studies [en línea] <<http://textred.spanport.lss.wisc.edu>> [20/04/2019].
- Lipski, J. M. (1996). *El español de América*. Madrid: Cátedra.
- Mejías, H. A. (1980). *Préstamos de las lenguas indígenas en el español americano del siglo XVII*. México, DF: Universidad Nacional Autónoma de México.
- Morínigo, M. A. (1998). *Nuevo Diccionario de Americanismos e Indigenismos*. Buenos Aires: Claridad.
- Oesterreicher, W. (2004). Textos entre inmediatez y distancia comunicativa: el problema de lo hablado escrito en el Siglo de Oro. En Cano Aguilar, R. (Ed.): *Historia de la lengua española*. Barcelona: Ariel. pp. 729-770.
- Polo Cano, N. (2005). Algunos indigenismos léxicos en el español de Guatemala del siglo XVIII. En Ana García Lenza y Ana Rodríguez Barreiro (Coords.). *Res Diachronicae Digital 4 (El contacto lingüístico desde una perspectiva histórica)*, pp. 184-202.

- Quesada Pacheco, M. Á. (1995). *Diccionario histórico del español de Costa Rica*. San José: Editorial Universidad Estatal a Distancia.
- Quesada Pacheco, M. Á. (2009). *Historia de la lengua española en Costa Rica*. San José: Editorial de la Universidad de Costa Rica.
- Quirós García, M y J. L. Ramírez Luengo (2015). Observaciones sobre el léxico del español de Yucatán (1650-1800), *Revista de Filología Española*, 95/1, pp. 183-210.
- Ramírez Luengo, J. L. (2007). *Breve historia del español de América*. España: Arco libros.
- Ramírez Luengo, J. L. (2017). *Textos para la historia del español, XI. Honduras y El Salvador (1650-1819)*. Alcalá de Henares: Universidad de Alcalá.
- Ramírez Luengo, J. L. (en prensa). La configuración fónica del español salvadoreño en la época colonial (1650-1803). *Boletín de la Real Academia Española*.
- Ramírez Luengo, J. L., Aquino Melchor, S. A. y J. J. Ramírez Vázquez (2017). Indigenismos léxicos en el español guatemalteco del siglo XVIII, *Lengua y habla*, 21, pp. 111-125.
- Ramírez Quintana, P. Á. (2008). Configuradores léxicos del español colonial en Campeche: la lengua maya. En C. Company Company y J. G. Moreno de Alba (Eds.), *Actas del VIII Congreso Internacional de Historia de la Lengua Española*, II. Madrid: Arco Libros, pp. 1471-1484.
- Real Academia Española (2014). *Diccionario de la Lengua Española*. Madrid: Espasa-Calpe.
- Santamaría, F. J. (1959). *Diccionario de mejicanismos*. México DF: Porrúa.



SEMINARIO DE LEXICOGRAFÍA HISPÁNICA (HUM 922)