

## ¿De qué hablamos cuando divulgamos sobre lingüística? Análisis de un corpus de textos divulgativos y aplicaciones al estudio terminológico de la semántica léxica

What do we talk about when we communicate about linguistics as a science?  
Analysis of a corpus of science communication texts and applications to the  
terminological study of lexical semantics

MARTA SÁNCHEZ-SAUS LASERNA

Instituto Universitario de Investigación en Lingüística Aplicada, Universidad de Cádiz, España  
marta.sanchezsaus@uca.es  
<https://orcid.org/0000-0002-1123-2873>

VANESA ÁLVAREZ TORRES

Instituto Universitario de Investigación en Lingüística Aplicada, Universidad de Cádiz, España  
vanesa.alvarez@uca.es  
<https://orcid.org/0000-0003-0477-4065>

### Resumen

En este trabajo presentamos la confección y el análisis de un corpus de textos digitales conformado por columnas periodísticas de corte divulgativo, escritas por lingüistas, que tratan sobre la lingüística como ciencia y que han sido publicadas en diarios digitales o revistas digitales de naturaleza divulgativa. Lo hacemos con dos objetivos: por un lado, observar qué aspectos del

### Abstract

In this paper we describe the compilation and analysis of a corpus of digital texts conformed by science communication articles written by linguists that discuss linguistic topics and that have been published in digital newspapers or digital magazines. The aims are twofold: on the one hand, to observe what aspects of linguistic work and research linguists are conveying

---

**Para citar este artículo:** Sánchez-Saus Laserna, M. y Álvarez Torres, V. (2022). ¿De qué hablamos cuando divulgamos sobre lingüística? Análisis de un corpus de textos divulgativos y aplicaciones al estudio terminológico de la semántica léxica. *ELUA*, (38), 73-98. <https://doi.org/10.14198/ELUA.22384>

Recibido: 31/03/2022 Aceptado: 14/05/2022

© 2022 Marta Sánchez-Saus Laserna, Vanesa Álvarez Torres



Este trabajo está sujeto a una licencia de Reconocimiento 4.0 Internacional de Creative Commons (CC BY 4.0)

trabajo y la investigación lingüística estamos trasladando los propios lingüistas a la sociedad en los textos divulgativos que aparecen en la prensa *online* y, por otro, analizar en este corpus la presencia de términos de la semántica léxica, lo que nos permite establecer para ellos un estatus terminológico propio de la divulgación e incluir, en el trabajo terminográfico, ejemplos de uso de estos términos en contextos divulgativos. Esto lo hacemos dentro del marco que nos proporcionan los proyectos sobre terminología de la semántica y bases de datos terminológicas que se están desarrollando en el Instituto Universitario de Investigación en Lingüística Aplicada de la Universidad de Cádiz<sup>1</sup>.

**PALABRAS CLAVE:** lingüística digital, lingüística de corpus, textos divulgativos, terminología, semántica léxica

to society in the informative texts that appear in the online press, and, on the other hand, to analyze the presence of terms related to lexical semantics in this corpus, which allows us to establish for them a terminological status specific to science communication and to include, in the terminographic work, examples of the use of these terms in science communication contexts. We do this within the framework provided by the projects on semantic terminology and terminological databases that are being developed at the Instituto Universitario de Investigación en Lingüística Aplicada of the University of Cádiz.

**KEYWORDS:** digital linguistics, corpus linguistics, science communication texts, terminology, lexical semantics

## 1. INTRODUCCIÓN

### 1.1. Comunicación mediada por ordenador y lingüística de corpus

El mundo desarrollado actual no se puede comprender sin la comunicación digital. En este contexto, tanto general como también especialmente en el académico, científico, técnico y profesional, es inevitable que surjan nuevos géneros digitales (Belda-Medina y Casañ-Pitarch 2020), que constituyen una línea de investigación lingüística emergente y novedosa (Luzón y Pérez-Llantada 2019). Las nuevas tecnologías y, sobre todo, las nuevas formas de transmisión de la información de manera electrónica han supuesto la aparición de nuevos géneros de carácter digital y la adaptación al entorno digital de géneros ya existentes, como el artículo periodístico en la prensa digital, que es el que abordamos en este trabajo. Los estudios lingüísticos no han sido ajenos a esta irrupción de lo digital y empieza a ser frecuente encontrar investigaciones que se sitúan dentro del campo de la “lingüística digital”, que Crespo Miguel (2020: 151) define así:

Cuando hablamos de lingüística digital debemos entender un conjunto de técnicas y herramientas enfocadas a la manipulación de datos lingüísticos en formato digital, lo que incluye el almacenamiento, la representación, la manipulación y la difusión de este tipo de datos. También se ocupa de cómo extraer y representar la información lingüística en formato electrónico, así como de las mejores prácticas para trabajar con esos datos, utilizando todos los beneficios de las tecnologías actuales.

<sup>1</sup> Los dos proyectos en los que se encuadra esta investigación, cuyo IP, en ambos casos, es Miguel Casas Gómez, son los siguientes: 1) “Lingüística y nuevas tecnologías de la información: la creación de un repositorio electrónico de documentación lingüística” (FEDER-UCA18-107788) de la Convocatoria 2018 de Ayudas a proyectos de I+D+i en el marco del Programa Operativo FEDER Andalucía 2014-2020, y 2) “Lingüística y Humanidades Digitales: base de datos relacional de documentación lingüística” (PY18-FR-2511) de la convocatoria 2018 de Ayudas a proyectos I+D+i (Modalidad “Frontera Consolidado”) del Plan Andaluz de Investigación, Desarrollo e Innovación.

Más concretamente, el marco de estudio de los géneros digitales está íntimamente vinculado con el empleo de la web como corpus, que se ha convertido en un tema de investigación muy fructífero en español, con distintos objetivos y desde distintas disciplinas de estudio. De Benito Moreno (2022: 481), en una interesante revisión de este campo a la que remitimos, señala que “la popularización de la comunicación digital ha conllevado el surgimiento de una nueva rama de la lingüística: el análisis de la comunicación y el discurso mediados por ordenador” y que la lingüística de corpus ha sido el marco metodológico en el que se ha desarrollado, debido a las grandes cantidades de texto que pone a nuestra disposición internet.

Los estudios que analizan las características propias del lenguaje empleado en las comunicaciones a través de internet comenzaron con el nacimiento de la propia internet. En inglés encontramos publicaciones que hacen referencia a la *computer mediated communication* (CMC) desde la década de 1980. Baron (1984) ya se centra en los posibles efectos que el empleo de ordenadores pueda tener para el lenguaje:

But there is yet a third dimension to the study of computers and language. That is the growing use of computers as conduits of natural language. Instead of speaking face-to-face with another or using traditional means of writing, we express ourselves at the computer keyboard. (Baron 1984: 119)

El acercamiento a la CMO supone, como hemos dicho, aproximarse a la lingüística de corpus. Como indican diversos autores (Parodi 2008, Zulaica Hernández 2015), la lingüística de corpus es una metodología de estudio de la lengua y el lenguaje que nos permite trabajar directamente en contextos auténticos con grandes cantidades de texto, asegurando la fiabilidad de los análisis cuantitativos y cualitativos y la comparación posterior con otros trabajos. Este enfoque es, además, de gran utilidad para el análisis de producciones orales y escritas de los hablantes de una lengua, como muestra la variedad de campos de estudio del ámbito de la lingüística que aplican los avances en la lingüística de corpus a sus investigaciones. En este trabajo vamos a enfocarnos hacia dos aplicaciones de los corpus: el análisis del discurso y el trabajo terminográfico.

Para el análisis del discurso son numerosísimos los trabajos que emplean corpus, tanto generales como especializados. Destacamos, como referencias básicas para el trabajo en español, los estudios de Mancera Rueda y Pano Alamán (2013) y (2014) y Pano Alamán y Moya Muñoz (2015), así como los de Cantamutto y Vela Delfa (2016) y Vela Delfa y Cantamutto (2015).<sup>2</sup> En nuestro caso, la perspectiva que adoptamos para el análisis de los temas tratados en el corpus es un análisis de contenido que parte de la extracción de las palabras clave o de las combinaciones léxicas significativamente más frecuentes en el corpus de estudio que en un corpus de comparación, asumiendo que existe una relación entre la frecuencia de la palabra y su relevancia en el discurso y en la construcción de representaciones (Scott y Tribble 2006; Rayson 2008; Duque 2014). En cuanto al uso de corpus para el trabajo terminográfico, como el que llevamos a cabo en la segunda parte de nuestro análisis, remitimos especialmente al trabajo de Cabré y Da Cunha (2022) –aunque nos centraremos en esto más abajo–.

El corpus que hemos confeccionado para este trabajo está formado por textos divulgativos escritos por especialistas en lingüística. Es cierto que ya existen para el español

<sup>2</sup> Llevamos a cabo una revisión más exhaustiva del uso de internet como corpus y, especialmente, de los textos extraídos de redes sociales, en Sánchez-Saus Laserna (2022).

proyectos en desarrollo muy interesantes basados en corpus que han tomado los textos periodísticos que tratan sobre la lengua como objeto de estudio, como los proyectos METAPRES y METAPRES-COLING<sup>3</sup> cuya IP es en ambos la profesora de la Universidad de Alicante Carmen Marimón. Estos proyectos están centrados en el análisis de las columnas sobre la lengua como género metalingüístico desde 1940 hasta la actualidad, aunque se encuentra aún en desarrollo<sup>4</sup>. Nuestro corpus es más modesto en tamaño, pero también es de naturaleza distinta: abarca solo textos de los últimos diez años, publicados directamente en prensa digital y escritos por lingüistas con una intención divulgativa.

## 1.2. La divulgación de la ciencia y los textos divulgativos como textos especializados

La divulgación científica tiene un papel fundamental para la mejora del conocimiento que la sociedad debe tener del mundo que le rodea. Es de gran relevancia que se conozca de primera mano la labor que realizan los investigadores, precisamente para que la sociedad valore la investigación en su justa medida. Esto, que es válido para cualquier disciplina, es especialmente importante para las humanidades y, en particular, para la lingüística, cuya imagen de cara a la sociedad está inexplicablemente alejada de la idea de investigación y se desconoce la aplicabilidad de muchas de sus disciplinas.

Sin embargo, la divulgación no es algo novedoso, pues, según indican Llácer Llorca y Ballesteros (2012: 58), se inició en 1632 con la obra *Dialogo sopra i due massimi sistemi del mondo tolemaico e copernicano* de Galileo Galilei, que, para llegar a un público más amplio, la redactó en italiano, y no en latín. Por su parte, Galán Rodríguez (2003: 138) señala que “la divulgación científica es una actividad indisoluble del proceso secular moderno que, a partir del siglo XVII, configuró en Europa un nuevo orden económico, político, científico, social y cultural”.

Aunque la divulgación de la ciencia no supla la formación realizada en las aulas por los docentes, ni convierta en especialista a un público lego en la materia, permite que “se privilegie la posibilidad de que el ciudadano pueda ubicarse racionalmente frente a la importancia de la ciencia y la tecnología en su devenir cotidiano, y que le permita recorrer y aprovechar la porción de conocimientos que le sean necesarios, útiles, o al menos interesantes” (Olmedo Estrada 2011: 139).

De este modo, difusión de la ciencia a la sociedad y los textos divulgativos están estrechamente relacionados, tanto que, como indica Galán Rodríguez (2021: 665), “el grado de acceso de la sociedad al conocimiento de una ciencia o una tecnología determina la posibilidad de producir textos divulgativos”. Esta autora emplea, en lugar de *ciencia*, un término más abarcador, *tecnociencia*, que genera dos tipos de discursos: 1) “de especialidad científica” y 2) “de divulgación” (Galán Rodríguez 2021: 657-658) y designa:

- 1) la actividad intelectual que trata de representar y explicar de forma objetiva el mundo y sus fenómenos;

<sup>3</sup> <http://metapres.es/>.

<sup>4</sup> El volumen de Marimón Llorca (2020) recoge resultados de las distintas direcciones en las que el equipo del Proyecto METAPRES ha abordado el examen del columnismo lingüístico como tradición discursiva. Por su parte, el proyecto de la Universidad de Cádiz “LinPePrensa. Ideas lingüísticas y pedagógicas en la prensa española del siglo XIX” analiza también las ideas lingüísticas en la prensa española, aunque se centra en textos del XIX. La polémica lingüística como tipología discursiva se analiza, por ejemplo, en Gaviño Rodríguez (2021).

- 2) el entramado del descubrimiento científico y sus aplicaciones sociales.

En cuanto a la divulgación de la lingüística, además de lo señalado anteriormente para la ciencia en general, si bien todos los hablantes de una comunidad no poseen un conocimiento específico de este campo de estudio, sí que cuentan con la facultad del lenguaje y un código lingüístico compartido por toda una comunidad de hablantes con el que se comunican entre sí: más que suficiente, por tanto, para que sea adecuado y pertinente transmitir a la sociedad los avances generados por la ciencia del lenguaje. Al fin y al cabo, como afirma Escandell Vidal (2011: XV):

El lenguaje está presente en todas las facetas de nuestra existencia: cada individuo es capaz de comunicarse en su lengua en multitud de situaciones y sin esfuerzo aparente. Sin embargo, es muy poco lo que el usuario común sabe de manera positiva y sistemática acerca de ese poderoso instrumento y de su propia capacidad de utilizarlo.

A partir de este razonamiento, el interés por los temas lingüísticos no solo debería surgir en lingüistas y filólogos u otros especialistas de otros ámbitos de estudio interesados en el lenguaje, sino también en los hablantes profanos en la materia. Así, si la sociedad actual apuesta por una ciencia al alcance de todos, en la que las nuevas tecnologías y el uso de nuevos medios de comunicación prestan servicio a la difusión de la misma, esto se presenta como una oportunidad para promulgar la proyección y el alcance del estudio científico del lenguaje a toda la sociedad.

Esta labor de divulgar el conocimiento científico se lleva a cabo, sobre todo, en los medios de comunicación. En ellos, divulgadores expertos en la materia transmiten conocimientos científicos al público en general de una manera más cercana y atrayente –adaptando un texto científico a un texto sobre ciencia para un público no experto en ese campo–, y, de este modo, se convierten en “traductores y propagandistas” (Galán Rodríguez y Montero Melchor 2002: 53). Sin embargo, hay que diferenciar la traducción realizada en el proceso adaptativo de los textos divulgativos de la traducción propiamente dicha, entre diferentes códigos lingüísticos, de ahí que Galán Rodríguez y Montero Melchor (2002: 59) establezcan las siguientes diferencias:

- a) mientras que en la traducción no hay ninguna referencia al texto primario (de hecho, deben ser dos textos distintos), en la divulgación se hace mención explícita a las fuentes originales porque, evidentemente, el divulgador actúa como un transmisor de la palabra de otros;
- b) el traductor es solo un intermediario sin voz propia que debe reflejar fielmente lo que se dice, pero no interpretarlo; el divulgador también es un intermediario, pero como su función es explicar un texto y adaptarlo (reformularlo), puede opinar;
- c) así como en la traducción no hay restricciones que afecten a la tipología o contenido de los textos, la divulgación está condicionada por los intereses del público.

Podemos añadir que los divulgadores, además de traductores y propagandistas, también son mediadores, en la línea que señalan Casas Gómez y Fernández Smith (2021: 724) cuando hacen referencia al género discursivo periodístico en el que se tratan temas específicos, como puede ser la ciencia.

De este modo, el redactor de textos divulgativos adecua a un público lego un contenido científico destinado en un principio a los propios científicos. Si bien los textos divulgativos

suelen estar redactados por expertos o periodistas con conocimientos específicos en la materia, ambos “se ajustan a un esquema semejante, con lo que las diferencias son de índole léxica y, en menor grado, gramaticales” (Galán Rodríguez y Montero Melchor 2002: 61). El divulgador, por tanto, se encuentra con la labor de mediar entre esos diferentes niveles de especialización, esto es, transmitir un conocimiento que cuenta con un alto grado de especialización a una sociedad que, por lo general, no cuenta con una formación tan específica en el ámbito en cuestión. Para ello, como indica Galán Rodríguez (2003: 150-152), el divulgador reformula o redefine haciendo uso de una serie de mecanismos metalingüísticos, “tipografía especial (comillas, letra itálica, paréntesis), verbos del ámbito semántico «llamar» (*llamar, llamarse, designar, nombrar, decir*, etc.), verbos que indican equivalencia o identidad (*ser, se define como, se considera como, consiste en, se trata de*), definiciones analíticas (por descomposición) y definiciones por la función”. Asimismo, el divulgador debe “reelaborar el sentido del trabajo científico original para que el texto divulgativo mantenga la coherencia semántica”, que implican reformulaciones no solo en el nivel léxico, sino también en otros niveles de análisis lingüístico, como el sintáctico.

Por tanto, si bien este tipo de textos se enmarca dentro de la comunicación especializada, caracterizada por el uso de un lenguaje especializado, debemos resaltar que el divulgador sí que hace uso de este tipo de lenguaje, pero los destinatarios del texto divulgativo, normalmente, al no contar con conocimientos específicos de la materia, carecen del lenguaje específico de ese ámbito. Esto constituye una muestra de la relación existente entre las dos dicotomías que algunos autores (Cabré 2004: 20) suelen relacionar: *lenguaje especializado/lenguaje general* y *conocimiento especializado/conocimiento general*.

### 1.2.1. Nivel de especialización de los textos de especialidad

El lenguaje especializado puede manifestarse en diferentes grados de especialización y, precisamente, el discurso divulgativo constituye, según Guantiva Acosta, Cabré y Castellà Lidon (2008: 18), el tercer nivel de especialización:

- 1) el discurso especializado (con distintos grados de especialización), dirigido a los especialistas;
- 2) el discurso de tipo didáctico, orientado al personal en formación, y
- 3) el discurso de carácter divulgativo, enfocado al público general.

En relación con los textos especializados y los textos divulgativos, Cabré (2002) establece algunas diferencias básicas entre ellos, que engloban la función, la temática, los usuarios, la situación comunicativa y el discurso y que pueden allanar el camino de nuestro estudio para comenzar con una base teórica fundamentada. No obstante, aunque es necesario determinar grados en el nivel de especialización de un texto, se ha señalado (Cabré 2002; Rodríguez Tapia 2016) que entre los distintos grados, en cuyos extremos están el texto especializado y el texto divulgativo, existe un *continuum*, en cuyo punto medio encontramos el texto semiespecializado, aunque con fronteras ciertamente difusas.

Los textos especializados pueden presentar mayor o menor grado de especialización dependiendo del conocimiento que tenga el destinatario sobre la materia. El destinatario puede tratarse de otro experto, un estudiante de universidad o una persona carente de conoci-

tos específicos del campo en cuestión, de ahí que, como afirman Cabré y da Cunha (2022: 198), “esta diferenciación entre niveles de especialización de los textos es clave a la hora de definir los géneros textuales que contendrá el corpus”. Para estas autoras la producción científica dirigida a expertos en la materia, como los artículos científicos y las monografías científicas, serían géneros textuales de nivel de especialización alto; el material generado por docentes en el ámbito universitario (por ejemplo, apuntes de clase) se enmarcaría en el nivel intermedio; y, por último, los textos divulgativos constituirían el nivel de especialización bajo. En estos tres niveles de especialización nos encontramos con una *comunicación especializada*, esto es, “el proceso mediante el cual se imparte o comparte conocimiento especializado” (Guantiva Acosta, Cabré y Castellà Lidon 2008: 17), pero, evidentemente, como ya hemos comentado, no del mismo grado de especialización, puesto que los factores que conciernen al propio acto comunicativo varían, como pueden ser los destinatarios. Así pues, en el lenguaje especializado empleado en este tipo de comunicación, estudiado en el ámbito de la terminología, encontramos una *variación de especialidad*, que forma parte de una lingüística del hablar (Casas Gómez y Fernández Smith 2021: 715).

### 1.3. Terminología y términos propios de la divulgación

La terminología, aunque surgiera fuera de la lingüística, se ha convertido en los últimos años en una disciplina fundamental para las nuevas directrices y orientaciones de la lingüística moderna, especialmente de corte aplicado, así como para el estudio de los nuevos contenidos y perspectivas de la semántica léxica (Casas Gómez y Hummel 2017, Casas Gómez 2020a). Desde los proyectos de investigación desarrollados por el grupo de investigación “Semaínein” (HUM-147) en el Instituto Universitario de Investigación en Lingüística Aplicada de la Universidad de Cádiz se trabaja en la elaboración de un diccionario especializado de los usos terminológicos de la semántica léxica. Y es fundamental recalcar esto último: el objetivo no es llevar a cabo un glosario de tecnicismos o unidades terminológicas, sino de usos técnicos de la semántica léxica, como señala Casas Gómez (2020b: 3), quien insiste en:

the necessity and usefulness of this type of specialised dictionary, since it is not a lexicon of technicisms or terminological units, but, rather, of the technical uses of lexical semantics. For this reason, starting from explanatory sub-indices, terminological datasheets must be created, for each acception of the different terms.

Con la extracción de los términos propios de la semántica presentes en el corpus de textos divulgativos pretendemos, en este contexto de investigación, por un lado, contribuir a uno de los objetivos que se plantean Casas Gómez y Varo Varo (2022, en prensa) en relación con la fase actual de trabajo de los proyectos citados:

Rescatar lingüísticamente una disciplina como la semántica, que parecía estar olvidada en el panorama de nuestra ciencia y que, sin embargo, se está convirtiendo, cada vez más, por su carácter básico, naturaleza interdisciplinar y dimensión transversal, en una materia central en las investigaciones lingüísticas actuales.

Y, por otro, ayudar a completar la base de datos de usos terminológicos sobre la que está construido el diccionario, que es metalingüística y multilingüe y ha conllevado la digitaliza-

ción de numerosas fuentes doctrinales y lexicográficas correspondientes a la semántica léxica (Álvarez Torres y Gómez Díaz 2022). Como señalan Casas Gómez y Varo Varo (2022, en prensa) resulta fundamental, como reto futuro para alcanzar resultados más representativos, ampliar el número de contextos metalingüísticos.

El documento terminográfico básico de estos proyectos es la ficha terminológica, patentada como *DOCUTERM: Modelo de documentación de usos terminológicos* (Casas Gómez, 2020), que consta de 24 campos. En este trabajo nos proponemos ampliar las posibilidades de dos de ellos: el estatus del término y los posibles contextos metalingüísticos significativos. En lo referente al estatus de los términos, los modelos de fichas no presentan información demasiado abundante sobre ellos. Cabré y Da Cunha (2022: 201) señalan, en la descripción de los posibles campos de una ficha: “Estatus del término (normativo, normalizado, etc.)”; Casas Gómez (2020c), por su parte, aporta como posibilidades dentro de este campo “estandarizado” / “no estandarizado”, “documentado en diccionarios o textos” / “no documentado en diccionarios o textos”, “propuesto para el traductor” / “no propuesto para el traductor”. Todas estas posibilidades dependen de su presencia en el corpus que se emplee y, como hemos indicado, el corpus empleado por los proyectos del ILA consta de diccionarios terminológicos y documentos de investigación científica. La propuesta que traemos nosotras en este trabajo es la de considerar un estatus más básico que todos estos, el estatus “documentado en textos divulgativos”, que manifiesta un uso mucho más extendido.

En cuanto a los contextos metalingüísticos, como se puede observar en la revisión de los modelos de documentación terminográfica que lleva a cabo Casas Gómez (2006), el apartado de contexto aparece en los modelos de Auger y Rousseau (1977), Dubuc (1999) y Cabré (1993), así como en el de Casas (2020c). Se prevé, para el diccionario de terminología de la semántica léxica, que estos contextos sean definitorios, representativos y adecuados para determinar y completar los diferentes usos terminológicos de las entradas léxicas. Sobre esto mismo, Casas (2020b: 5) indica:

It is not a mere testimonial use in which its presence is documented in a text, but, rather, a definitory or metalinguistic context, providing information on the content or giving a linguistic explanation of its use as a terminological unit of the discipline.

Como desarrollamos en el apartado de análisis, en este trabajo aportamos, para cada uno de los términos de la semántica léxica extraídos del corpus de artículos divulgativos, tres contextos metalingüísticos significativos que sean definitorios (del tipo “este término significa X”), que se empleen para denominar ejemplos de fenómenos lingüísticos (del tipo “este fenómeno X se denomina ‘término’”) o bien que, en el contexto en que aparecen, se vinculen con otros términos o sean aclaradas cuestiones de su definición o de su uso que puedan resultar interesantes desde el punto de vista terminológico.

## 2. METODOLOGÍA

El corpus creado para este trabajo está compuesto por los textos de 226 columnas de opinión de corte divulgativo, sobre temas relacionados con la lingüística y escritas por lingüistas. Están publicadas en prensa digital o en publicaciones periódicas digitales especializadas en divulgación, de acceso libre (es decir, para cuyo acceso no es necesario estar registrado en

la web del medio que las publica) en los últimos diez años (desde 2012). No se han tomado propiamente noticias, sino columnas, es decir, escritos de opinión redactados por lingüistas especializados. Para este estudio, que en principio se limita a la terminología en español, se han tomado solo publicaciones de España y en lengua española.

El corpus empleado está formado por 229.927 palabras extraídas de las siguientes fuentes:

- columnas de Elena Álvarez Mellado en *elDiario.es* ([https://www.eldiario.es/autores/elena\\_alvarez\\_mellado/](https://www.eldiario.es/autores/elena_alvarez_mellado/), desde el 15 de enero de 2017 hasta el 7 de febrero de 2022, 53 columnas): 50.785 palabras;
- columnas de María del Carmen Horno Chéliz en *Letras Libres* (<https://letraslibres.com/author/mamen-horno/>, desde el 10 de abril de 2012 hasta el 8 de marzo de 2022, 34 columnas): 27.622 palabras;
- columnas de Lola Pons en el suplemento de ciencia de *El País*, *Verne* ([https://verne.elpais.com/autor/maria\\_dolores\\_pons\\_rodriguez/a/](https://verne.elpais.com/autor/maria_dolores_pons_rodriguez/a/), desde el 2 de julio de 2017 hasta el 20 de febrero de 2021, 71 columnas): 71.220 palabras<sup>5</sup>;
- artículos publicados en la revista *Archiletras* en su sección de firmas (<https://www.archiletras.com/firmas/>, 68 artículos, de diversos autores, desde el 17 de diciembre de 2018 hasta el 11 de febrero de 2022): 80.300 palabras.

Para la extracción de los textos se ha empleado la herramienta de Sketch Engine para la confección de corpus desde la web. En concreto, se le han proporcionado los enlaces de cada uno de los artículos de cada periódico o revista, cuyo contenido ha sido descargado y con él se ha creado el corpus de trabajo. Se han añadido como metadatos de cada artículo el nombre del autor, la fecha de publicación, el medio en que fue publicado y el enlace al artículo. De nuevo con Sketch Engine, se han extraído las palabras clave del texto, así como los denominados por esta plataforma *multi-word terms*. Estos son combinaciones léxicas frecuentes “which appear more frequently in the focus corpus than in the reference corpus and, additionally, match the typical format of terminology in the language. The format is defined in the term grammar”<sup>6</sup>. El formato para el español es nombre + adjetivo, nombre + *de* + nombre o adjetivo + nombre + *de* + nombre, es decir, las estructuras que generalmente presentan los términos complejos, tal y como señalan Estopà (1999) y Vivaldi (2001), entre otros.

Estas palabras clave y *multi-word terms* son elementos cuya frecuencia de aparición en el corpus de análisis es significativamente mayor que en el corpus de comparación, a partir de un *keyness score*. El cálculo de este *keyness score* lo lleva a cabo Sketch Engine de la siguiente manera (Kilgarriff *et al.* 2014):

Key words are words typical of a focus corpus (a corpus we are interested in) in contrast to a reference corpus (usually a general corpus in the same language as the focus corpus). The *keyness score* of a word is calculated according to the following formula:

$$\frac{fpm_{focus} + n}{fpm_{ref} + n}$$

5 Las columnas de Lola Pons han continuado publicándose en *El País* tras el cierre del suplemento *Verne* en 2021, pero sujetas a suscripción al diario, por lo que no nos ha sido posible incluirlas en el corpus.

6 <https://www.sketchengine.eu/guide/keywords-and-term-extraction/#toggle-id-2>.

where  $f_{m_{\text{focus}}}$  is the normalized (per million) frequency of the word in the focus corpus,  $f_{m_{\text{ref}}}$  is the normalized (per million) frequency of the word in the reference corpus,  $n$  is the simple Maths (smoothing) parameter ( $n = 1$  is the default value).

El corpus de comparación que empleamos en la plataforma es el *Spanish Web 2018* (esTenTen18), que cuenta con más de 20.000 millones de palabras, extraídas de la web a partir de textos tanto de español de España como americano.

Esta selección de elementos léxicos a partir de su frecuencia en el corpus nos da un listado de elementos significativamente frecuentes, que evaluamos de manera manual para poder analizar con ellos las temáticas más frecuentes en los textos divulgativos sobre lingüística. Eliminamos de la lista apellidos (como Chomsky, Moliner o Nebrija), títulos de obras (*Diccionario de Autoridades*), palabras relacionadas con el contexto inmediato en que se redactó el artículo (*pandemia, confinamiento*) y, sobre todo, los ejemplos de fenómenos de los que se habla en los artículos del corpus, que, por su excepcionalidad, su carácter neológico o su pertenencia al registro coloquial o al nivel de lengua vulgar no tienen presencia en el corpus de comparación y, por tanto, aparecen entre las primeras palabras clave del corpus (*almóndiga, malamente, kellys, güeno, verdul, portavoza, señor, piolines, habemos* y otros muchos).

Para el análisis terminológico que llevamos a cabo en este trabajo vamos a centrarnos en la lista de palabras clave (pues los términos propios de la semántica que aparecen en el corpus son simples en todos los casos). Emplearemos las combinaciones léxicas frecuentes (los *multi-word terms*) únicamente para la descripción temática del corpus de divulgación, que tratamos en el epígrafe 3.1. Para esta descripción temática no necesitamos discriminar entre términos propiamente dichos (*principio de economía lingüística*, por ejemplo) y simples combinaciones de término (sustantivo) + modificador (*hablante nativo*), pues no pretendemos hacer un análisis terminológico con ellas.

La plataforma Sketch Engine se ha empleado, también, para extraer las concordancias de los términos que nos interesan, es decir, los contextos en los que los términos se encuentran dentro del corpus, y seleccionar de entre ellos los que sean apropiados para el trabajo terminográfico. La herramienta de correspondencias de esta plataforma incluye una opción especialmente interesante cuando se quieren extraer contextos significativos de cara a una base de datos terminológica: el sistema de evaluación de oraciones llamado GDEX, “Good Dictionary Examples” (Kilgarriff *et al.* 2008)<sup>7</sup>:

It is a system for evaluation of sentences with respect to their suitability to serve as dictionary examples or good examples for teaching purposes. Sentences are evaluated with respect to their length, use of complicated vocabulary, presence of controversial topics (politics, religion...), sufficient context, references pointing outside of the sentence (e.g. pronouns), brand names and other criteria.

7 <https://www.sketchengine.eu/guide/gdex/>.

### 3. ANÁLISIS

#### 3.1. Análisis de contenido del corpus: ¿de qué se habla en los artículos de divulgación sobre lingüística?

En el primer apartado del análisis nos centramos en las palabras clave y las combinaciones léxicas frecuentes (*multi-word terms*) del corpus que nos permite extraer Sketch Engine. Estos, recordamos, son los elementos léxicos simples y complejos significativamente frecuentes en nuestro corpus en comparación con un corpus normal. Como se ha señalado más arriba, conocer sobre qué temas se habla cuando se hace divulgación de la lingüística tiene una gran importancia de cara a la visibilidad y la percepción que la sociedad tiene de la disciplina: ¿qué estamos transmitiendo a la sociedad sobre lo que se investiga en lingüística?, ¿qué temas aparecen más a menudo?

	Palabras clave	FAce	FRce	FRce	Score		Combinaciones léxicas frecuentes	FA	FRce	FRce	Score
1	<b>hablante</b>	476	1784,31	3,87	366,87	1	<b>lengua materna</b>	71	266,15	1,70	99,03
2	<b>anglicismo</b>	63	236,16	0,44	164,36	2	<b>género gramatical</b>	28	104,96	0,07	98,87
3	<b>léxico</b>	134	502,31	3,26	118,02	3	<b>lenguaje inclusivo</b>	22	82,47	0,08	77,39
4	<b>lingüista</b>	63	236,16	1,27	104,41	4	<b>lengua estándar</b>	19	71,22	0,05	68,50
5	<b>lingüístico</b>	409	1533,16	14,19	100,98	5	<b>cambio lingüístico</b>	19	71,22	0,07	67,41
6	<b>gramatical</b>	102	382,35	3,09	93,69	6	<b>norma culta</b>	17	63,73	0,03	62,79
7	<b>uve</b>	30	112,46	0,27	89,69	7	<b>fenómeno lingüístico</b>	17	63,73	0,07	60,69
8	<b>diminutivo</b>	39	146,19	0,66	88,91	8	<b>hablante nativo</b>	20	74,97	0,26	60,07
9	<b>hache</b>	33	123,70	0,41	88,30	9	<b>máxima de Grice</b>	12	44,98	0	45,88
10	<b>leísmo</b>	23	86,22	0,05	83,31	10	<b>uso lingüístico</b>	12	44,98	0,11	41,44
11	<b>monolingüe</b>	27	101,21	0,27	80,42	11	<b>sexismo lingüístico</b>	11	41,23	0,02	41,42
12	<b>RAE</b>	121	453,57	4,82	78,13	12	<b>lenguaje juvenil</b>	11	41,23	0,02	41,39
13	<b>participio</b>	29	108,71	0,41	77,97	13	<b>español europeo</b>	11	41,23	0,04	40,74
14	<b>hispanohablante</b>	46	172,43	1,24	77,46	14	<b>diccionario académico</b>	11	41,23	0,06	39,67
15	<b>espanghish</b>	21	78,72	0,04	77,02	15	<b>castellano antiguo</b>	11	41,23	0,09	38,83
16	<b>latín</b>	178	667,24	7,77	76,17	16	<b>variedad lingüística</b>	11	41,23	0,10	38,43
17	<b>presuposición</b>	24	89,97	0,24	73,57	17	<b>lengua hija</b>	10	37,49	0	38,38

	Palabras clave	FAce	FRce	FRcc	Score		Combinaciones léxicas frecuentes	FA	FRce	FRcc	Score
18	<b>multilingüismo</b>	22	82,47	0,15	72,40	18	<b>historia del español</b>	10	37,49	0,03	37,37
19	<b>diccionario</b>	268	1004,61	13,03	71,69	19	<b>lengua de signos</b>	16	59,98	0,64	37,27
20	<b>morfema</b>	25	93,71	0,33	71,09	20	<b>español americano</b>	10	37,49	0,07	36,09
21	<b>eñe</b>	22	82,47	0,19	70,35	21	<b>español de España</b>	10	37,49	0,09	35,21
22	<b>lengua</b>	1311	4914,35	71,75	67,57	22	<b>signo de puntuación</b>	13	48,73	0,44	34,51
23	<b>yeísmo</b>	18	67,47	0,03	66,41	23	<b>lengua natural</b>	10	37,49	0,12	34,35
24	<b>subjuntivo</b>	25	93,71	0,45	65,20	24	<b>economía del lenguaje</b>	9	33,74	0,03	33,82
25	<b>gramática</b>	91	341,12	4,25	65,14	25	<b>complemento directo</b>	10	37,49	0,16	33,15
26	<b>multilingüe</b>	29	108,71	0,73	63,49	26	<b>lengua del mundo</b>	9	33,74	0,11	31,39
27	<b>preposición</b>	33	123,70	0,97	63,40	27	<b>idioma nacional</b>	9	33,74	0,11	31,38
28	<b>vocativo</b>	18	67,47	0,11	61,72	28	<b>lengua romance</b>	11	41,23	0,35	31,33
29	<b>neologismo</b>	27	101,21	0,67	61,22	29	<b>lengua española</b>	38	142,44	3,58	31,31
30	<b>consonante</b>	37	138,70	1,29	61,01	30	<b>máxima cooperativa</b>	8	29,99	0	30,99
31	<b>verbo</b>	166	622,26	9,55	59,06	31	<b>economía lingüística</b>	8	29,99	0,01	30,59
32	<b>lingüística</b>	52	194,92	2,35	58,54	32	<b>lenguaje claro</b>	10	37,49	0,29	29,90
33	<b>sufijo</b>	33	123,70	1,14	58,39	33	<b>género neutro</b>	8	29,99	0,05	29,54
34	<b>pronombre</b>	36	134,95	1,33	58,33	34	<b>grupo de palabra</b>	8	29,99	0,10	28,20
35	<b>lexicográfico</b>	18	67,47	0,24	55,07	35	<b>signo de apertura</b>	7	26,24	0,02	26,59
36	<b>lexicógrafo</b>	16	59,98	0,12	54,57	36	<b>área de Broca</b>	7	26,24	0,04	26,09
37	<b>sintáctico</b>	37	138,70	1,60	53,79	37	<b>español peninsular</b>	7	26,24	0,07	25,42
38	<b>blending</b>	14	52,48	0,00	53,41	38	<b>unidad léxica</b>	7	26,24	0,11	24,57
39	<b>pronunciación</b>	48	179,93	2,43	52,71	39	<b>sistema lingüístico</b>	7	26,24	0,13	24,09
40	<b>filólogo</b>	27	101,21	1,01	50,97	40	<b>sexo débil</b>	7	26,24	0,13	24,05
41	<b>tilde</b>	32	119,95	1,38	50,88	41	<b>desdoblamiento de género</b>	6	22,49	0	23,48
42	<b>adjetivo</b>	97	363,61	6,25	50,30	42	<b>principio de economía lingüística</b>	6	22,49	0	23,47
43	<b>dialecto</b>	48	179,93	2,78	47,90	43	<b>letra de mujer</b>	6	22,49	0,01	23,37

	Palabras clave	FAce	FRce	FRcc	Score		Combinaciones léxicas frecuentes	FA	FRce	FRcc	Score
44	<b>etimológico</b>	23	86,22	0,86	46,82	44	<b>español caribeño</b>	6	22,49	0,01	23,33
45	<b>ortografía</b>	64	239,91	4,47	44,00	45	<b>enunciado performativo</b>	6	22,49	0,01	23,29
46	<b>seseo</b>	12	44,98	0,08	42,54	46	<b>lengua actual</b>	6	22,49	0,02	23,08
47	<b>prefijo</b>	32	119,95	1,84	42,53	47	<b>diversidad lingüística</b>	7	26,24	0,20	22,67
48	<b>ideófono</b>	11	41,23	0,00	42,18	48	<b>niño bilingüe</b>	6	22,49	0,04	22,62
49	<b>performativo</b>	16	59,98	0,49	41,03	49	<b>lengua vehicular</b>	7	26,24	0,21	22,44
50	<b>aranés</b>	12	44,98	0,16	39,56	50	<b>función sintáctica</b>	6	22,49	0,07	22,03

Tabla 1. Palabras clave y combinaciones léxicas frecuentes en el corpus de estudio (FAce: frecuencia absoluta en el corpus de estudio, FRce: frecuencia relativa por millón de palabras en el corpus de estudio, FRcc: frecuencia relativa por millón de palabras en el corpus de comparación).

Más allá de términos muy generales, empleados en cualquier tema que esté relacionado con la lingüística (*hablante, lingüista, lingüístico, lengua, fenómeno lingüístico*, etc.), con los niveles del lenguaje (*sintáctico*) o con las categorías gramaticales (*verbo, preposición, adjetivo*, etc.), observamos que los temas preponderantes son varios. En primer lugar, es muy frecuente que se traten cuestiones relacionadas con los diccionarios y la labor lexicográfica (*diccionario, lexicógrafo, lexicográfico*), y, por extensión, con el papel de la RAE, así como la norma culta y la lengua estándar. Este papel de la RAE y del diccionario académico está también vinculado a los neologismos, de cuya inclusión en el diccionario o existencia fuera de él se habla a menudo (especialmente cuando la RAE publica la lista de palabras y acepciones nuevas incluidas en el *DLE*). Es recurrente en el corpus, en este sentido, la idea de que las palabras existen más allá de su registro en los diccionarios, que los diccionarios no incluyen solo palabras de la norma culta y que los usos de las palabras pueden ir más allá de la información que registre de ellas el diccionario, particularmente el *DLE*:

- (1) Los diccionarios no incluyen solo las formas que consideramos prestigiosas sino también otras que están marcadas por ser restringidas, por ejemplo, porque están en desuso, se limitan a una zona hispanohablante o tienen un empleo restringido a determinados contextos muy formales o muy informales. (“*Toda la verdad sobre ‘almóndiga’*”, Lola Pons, *Verne*, 10/03/2018, [https://verne.elpais.com/verne/2018/03/06/articulo/1520325761\\_808313.html](https://verne.elpais.com/verne/2018/03/06/articulo/1520325761_808313.html)).
- (2) Muchos han impugnado la palabra *desescalada* porque no está en el *DLE* (ahora presente en la última actualización), olvidando que nuestro diccionario es un puzzle que, además de registrar palabras concretas, incluye también las piezas necesarias para construir otras nuevas. (“*Desescalando la nueva normalidad*”. Isaías Lafuente, *Archiletras*, 02/12/2020, <https://www.archiletras.com/firma/desescalando-nueva-normalidad/>).
- (3) Nada más lejos de la realidad: lo que dota de significado a las palabras no es su etimología, ni lo que significaron en el pasado, ni siquiera su definición en un diccionario; lo que dota

de significado a una palabra es el uso que los hablantes hagan de ella. (“La falacia etimológica”, Elena Álvarez Mellado, *Archiletras*, 04/10/2019, <https://www.archiletras.com/firma/la-falacia-etimologica/>).

Por tanto, si lo lexicográfico y lo académico tienen esa presencia en el corpus, es comprensible que sean muy frecuentes términos como *lengua estándar*, *norma culta*, así como *neologismo* y, en relación con esto, *anglicismo* o *blending* como procedimientos para crear palabras nuevas. En relación, asimismo, con la norma culta y las cuestiones normativas otro tema recurrente es la ortografía, sobre la que encontramos términos como *uve*, *hache*, *ortografía*, *signo de puntuación*.

El enorme debate que suscita el lenguaje inclusivo está, como era de esperar, en el corpus: *género gramatical*, *lenguaje inclusivo*, *sexismo lingüístico*, *género neutro*, *desdoblamiento de género* muestran la relevancia de este tema en los artículos de divulgación. Las posturas son diversas y, en algunos casos, muy críticas:

- (1) Finalmente, desde el punto de vista sociolingüístico y estilístico, teniendo en cuenta que la mayoría de los registros de estas voces han partido de chicas jóvenes adictas a los clubs de fans, al ser un morfema no marcado en cuanto a género gramatical ofrece una ventaja a las hablantes sensibilizadas –y obsesionadas– con la cuestión del sexismo y del género inclusivo por mor de una mal entendida corrección política. (“Un nuevo sufijo en español: -er”, Félix Rodríguez, *Archiletras*, 23/03/2021, <https://www.archiletras.com/firma/un-nuevo-sufijo-en-espanol-er/>).
- (2) Pues bien, si los fenómenos sexistas discursivos reflejan estas formas de discriminación más allá del género gramatical, las recomendaciones no pueden ser ni la duplicación sistemática ni la feminización de cualquier apelativo, ni tampoco el uso de términos neutros. (“Una mirada otra al sexismo lingüístico”, Aguas-Vivas Catalá González y Enriqueta García Pascual, *Archiletras*, 05/03/2019, <https://www.archiletras.com/firma/sexismo-linguistico/>).

En otros, en cambio, se justifica el uso de, por ejemplo, el morfema “-e”:

- (1) La propuesta de construir un género neutro en -e soluciona muchos de los escollos que las anteriores propuestas dejaban sin resolver: fácil de pronunciar, morfológicamente claro, lingüísticamente económico, socialmente inclusivo. (“Todas, tod@s, todxs, todes: historia de la disidencia gramatical”, Elena Álvarez Mellado, *elDiario.es*, 27/06/2017, [https://www.eldiario.es/opinion/zona-critica/todas-todes-historia-disidencia-gramatical\\_129\\_3307468.html](https://www.eldiario.es/opinion/zona-critica/todas-todes-historia-disidencia-gramatical_129_3307468.html)).
- (2) Lo interesante del “todes” radica en que no solo se está proponiendo un morfema novedoso, sino que introduce un nuevo valor en oposición al masculino y femenino. Es decir, la morfología del castellano distingue masculino y femenino, y esta distinción es encarnada por diversos pares de morfemas (-o/-a, or/riz, -és/esa, etc.). La propuesta del género en -e no solo propone un nuevo morfema, sino que introduce todo un valor morfológico nuevo que sí tendría un valor neutro en oposición al masculino y femenino (“Lenguaje inclusivo: algunas claves lingüísticas”, Elena Álvarez Mellado, *elDiario.es*, 30/07/2021, [https://www.eldiario.es/opinion/zona-critica/lenguaje-inclusivo-claves-linguisticas\\_129\\_8021559.html](https://www.eldiario.es/opinion/zona-critica/lenguaje-inclusivo-claves-linguisticas_129_8021559.html)).

Las cuestiones que rodean a la lengua española son, pues, preponderantes. Se habla a menudo sobre sus variedades (*español americano*, *español de España*, *español europeo*,

*español peninsular, español caribeño*). Y no solo en su sincronía, sino también en su diacronía, lo que hace que términos como *castellano antiguo, historia del español* y *latín* se encuentren en la lista.

La diversidad lingüística, por tanto, tiene poca presencia. Como decimos, mayoritariamente se habla sobre el español, sus variedades, sus fenómenos, sus diccionarios, su academia y su historia. Solo en casos puntuales se adopta una perspectiva que recoja otras lenguas, normalmente asociada al multilingüismo y la convivencia de lenguas, especialmente en España, y en alguna ocasión puntual a cuestiones tipológicas y, también, a la importancia del reconocimiento de las lenguas maternas, especialmente cuando son minoritarias (y de ahí que aparezcan en el listado elementos como *monolingüe, multilingüe, lengua del mundo, diversidad lingüística* o *lengua materna*):

- (1) Cada cierto tiempo suelen aparecer noticias en diversos medios de comunicación y voces en redes sociales que se dedican a exponer los peligros y las desventajas de la diversidad lingüística, es decir, del hecho de que en el mundo coexistan lenguas diferentes, en la mayoría de los casos, en un mismo espacio geográfico (“Las bondades de las lenguas”, Iraide Ibarretxe-Antuñano, *Archiletras*, 20/01/2022, <https://www.archiletras.com/firma/las-bondades-de-las-lenguas/>).
- (2) El respeto a la diversidad cultural y lingüística es, sin duda, un patrimonio cultural que hay que cuidar y proteger; sin embargo, si bien es cierto que, desde una dimensión sociopolítica, institucionalmente se reconoce el multiculturalismo y el multilingüismo, no siempre se concibe la convivencia de lenguas y de culturas como un valor positivo. (“Los mixtecos, más allá de ‘Roma’”, Azucena Palacios, *Archiletras*, 10/07/2019, <https://www.archiletras.com/firma/los-mixtecos-mas-alla-de-roma/>).
- (3) Fantaseo con cómo de diferente sería este país (y cómo mejoraría nuestra capacidad para entender idiomas) si el multilingüismo impregnase los distintos ámbitos de la vida cotidiana, también en las regiones supuestamente monolingües: que la tele, la ficción audiovisual, los letreros en las calles o los discursos de las Cortes reflejasen la diversidad de lenguas que se hablan en el territorio. (“La fantasía de la España monolingüe”, Elena Álvarez Mellado, *elDiario.es*, [https://www.eldiario.es/opinion/zona-critica/fantasia-espana-monolingue\\_129\\_8530274.html](https://www.eldiario.es/opinion/zona-critica/fantasia-espana-monolingue_129_8530274.html)).
- (4) A través de la diversidad de las lenguas podemos intentar diferenciar qué aspectos de las lenguas son principios universales del lenguaje comunes a todas los idiomas y qué valores son variables y están sujetos a cambios. (“¿Para qué sirven las lenguas?”, Elena Álvarez Mellado, *elDiario.es*, [https://www.eldiario.es/opinion/zona-critica/sirven-lenguas\\_129\\_7283273.html](https://www.eldiario.es/opinion/zona-critica/sirven-lenguas_129_7283273.html)).
- (5) El tiempo vuela, sinceramente, y este mes celebramos, de nuevo, el día de la Lengua Materna. Hace un año os intenté convencer aquí de la naturaleza íntima y trascendente de esta lengua, que nos hace ser como somos, pensar como pensamos y sentir como sentimos. Entonces os animé a defender el derecho de los hablantes nativos a usar su lengua en los momentos más complicados de la vida. (“El derecho a tener lengua materna”, Mamen Horno, *Letras Libres*, 08/02/2021, <https://letraslibres.com/ciencia-y-tecnologia/el-derecho-a-tener-lengua-materna/>).

La referencia a la lengua materna también se hace en artículos de corte cognitivista que abordan su adquisición, el aprendizaje de las lenguas extranjeras y, también, las diferencias

en los procesos psicolingüísticos cuando se hablan unas u otras:

- (1) Otro de los beneficios con los que se relaciona el uso del área de Broca del hemisferio izquierdo por parte de los bilingües equilibrados tiene que ver con el cambio de código. Tal y como veíamos antes, el cambio de una lengua a otra supone un coste atencional que, en general, resulta mayor cuando pasamos de la L2 a la lengua materna que viceversa (“Bilingüismo, cognición y ciencia”, Mamen Horno, *Letras Libres*, 17/05/2016, <https://letraslibres.com/revista-espana/bilinguismo-cognicion-y-ciencia/>).

Por otro lado, términos como *presuposición*, *enunciado performativo*, *máxima de Grice*, *máxima cooperativa* muestran que las cuestiones pragmáticas tienen, asimismo, relevancia:

- (1) De hecho, quizá donde mejor se vea operar las máximas de Grice es cuando no se respetan, porque su incumplimiento tiende a producir comunicaciones insatisfactorias (o, como se dice en términos pragmáticos, infelices): la violación de las máximas de Grice explica la risa que nos produce cuando en un paquete de nueces se anuncia solemnemente que “contiene nueces” o la frustración que nos causa cuando alguien intenta hacerse el gracioso y responde con un escueto “sí” a la pregunta “¿Tienes hora?” (“Un cerdo a la izquierda”, Elena Álvarez Mellado, *elDiario.es*, 23/08/2020, [https://www.eldiario.es/opinion/zona-critica/cerdo-izquierda\\_129\\_6178302.html](https://www.eldiario.es/opinion/zona-critica/cerdo-izquierda_129_6178302.html)).
- (2) La elección de este titular (cuyo contenido es en apariencia obvio, pero periodísticamente atractivo) es en realidad un buen ejemplo de lo que en Lingüística se conoce como la teoría de la relevancia, una teoría formulada por los lingüistas Sperber y Wilson en 1986 y que supone una vuelta de tuerca sobre las máximas de Grice (de las que ya hablamos aquí en su día). (“El arte de escribir titulares relevantes”, Elena Álvarez Mellado, *elDiario.es*, 18/09/2021, [https://www.eldiario.es/opinion/zona-critica/arte-escribir-titulares-relevantes\\_129\\_8307960.html](https://www.eldiario.es/opinion/zona-critica/arte-escribir-titulares-relevantes_129_8307960.html)).

Así pues, la divulgación de la lingüística, en nuestro país, en español y en la última década, está especialmente enfocada a hablar sobre las cuestiones de las que más se habla también por parte de los no especialistas y que son, generalmente, los temas que el destinatario no especializado suele asociar con la lengua: lo académico (la RAE, el *DLE* –que son, sin duda, el referente y el contacto, respectivamente, más cercanos para el lector no especializado con el lenguaje y su lengua–) y los diccionarios en general, así como las nuevas palabras que se incluyen en ellos. Se habla sobre esto con una perspectiva descriptivista y con la intención de desterrar mitos, específicamente. El otro tema principal es el sexismo lingüístico y las estrategias morfológicas de creación de un género neutro, en el que sí hay posturas enfrentadas entre los autores de las columnas: desde quienes lo rechazan frontalmente hasta quienes acogen las propuestas con, incluso, entusiasmo, aunque también existen posturas intermedias. Otros temas están ya más vinculados al autor que escribe la columna, pues, como es de esperar, cada autor suele escribir sobre los temas de los que es especialista. En las columnas de Lola Pons es habitual encontrar artículos sobre historia del español o sobre variedades del español; en la de Mamen Horno, sobre cuestiones cognitivas y adquisición aprendizaje de lenguas y, en el caso de Elena Álvarez Mellado, aunque sus temas son más variados, son frecuentes las cuestiones sobre pragmática.

### 3.2. Corpus de textos digitales y terminografía. Establecimiento de un estatus terminológico propio de la divulgación y de contextos divulgativos

El análisis de la temática del corpus confeccionado pone de manifiesto que las cuestiones relacionadas con la semántica léxica (salvo en lo que se refiere a la lexicografía y, en menor medida, cuando se tratan cuestiones de semántica cognitiva) no tienen una presencia muy habitual en la divulgación de la lingüística. No obstante, se detecta un número suficiente de términos que nos permite abrir la puerta al establecimiento, en las fichas terminológicas, de un estatus terminológico propio de la divulgación para algunos términos tan habituales que aparecen con suficiente frecuencia incluso en un corpus no centrado en cuestiones semánticas.

Los términos propios de la semántica léxica presentes en el corpus y que hemos podido extraer con el análisis de palabras clave son los siguientes:

Orden	Término	FAce	FRce	FRcc	Score
18	<b>diccionario</b>	268	1004,61	13,03	71,69
28	<b>neologismo</b>	27	101,21	0,67	61,22
34	<b>lexicográfico</b>	18	67,47	0,24	55,07
35	<b>lexicógrafo</b>	16	59,98	0,12	54,57
51	<b>prototípico</b>	13	48,73	0,32	37,64
61	<b>lexicón</b>	9	33,74	0,09	31,82
67	<b>eufemismo</b>	21	78,72	1,61	30,57
69	<b>metafórico</b>	20	74,97	1,61	29,06
71	<b>polisemia</b>	9	33,74	0,22	28,41
73	<b>despectivo</b>	20	74,97	1,70	28,16
92	<b>significado</b>	226	847,17	33,49	24,59
97	<b>acepción</b>	28	104,96	3,56	23,24
180	<b>lexicografía</b>	5	18,74	0,19	16,65
315	<b>metonimia</b>	4	14,99	0,27	12,62

Tabla 2. Términos propios de la semántica léxica presentes en el corpus con una frecuencia significativa (FAce: frecuencia absoluta en el corpus de estudio, FRce: frecuencia relativa por millón de palabras en el corpus de estudio, FRcc: frecuencia relativa por millón de palabras en el corpus de comparación).

Aportamos, para cada uno de ellos, tres contextos metalingüísticos lo más definitorios, representativos y adecuados posible, de entre los disponibles en el corpus, para determinar los usos terminológicos de estos términos propios de la divulgación. Es cierto que no siempre se pueden encontrar contextos que definan el término, pero sí aportan, en todos los casos, información relevante del uso de estos términos y, también, del vínculo que presentan con otros términos de la lingüística o bien muestran información aclaratoria sobre el concepto manejado que no siempre encontramos en textos escritos por especialistas para especialistas.

- diccionario:

- (1) Los diccionarios no incluyen solo las formas que consideramos prestigiosas sino también otras que están marcadas por ser restringidas, por ejemplo, porque están en desuso, se limitan a una zona hispanohablante o tienen un empleo restringido a determinados contextos muy formales o muy informales. (“Toda la verdad sobre ‘almóndiga””, Lola Pons, *Verne*, 10/03/2018, [https://verne.elpais.com/verne/2018/03/06/articulo/1520325761\\_808313.html](https://verne.elpais.com/verne/2018/03/06/articulo/1520325761_808313.html)).
- (2) La lengua se crea porque los hablantes la usan y la función de los diccionarios y gramáticas es (o debería ser) recoger fielmente cómo nos expresamos los hablantes, no decirnos cómo debemos hablar. (“Iros’: el bueno, el feo y el malo”, Elena Álvarez Mellado, *elDiario.es*, 18/07/2017, [https://www.eldiario.es/opinion/zona-critica/iros-bueno-feo-malo\\_129\\_3270024.html](https://www.eldiario.es/opinion/zona-critica/iros-bueno-feo-malo_129_3270024.html)).
- (3) Es exactamente lo mismo que debe hacer el diccionario: no eliminar palabras, sino recogerlas, definir las e indicar siempre la connotación y el uso que tienen. (“Dejad de pedirle a la RAE que elimine palabras”, Elena Álvarez Mellado, *elDiario.es*, 20/05/2017, [https://www.eldiario.es/opinion/zona-critica/dejad-pedirle-rae-elimine-palabras\\_129\\_3381428.html](https://www.eldiario.es/opinion/zona-critica/dejad-pedirle-rae-elimine-palabras_129_3381428.html)).

- metáfora / metafórico:

- (1) Son las metáforas: el mecanismo lingüístico por el que nos referimos a algo utilizando palabras que, en principio, usamos para denominar a otro objeto simplemente porque creemos ver entre ellos una cierta semejanza o paralelismo. (“Metáforas peligrosas: el cáncer como lucha”, Elena Álvarez Mellado, *elDiario.es*, 23/10/2017, [https://www.eldiario.es/opinion/zona-critica/metaporas-peligrosas-cancer-lucha\\_129\\_3108779.html](https://www.eldiario.es/opinion/zona-critica/metaporas-peligrosas-cancer-lucha_129_3108779.html)).
- (2) Generalmente, lo que sucede cuando se emplea una metáfora es que los hablantes ponemos en relación dos conceptos que aparentemente no tienen nada que ver entre sí, pero que sentimos que comparten algunas características, a veces físicas (la estructura familiar basada en las raíces y las ramas de un árbol), a veces funcionales (la medición de la potencia de un coche según el número de caballos). (“El poder de la metáfora en tiempos de pandemia”, Iraide Ibarretxe-Antuñano, *Archiletras*, 04/11/2020, <https://www.archiletras.com/firma/el-poder-de-la-metapora-en-tiempos-de-pandemia/>).
- (3) Lo interesante reside en que, en este tipo de juegos metafóricos, la creatividad funciona en el sentido inverso al que suele hacerlo en la lengua: en lengua, el escenario habitual suele consistir en una situación o realidad novedosa a la que le asignamos una imagen que la representa y que utilizamos para nombrar: los grupos de convivencia aislados frente a la COVID-19 pasan a ser burbujas, porque están cerrados como las burbujas. (“600 formas de mirar una pandemia”, Elena Álvarez Mellado, *elDiario.es*, 08/04/2021, [https://www.eldiario.es/opinion/zona-critica/600-formas-mirar-pandemia\\_129\\_7378281.html](https://www.eldiario.es/opinion/zona-critica/600-formas-mirar-pandemia_129_7378281.html)).

- prototípico:

- (1) Lo que se propone desde la psicolingüística es que cuando los hablantes manejamos categorías y conceptos tenemos inconscientemente en la cabeza uno o varios ejemplares concretos que consideramos prototípicos, es decir, ejemplos que los hablantes consideran buenos representantes de esa categoría y que son por lo general compartidos por los hablantes. (“Hambur-

guesa de espinacas, leche de soja”, Elena Álvarez Mellado, *elDiario.es*, 07/11/2020, [https://www.eldiario.es/opinion/zona-critica/hamburguesa-espinacas-leche-soja\\_129\\_6394604.html](https://www.eldiario.es/opinion/zona-critica/hamburguesa-espinacas-leche-soja_129_6394604.html)).

- (2) Así, por ejemplo, si hablamos de aves, un ruiseñor o un cóndor probablemente serán ejemplos más prototípicos en la cabeza de los hablantes que un pingüino. (“Hamburguesa de espinacas, leche de soja”, Elena Álvarez Mellado, *elDiario.es*, 07/11/2020, [https://www.eldiario.es/opinion/zona-critica/hamburguesa-espinacas-leche-soja\\_129\\_6394604.html](https://www.eldiario.es/opinion/zona-critica/hamburguesa-espinacas-leche-soja_129_6394604.html)).
- (3) Es decir, no se trata de que un objeto cumpla con una serie de requisitos o definiciones abstractas para definirlo como sopa, sino que es la comparación respecto al ejemplo prototípico de la clase (en este caso, la sopa de fideos) lo que establece su sopidad. (“Hamburguesa de espinacas, leche de soja”, Elena Álvarez Mellado, *elDiario.es*, 07/11/2020, [https://www.eldiario.es/opinion/zona-critica/hamburguesa-espinacas-leche-soja\\_129\\_6394604.html](https://www.eldiario.es/opinion/zona-critica/hamburguesa-espinacas-leche-soja_129_6394604.html)).

- neologismo:

- (1) Cualquier innovación que se registra en una lengua pasa por dos fases: una es la propia innovación y otra, la verdaderamente clave, es la difusión. Innovar lingüísticamente no es tan complejo, pero esa creación queda en nada si no es adoptada por otros hablantes y se generaliza más allá de la persona que inventa. Esa nueva palabra que se crea y, si tiene suerte, se extiende, recibe el nombre de neologismo. (“‘Covidiota’, ‘balconazis’, ‘cuarenpena’... los neologismos que nos ha traído la pandemia”, Lola Pons, *Verne*, 08/04/2020, [https://verne.elpais.com/verne/2020/04/07/articulo/1586246728\\_179666.html](https://verne.elpais.com/verne/2020/04/07/articulo/1586246728_179666.html)).
- (2) O en uno de mis neologismos favoritos: el brillibrilli o el efecto del uso (o abuso, según se mire) de purpurina o cualquier elemento que brille y cause destellos. Neologismo, por cierto, muy convencionalizado ya en diversas redes sociales y que sigue el patrón reduplicativo e ideofónico tan característico de otras lenguas con grandes dosis de este tipo de iconicidad como el japonés, donde se dice kira-kira, o el euskera, donde se dice dizdiz o briz-briz. (“Ser o no ser en lingüística”, Iraide Ibarretxe-Antuñano, *Archiletras*, 14/06/2021, <https://www.archiletras.com/firma/ser-o-no-ser-en-linguistica/>).
- (3) Todos estos neologismos se pueden usar en la conversación, los hablantes saben crearlos sin necesidad de que nadie les enseñe y sus interlocutores los entienden. Está claro que no todos tendrán una larga vida, no muchos se convertirán en palabras de uso generalizado y solo unos pocos entrarán en el diccionario, pero todos ellos, con más o menos éxito, son palabras de nuestro idioma. (“Cread, cread, benditos. Desterrando el mito del diccionario”, Mamen Horno, *Letras Libres*, 07/12/2020, <https://letraslibres.com/ciencia-y-tecnologia/cread-cread-benditos-desterrando-el-mito-del-diccionario/>).

- lexicón:

- (1) Así, por ejemplo, se ha comprobado en laboratorio que los hablantes multilingües activan simultáneamente todos sus lexicones con independencia de la lengua en la que estén hablando, incluso en entornos marcadamente monolingües. (“Bilingüismo, cognición y ciencia”, Mamen Horno, *Letras Libres*, 17/05/2016, <https://letraslibres.com/revista-espana/bilinguismo-cognicion-y-ciencia/>).
- (2) Este efecto de interferencia apunta, como vemos, a que en esta prueba de reconocimiento de palabras no solo está activado el lexicón de la lengua que se está utilizando, sino que se

activan todos los lexicones al mismo tiempo. (“Bilingüismo, cognición y ciencia”, Mamen Horno, *Letras Libres*, 17/05/2016, <https://letraslibres.com/revista-espana/bilinguismo-cognicion-y-ciencia/>).

- (3) Esta activación de todos los lexicones a un mismo tiempo en este tipo de pruebas se da tanto cuando la prueba se realiza en la L2 (y se activa a la vez la lengua materna) como cuando la prueba se realiza en la L1 (y se activan las lenguas segundas). (“Bilingüismo, cognición y ciencia”, Mamen Horno, *Letras Libres*, 17/05/2016, <https://letraslibres.com/revista-espana/bilinguismo-cognicion-y-ciencia/>).

- eufemismo:

- (1) Si el suavizar es la función del eufemismo y lo políticamente correcto, el disfemismo juega con la intensificación que provoca el término interdicto. [...] de igual manera, se conservan las ideas que guardan las palabras aunque usemos otras voces sustitutas que las encubran... (“Eufemismo de antes y tabúes de ahora: dos maneras de decir lo mismo”, Mercedes de la Torre García, *Archiletras*, 25/10/2021, <https://www.archiletras.com/firma/eufemismo-de-antes-y-tabues-de-ahora-dos-maneras-de-decir-lo-mismo/>).

- (2) El contexto y el paso del tiempo obligan a una continua sustitución, por ejemplo, los místicos usaron en sus obras el giro “retrete de amor”, donde retrete tiene el significado de ‘cuarto pequeño en la casa o habitación destinada para retirarse’(a orar), y a la postre se resemantiza para referirse al ‘lugar donde vamos a hacer aguas menores y aguas mayores’(otro eufemismo) que sustituye, a su vez, a letrina, y así sucesivamente (urinario, aseos, servicios, W.C., etc.). (“Eufemismo de antes y tabúes de ahora: dos maneras de decir lo mismo”, Mercedes de la Torre García, *Archiletras*, 25/10/2021, <https://www.archiletras.com/firma/eufemismo-de-antes-y-tabues-de-ahora-dos-maneras-de-decir-lo-mismo/>).

- (3) Pero *kelly* ha acabado funcionando también como eufemismo.[...] Así, la palabra sustituye no solo a «camarera», sino a otras denominaciones tradicionales y más claras, como «limpiadora», «señora de la limpieza» o «fregona». [...]. Es curioso, porque «camarera», en su sentido actual, ya es un eufemismo (y muy enfático), sacado de la terminología cortesana, como sucede también en España con la palabra «azafata»: en ambos casos se alude a oficios desemeñados antiguamente por grandes damas que tenían el privilegio de servir a la reina, en un caso organizando el servicio femenino en el palacio y, en el segundo, custodiando su joyero. (“Limpian, fijan y dan esplendor: las *kellys*”, Óscar Esquivias, *Archiletras*, 25/09/2020, <https://www.archiletras.com/firma/limpian-fijan-y-dan-esplendor-las-kellys/>).

- lexicografía / lexicográfico / lexicógrafo

- (1) Pero el caso es que la labor lexicográfica (esto es, el noble arte de hacer diccionarios) no funciona así. Las palabras no pertenecen a la RAE ni a los diccionarios, pertenecen a los hablantes. (“El mito de las palabras que no están en la RAE”, Elena Álvarez Mellado, *elDiario.es*, 13/03/2017, [https://www.eldiario.es/opinion/zona-critica/mito-palabras-rae-diccionario-drae\\_129\\_3529181.html](https://www.eldiario.es/opinion/zona-critica/mito-palabras-rae-diccionario-drae_129_3529181.html)).

- (2) Esta respuesta tan poco entusiasta le ha granjeado a la institución fama de antipática y carca, pero es, en realidad, bastante prudente y está plenamente justificada desde el punto de vista técnico: las definiciones deben confeccionarse atendiendo al uso (es decir, a datos empíricos), no a intuiciones o preferencias personales, así que conviene ser cautelosos para garantizar el rigor del trabajo lexicográfico. (“La RAE y el lujo de definir”, Elena Álvarez Mellado, *elDiario.es*, 12/03/2018, [https://www.eldiario.es/opinion/zona-critica/rae-lujo-definir\\_129\\_2225819.html](https://www.eldiario.es/opinion/zona-critica/rae-lujo-definir_129_2225819.html)).

- (3) La misión de los lexicógrafos es decirnos qué palabras se usan, con qué sentido, en qué contextos y con qué consecuencias; nos alertan sobre si una palabra se considera actualmente vulgar o si se usa poco, o la usa solo un determinado sector de la población. (“Cread, cread, benditos. Desterrando el mito del diccionario”, Mamen Horno, *Letras Libres*, 07/12/2020, <https://letraslibres.com/ciencia-y-tecnologia/cread-cread-benditos-desterrando-el-mito-del-diccionario/>).

- polisemia:

- (1) El motivo por el que estos principios son fundamentales en la conversación reside en que, cuando nos expresamos, casi nada de lo que decimos es unívoco o plenamente explícito, sino que está cuajado de ambigüedad, de polisemia o de información implícita que se presupone. (“Las trabajadoras y la cooperación”, Elena Álvarez Mellado, *elDiario.es*, 25/06/2018, [https://www.eldiario.es/opinion/zona-critica/trabajadoras-cooperacion\\_129\\_2055070.html](https://www.eldiario.es/opinion/zona-critica/trabajadoras-cooperacion_129_2055070.html)).
- (2) La polisemia y la ambigüedad son el pan nuestro de cada día en lengua. De hecho, la mayoría de palabras de un idioma acaparan varios significados. Y tiene sentido: al fin y al cabo, una lengua que asignase un único significado a cada palabra acumularía un repertorio de vocabulario inabarcable. Habitualmente, la polisemia no genera demasiados conflictos a los hablantes porque sabemos discernir sin mucho esfuerzo el significado con el que nuestro interlocutor está usando una palabra polisémica, ya sea a través de las palabras que la acompañan o de la situación que nos rodea. (“De qué hablamos cuando hablamos de consentir”, Elena Álvarez Mellado, *elDiario.es*, 20/11/2017, [https://www.eldiario.es/opinion/zona-critica/hablamos-consentir\\_129\\_3049368.html](https://www.eldiario.es/opinion/zona-critica/hablamos-consentir_129_3049368.html)).
- (3) Estas posiciones (que recuerdan a la de los solotildistas recalcitrantes que buscan ambigüedades imposibles para justificar su nostalgia tildica) olvidan que la polisemia es una característica inherente a las lenguas naturales y que esa ausencia de distinción ya ocurre de facto en la oralidad para millones de hispanohablantes, sin que reine por ello el caos lingüístico entre ellos. (“Sobre la ortografía”, Elena Álvarez Mellado, *elDiario.es*, 12/09/2020, [https://www.eldiario.es/opinion/zona-critica/ortografia\\_129\\_6217681.html](https://www.eldiario.es/opinion/zona-critica/ortografia_129_6217681.html)).

- despectivo:

- (1) Pero los diminutivos, y esto lo sabemos bien todos los hispanohablantes, también pueden ser muy despectivos, y lejos de atenuar el significado de la palabra, acentúan lo que pueda tener de negativo. (“Queridos y temidos diminutivos”, Óscar Esquivias, *Archiletras*, 04/12/2018, <https://www.archiletras.com/firma/queridos-y-temidos-diminutivos/>).
- (2) ‘Señoro’ tiene un sentido despectivo, señala a los varones que tratan de forma condescendiente a las mujeres. (“La vida empuja a la lengua: de señora a señor”, Lola Pons, *Verne*, 05/10/2018, [https://verne.elpais.com/verne/2018/10/05/articulo/1538748293\\_739942.html](https://verne.elpais.com/verne/2018/10/05/articulo/1538748293_739942.html)).
- (3) Pongamos por caso que un extranjero viese u oyese la expresión “sexo débil” y nos preguntase por su significado. O por “sudaca”. O por “maricón”. O por cualquier otra palabra ofensiva o denigrante. ¿Acaso le responderíamos que no lo sabemos? ¿Que esas palabras no existen? No. Lo más razonable sería que le explicásemos el significado, pero advirtiéndole de que se trata de una palabra muy despectiva que suele usarse como insulto. (“Dejad de pedirle a la RAE que elimine palabras”, Elena Álvarez Mellado, *elDiario.es*, 20/05/2017, [https://www.eldiario.es/opinion/zona-critica/dejad-pedirle-rae-elimine-palabras\\_129\\_3381428.html](https://www.eldiario.es/opinion/zona-critica/dejad-pedirle-rae-elimine-palabras_129_3381428.html)).

- significado:

- (1) Lo que dota de significado a las palabras no es su etimología, ni lo que significaron en el pasado, ni siquiera su definición en un diccionario; lo que dota de significado a una palabra es el uso que los hablantes hagan de ella. Las palabras no tienen un significado natural inherente porque el significado de las palabras es pura convención entre hablantes. (“La falacia etimológica”, Elena Álvarez Mellado, *Archiletras*, 04/10/2019, <https://www.archiletras.com/firma/la-falacia-etimologica/>).
- (2) Las palabras no significan según el significado primigenio que tenían en origen, sino según el significado que le atribuye el conjunto de hablantes de una comunidad en un momento dado. (“Lenguaje inclusivo: algunas claves lingüísticas”, Elena Álvarez Mellado, *elDiario.es*, 30/07/2021, [https://www.eldiario.es/opinion/zona-critica/lenguaje-inclusivo-claves-linguisticas\\_129\\_8021559.html](https://www.eldiario.es/opinion/zona-critica/lenguaje-inclusivo-claves-linguisticas_129_8021559.html)).
- (3) La función mágica de las palabras: aunque la relación entre significante y significado sea arbitraria, a veces los sonidos que conforman una palabra tienen una carga especial, puede ser mágica o también convertirse en un tabú. (“La función mágica de las palabras”, Mamen Horno, *Letras Libres*, 09/11/2020, <https://letraslibres.com/ciencia-y-tecnologia/la-funcion-magica-de-las-palabras/>).

- acepción:

- (1) Hablamos de “bizarro”, la palabra que, tras muchos años de lucha, ha visto reconocida su largamente reclamada acepción de ‘raro o extravagante’ en el diccionario de la RAE. (“Bizarro: Las palabras no somos estancas, evolucionamos”, Elena Álvarez Mellado, *elDiario.es*, 25/12/2021, [https://www.eldiario.es/opinion/zona-critica/bizarro-palabras-no-estancas-evolucionamos\\_129\\_8601427.html](https://www.eldiario.es/opinion/zona-critica/bizarro-palabras-no-estancas-evolucionamos_129_8601427.html)).
- (2) Casi todos los diccionarios otorgan una acepción a *leer* como verbo de lengua y otra diferente como verbo de percepción, pero ninguno –si no estoy equivocado– le asigna una acepción como verbo de consumición. (“Las palabras como prismas”, Ignacio Bosque, *Archiletras*, 01/02/2021, <https://www.archiletras.com/firma/palabras-como-prismas/>).
- (3) Y hace cuatro décadas, en 1984, incluyó una nueva acepción para nombrar cualquier proceso de “aumento rápido y por lo general alarmante de algo”, ya sean precios, delitos o, en este caso concreto, contagios y muertes por una enfermedad, sin que se señale su condición de anglicismo. (“Desescalando la nueva normalidad”, Isaías Lafuente, *Archiletras*, 02/12/2020, <https://www.archiletras.com/firma/desescalando-nueva-normalidad/>).

- metonimia:

- (1) El mecanismo de creación de palabras en todos estos casos es siempre el mismo: la metonimia, es decir, cuando utilizamos el nombre de una parte (el color del uniforme, la forma del sombrero) para denominar al todo. (“‘Piolines’: el nacimiento de una palabra”, Elena Álvarez Mellado, *elDiario.es*, 09/10/2017, [https://www.eldiario.es/opinion/zona-critica/piolines-nacimiento-palabra\\_129\\_3138250.html](https://www.eldiario.es/opinion/zona-critica/piolines-nacimiento-palabra_129_3138250.html)).
- (2) La metonimia es un proceso muy vivo en español que resulta particularmente agradecido para generar motes o al menos apelativos no demasiado bienintencionados cuando lo aplicamos a personas (“el Coletas”, “Manolito Gafotas”). “Piolines” es también un ejemplo de metonimia

nacida a partir de una anécdota un tanto estrambótica y acuñada con mucha sorna. (“Piolines”: el nacimiento de una palabra”, Elena Álvarez Mellado, *elDiario.es*, 09/10/2017, [https://www.eldiario.es/opinion/zona-critica/piolines-nacimiento-palabra\\_129\\_3138250.html](https://www.eldiario.es/opinion/zona-critica/piolines-nacimiento-palabra_129_3138250.html)).

- (3) En realidad, la expresión *oír el río* contiene una metonimia, como sucede en *oír un piano* u *oír una moto*, ya que el sustantivo designa en estos casos el sonido que produce aquello a lo que se hace referencia. (“Las palabras como prismas”, Ignacio Bosque, *Archiletras*, 01/02/2021, <https://www.archiletras.com/firma/palabras-como-prismas/>).

#### 4. CONCLUSIONES

El presente trabajo vincula la lingüística de corpus, la terminografía, los textos digitales y la divulgación de la lingüística. La terminología como disciplina teórica, la terminografía como aplicación práctica y la lingüística de corpus como metodología para el análisis de grandes cantidades de texto destacan por su relevancia en el panorama investigador de la lingüística actual. Además, la importancia que la divulgación de la ciencia tiene en la sociedad actual y la relativamente poca atención que se les ha prestado a los textos de corte divulgativo como fuentes que deben ser tenidas en cuenta para la confección de bases de datos terminológicas hacen de este trabajo una propuesta original y novedosa.

El tratamiento de los textos de corte divulgativo –escritos por especialistas, no por periodistas– nos ha permitido poner de manifiesto, en primer lugar, la concentración de este tipo de textos en unos cuantos temas: la lexicografía, el papel de las academias y la ortografía, así como cuestiones etimológicas e históricas, junto con, cierto es, la pragmática y, en ocasiones y siempre dependiendo del autor de los textos, el multilingüismo o la diversidad lingüística. Creemos que el análisis que llevamos a cabo en este trabajo puede servir como punto de partida para reflexionar, como especialistas, sobre qué otros temas podríamos divulgar (al menos en el medio que aquí tratamos, la prensa escrita digital) y qué conviene que la sociedad conozca sobre nuestras investigaciones además de lo que ya estamos contando.

La segunda parte del trabajo se ha centrado en el aprovechamiento que se puede hacer desde la terminografía de un corpus digital de textos divulgativos como el que presentamos. Se han detectado doce términos propios de la semántica (en algunas ocasiones únicos; en otras hemos unido un término y sus derivados adjetivales o de la misma familia léxica): *diccionario*, *metáfora* / *metafórico*, *prototípico*, *neologismo*, *lexicón*, *eufemismo*, *lexicografía* / *lexicográfico* / *lexicógrafo*, *polisemia*, *despectivo*, *significado*, *acepción* y *metonimia*. Para estos términos proponemos, pues, que se incluya su estatus de propios de la divulgación como estatus que revela una alta frecuencia de aparición en los textos de cualquier grado de especialidad. Además, aportamos para cada uno de ellos tres contextos extraídos del corpus, que estimamos relevantes como ejemplos del uso que se hace de ellos en este tipo de textos. Aunque el corpus es susceptible de ser ampliado, no solo con textos más antiguos o de otras zonas hispanohablantes, sino también con textos en otras lenguas, consideramos que este acercamiento a la inclusión de los textos divulgativos en el trabajo terminográfico y el establecimiento de un estatus propio de la divulgación y de contextos extraídos de un corpus de textos divulgativos contribuyen a enriquecer el futuro diccionario de términos de la semántica léxica elaborado en el Instituto de Lingüística Aplicada de la Universidad de Cádiz con datos de frecuencia de uso y contextos representativos que no son tenidos en cuenta de manera habitual en la terminografía.

## 5. REFERENCIAS BIBLIOGRÁFICAS

- Álvarez Torres, V. y S. Gómez Díaz. (2022). Terminología y documentación: la confección de un corpus metalingüístico y multilingüe de semántica léxica. En M. Casas Gómez y C. Varo Varo (Coords.). *La terminología de la semántica* (en prensa). Editorial Universidad de Sevilla.
- Auger, P. y L. J. Rousseau. (1977). *Méthodologie de la recherche terminologique*. Québec: Régie de la langue française. Traducción de J. M.<sup>a</sup> Bermúdez y G. Guerrero Ramos. Universidad de Málaga, 2002.
- Baron, N. (1984). Computer Mediated Communication as a Force in Language Change, *Visible Language*, XVIII(2), 118-141.
- Belda-Medina, J. y R. Casañ-Pitarch (Eds.) (2020). *Análisis del discurso en la era digital. Una recopilación de casos de estudio*. Comares.
- Cabré, M.<sup>a</sup> T. (1993). *La terminología. Teoría, metodología, aplicaciones*. Antártida-Empúries.
- Cabré, M.<sup>a</sup> T. (2002). Textos especializados y unidades de conocimiento: metodología y tipologización. En J. García Palacios y M. T. Fuentes Morán (Eds.). *Texto, terminología y traducción* (pp. 15-36). Almar.
- Cabré, M.<sup>a</sup> T. (2004). ¿Lenguajes especializados o lenguajes para propósitos específicos?, *Revista hispánica de Flandes y Holanda*, 26, número dedicado a: Textos y discursos de especialidad: el español de los negocios, 19-33. [https://doi.org/10.1163/9789401202893\\_003](https://doi.org/10.1163/9789401202893_003)
- Cabré, M.<sup>a</sup> T. e I. da Cunha (2022). El papel de los corpus en la terminología. Una mirada específica a la terminología del español. En G. Parodi, P. Cantos-Gómez y Ch. Howe (Eds.). *Lingüística de corpus en español* (pp. 189-203). Routledge. <https://doi.org/10.4324/9780429329296-16>
- Cantamutto, L. y C. Vela Delfa. (2016). El discurso digital como objeto de estudio: de la descripción de interfaces a la definición de propiedades, *Aposta. Revista de Ciencias Sociales*, 69, 296-323. <http://apostadigital.com/revistav3/hemeroteca/cvela2.pdf>
- Casas Gómez, M. (2006). Modelos representativos de documentación terminográfica y su aplicación a la terminología lingüística, *Revista de Lingüística y Lenguas Aplicadas*, 1, 25-36. <https://doi.org/10.4995/rlyla.2006.680>
- Casas Gómez, M. (2020a). Perspectivas actuales en el estudio del léxico, *Lynx: Panorámica de estudios lingüísticos*, 19, 5-72.
- Casas Gómez, M. (2020b). Conceptual Relationships and their Methodological Representation in a Dictionary of the Terminological Uses of Lexical Semantics, *Fachsprache. Journal of Professional and Scientific Communication*, 42(1-2), 2-26. <https://doi.org/10.24989/fs.v42i1-2.1789>
- Casas Gómez, M. (2020c). *DOCUTERM: Modelo de documentación de usos terminológicos*. Patente registrada con número 202099908175522, número de expediente: CA-00256-2020 y número de asiento registral 04/2020/4177.
- Casas Gómez, M. y M. Hummel (Eds.) (2017). *Semántica léxica*. Número monográfico extraordinario de *RILCE. Revista de Filología Hispánica*, 33(3). <https://doi.org/10.15581/008.33.3.869-911>
- Casas Gómez, M. y G. Fernández Smith (2021). El hablar y la variación de especialidad. En O. Loureda y A. Schrott (Eds.). *Manual de lingüística del hablar* (pp. 715-732). Walter de Gruyter, Colección Manuals of Romance Linguistics (MLR). <https://doi.org/10.1515/9783110335224-036>
- Casas Gómez, M. y C. Varo Varo. (2022). Redes terminológicas de la semántica léxica: bases metodológicas y recursos digitales. En M. Casas Gómez y C. Varo Varo (Coords.). *La terminología de la semántica* (en prensa). Editorial Universidad de Sevilla.
- Crespo Miguel, M. (2020). Lingüística digital: revisión de su estado actual y retos en el Instituto Universitario de Investigación en Lingüística Aplicada de la Universidad de Cádiz. *Pragmalingüística*, 28, 148-165. <https://doi.org/10.25267/Pragmalinguistica.2020.i28.08>
- De Benito Moreno, C. (2022). Uso de los medios digitales de comunicación como corpus de español. En G. Parodi, P. Cantos-Gómez y Ch. Howe (Eds.) *Lingüística de corpus en español* (pp. 480-492). Routledge. <https://doi.org/10.4324/9780429329296-36>

- Dubuc, R. (1999, 3ª ed. corregida y actualizada). *Manual práctico de terminología*. Traducción de I. Cabrera. Unión Latina. RiL editores.
- Duque, E. (2014). Análisis de contenido mediante análisis de palabras clave: La representación de los participantes en los discursos de Esperanza Aguirre. *Mediaciones Sociales. Revista de Ciencias Sociales y de la Comunicación*, 13, 39-73. <https://doi.org/10.5209/revMESO.2014.n13.49432>
- Escandell Vidal, M.ª V. (Coord.) (2011). *Invitación a la lingüística*. Editorial Universitaria Ramón Areces/UNED.
- Estopà, R. (1999). *Extracció de terminologia: elements per a la construcció d'un SEACUSE (Sistema d'Extracció Automàtica de Candidats a Unitats de Significació Especialitzada)*. Tesis doctoral. IULA, Universitat Pompeu Fabra.
- Galán Rodríguez, C. (2003). “La ciencia en zapatillas”: análisis del discurso de divulgación científica. *Anuario de estudios filológicos*, 26, 137-156.
- Galán Rodríguez, C. (2021). El discurso tecnocientífico. En O. Loureda y A. Schrott (Eds.). *Manual de lingüística del hablar (657-673)*. Walter de Gruyter, Colección Manuals of Romance Linguistics (MLR), 28. <https://doi.org/10.1515/9783110335224-033>
- Galán Rodríguez, C. y J. Montero Melchor (2002). *El discurso tecnocientífico: la caja de herramientas del lenguaje*. Arco/Libros.
- Gaviño Rodríguez, V. (2021). La polémica lingüística como tipología discursiva en la prensa española del siglo XIX. Los discursos polémicos de Fernando Gómez de Salazar. *Pragmalingüística*, (29), 173-189. <https://doi.org/10.25267/Pragmalinguistica.2021.i29.09>
- Guantiva Acosta, R., M.ª T. Cabré y J. M. Castellà Lidon (2008). Clasificación de textos especializados a partir de su terminología. *Íkala, revista de lenguaje y cultura*, 13(19), 15-39.
- Kilgarrif, A., M. Husák, K. McAdam, M. Rundell y P. Rychlý (2008). GDEX: Automatically finding good dictionary examples in a corpus. En *Proceedings of the 13th EURALEX International Congress (Barcelona, 15-19 July 2008)*, (pp. 425-432). IULA, Universitat Pompeu Fabra, Documenta Universitaria.
- Kilgarrif, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý y V. Suchomel. (2014). The Sketch Engine: ten years on, *Lexicography*, 1, 7-36. <https://doi.org/10.1007/s40607-014-0009-9>
- Llácer Llorca, E. V. y F. J. Ballesteros Roselló (2012). El lenguaje científico, la divulgación de la ciencia y el riesgo de las pseudociencias, *Quaderns de filologia. Estudis lingüístics*, 17, 51-67.
- Luzón, M. J. y C. Pérez-Llantada (Eds.) (2019). *Science Communication on the Internet. Old genres meet new genres*. John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.308>
- Mancera Rueda A. y A. Pano Alamán (2013). Nuevas dinámicas discursivas en la comunicación política en Twitter. *Círculo de Lingüística Aplicada a la Comunicación*, 56, 53-80. <https://doi.org/10.5209/revCLAC.2013.v56.43867>
- Mancera Rueda A. y A. Pano Alamán (2014). Las redes sociales como corpus de estudio para el análisis del discurso mediado por ordenador. En S. López Poza y N. Pena Sueiro (Eds.). *Humanidades Digitales: desafíos, logros y perspectivas de futuro* (pp. 305-315). Janus, Anexo 1. Recuperado el 20 de marzo de 2022, de <https://www.janusdigital.es/anexos/contribucion/descargar.htm?id=28>
- Marimón Llorca, C. (Ed.) (2020). *El columnismo lingüístico en España desde 1940. Análisis multidimensional y caracterización genérica*. Arco / Libros.
- Olmedo Estrada, J. C. (2011). Educación y Divulgación de la Ciencia: Tendiendo puentes hacia la alfabetización científica, *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 8(2), 137-148. [https://doi.org/10.25267/Rev\\_Eureka\\_ensen\\_divulg\\_cienc.2011.v8.i2.01](https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2011.v8.i2.01)
- Pano Alamán, A. y Moya Muñoz, P. (2015). CorpusRedEs. Proyecto de creación y anotación de un corpus de comunicación mediada por ordenador en español. *CHIMERA: Romance Corpora and Linguistic Studies*, 2, 117-129. <https://doi.org/10.15366/chimera2015.2>
- Parodi, G. (2008). Lingüística de corpus: una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada*, 46(1), 93-119. <https://doi.org/10.4067/S0718-48832008000100006>

- Rayson, P. (2008). From key words to key semantic domains, *International Journal of Corpus Linguistics*, 13(4), 519-549. <https://doi.org/10.1075/ijcl.13.4.06ray>
- Rodríguez Tapia, S. (2016). Los textos especializados, semiespecializados y divulgativos: una propuesta de análisis cualitativo y clasificación cuantitativa, *Signa*, 25, 987-1006. <https://doi.org/10.5944/signa.vol25.2016.16926>
- Sánchez-Saus Laserna, M. (2022). Las redes sociales como corpus para el estudio lingüístico. Revisión bibliográfica y catálogo de herramientas digitales. En L. Mañas Viniegra, I. Sacaluga Rodríguez y S. Mariscal Vega (Coords.), *Manifestaciones del humanismo en el siglo XXI* (pp. 455-464). Tirant lo Blanch.
- Scott, M. y C. Tribble. (2006). *Textual Patterns: key words and corpus analysis in language education*. John Benjamins Publishing Company. <https://doi.org/10.1075/scl.22>
- Sinclair, J. (2004). *Trust the text: language, corpus and discourse*. London/New York: Routledge. <https://doi.org/10.4324/9780203594070>
- Vela Delfa, C. y L. Cantamutto (2015). Methodological approach to the design of digital discourse corpora in Spanish. Proposal of the CÓDICE project. En P. Fuertes Olivera *et al.* (Eds.). *Current work in corpus linguistics: Working with traditionally-conceived corpora and beyond. Selected papers from the 7th International Conference on Corpus Linguistics (CILC2015)* (pp. 494-499). <https://doi.org/10.1016/j.sbspro.2015.07.472>
- Vivaldi, J. (2001). *Extracción de candidatos a término mediante combinación de estrategias heterogéneas*. Tesis doctoral. IULA, Universitat Pompeu Fabra.
- Zulaica Hernández, I. (2015). Lingüística de corpus. En J. Gutiérrez Rexac (Ed.). *Enciclopedia de Lingüística Hispánica*, volumen 1 (pp. 216-224). Routledge. <https://doi.org/10.4324/9781315713441-20>