

Nieto Caballero, G. (2021): Bloques textuales en la creación de universos ficticios de la novela realista española: un estudio de corpus. *Cultura, Lenguaje y Representación*, Vol. xxv, 245–264
ISSN 1697-7750 · E-ISSN 2340-4981
DOI: <http://dx.doi.org/10.6035/clr.2021.25.14>

Bloques textuales en la creación de universos ficticios de la novela realista española: un estudio de corpus

Textual building blocks and the shaping of fictional universes in Spanish Realism: A corpus approach

GUADALUPE NIETO CABALLERO
UNIVERSIDAD COMPLUTENSE DE MADRID
ORCID: <https://orcid.org/0000-0001-5166-7057>

Artículo recibido el / *Article received*: 2020-06-26

Artículo aceptado el / *Article accepted*: 2021-04-19

RESUMEN: En este artículo se presenta un análisis sobre bloques textuales en la novela realista española y su función como elemento que ayuda a configurar los universos ficticios de las historias. El análisis se ha llevado a cabo sobre un corpus de cien novelas de nueve autores canónicos (c. 8 millones de palabras). Para identificar los bloques textuales hemos empleado una metodología de estilística de corpus, que nos ha permitido identificar los ejemplos de manera automática y calibrar su representatividad. Los resultados obtenidos han sido clasificados según grandes bloques de significado: espacio, tiempo, lenguaje corporal, discurso. El análisis de los ejemplos revela la existencia de una serie de hábitos lingüísticos que podríamos considerar transversales y propios de la novela realista como movimiento literario y que van más allá de la idiosincrasia estilística de cada autor.

Palabras clave: bloques textuales, universos ficticios, novela realista, estilo, enfoque de corpus.

ABSTRACT: This paper analyses textual building blocks in fictional narratives of Spanish Realism, focusing on their function as a textual device with which authors shape their fictional narratives. The analysis has been carried out in a corpus of one hundred novels by nine canonical authors (c. 8 million words). In order to identify textual building blocks a corpus-stylistic approach has been used, which has made it possible to automatically retrieve the examples and gauge their textual saliency. Textual building blocks have been classified in four main groups: space, time, body language and discourse. The analysis of the

examples reveals a series of linguistic habits that may be considered common to Spanish Realism in general, going well beyond the idiosyncratic style of individual authors.

Key words: textual building blocks, fictional universes, Spanish Realism, style, corpus approach.

1. INTRODUCCIÓN

En este artículo se analiza el uso de bloques textuales en la creación de universos ficticios en la novela realista española. Los bloques textuales o *textual building blocks* (Mahlberg 2013: 26)¹ son unidades que contribuyen a dar forma a la realidad literaria que se nos presenta. Estas unidades pueden estar tan interiorizadas dentro del registro al que pertenecen (en nuestro caso, la novela realista española) que ni los lectores (incluyendo críticos) ni los propios autores son necesariamente conscientes de su valor (Mahlberg, Wiegand, Stockwell y Hennessey, 2019: 349). Sin embargo, más allá de la idiosincrasia de cada autor o de las particularidades estilísticas que dominan un movimiento literario —en el caso de la novela realista, por ejemplo, podríamos mencionar la variación interna de la lengua según los ejes clásicos: diatópico, diastrático y diafásico (Rodríguez Marín, 2005)—, también existen hábitos menos llamativos pero igualmente significativos que pueden contribuir a configurar la base sobre la que se construyen la realidad literaria de los textos y, por qué no, sobre la que se definen estos movimientos. Con esta hipótesis, este artículo pretende investigar y ver cómo los universos ficticios de la novela realista se construyen sobre bloques textuales que van más allá de la idiosincrasia estilística de cada autor y que no han sido explorados por la crítica literaria tradicional.

Para llevar a cabo el análisis hemos utilizado un enfoque de corpus, cuyas metodologías ofrecen la ventaja de poder trabajar con cantidades de texto inasumibles para disciplinas de corte más tradicional. En este artículo hemos utilizado un corpus de más de ocho millones de palabras formado por 100 novelas publicadas por nueve autores canónicos (véase apartado 3). Sobre este corpus hemos identificado segmentos de texto que se repiten en varios textos de varios autores atendiendo a parámetros que garanticen su representatividad y sugieran una relevancia estilística. Esta metodología podría englobarse dentro de la estilística de corpus (McIntyre y Walker, 2019), una disciplina que aplica «corpus methods to the analysis literary texts, giving particular emphasis to the relationship between linguistic description and literary appreciation» (Mahlberg, 2014: 378). La eficacia de los estudios de estilística de corpus ha quedado demostrada en trabajos sobre autores de referencia que se han realizado en la última década. Dentro de la literatura anglosajona, por ejemplo, las obras de figuras de referencia como William Shakespeare (Culpeper, 2009), Jane Austen (Fischer-Starcke, 2010), Virginia Woolf (Balossi, 2014) o Charles Dickens (Mahlberg, 2013; Ruano San Segundo, 2016), entre otras, han sido analizadas utilizando enfoques de corpus. Estos ejemplos, sin embargo, contrastan con la falta de trabajos que adopten este tipo de enfoques en el análisis de textos literarios en lengua española, donde la estilística de corpus constituye una disciplina aún por explotar. En este sentido, podría decirse que el presente artículo

¹ Además de *textual building blocks*, los bloques textuales también son conocidos, en inglés, como *basic building-blocks* o *world-building elements* (Gavins, 2007: 36).

también pretende contribuir a ilustrar el potencial de los estudios de corpus en el análisis de textos literarios en español, un área todavía escasa de estudios que adopten este tipo de enfoques.

El artículo se organiza como sigue. En primer lugar, se realiza una breve contextualización del marco de estudio, en la que nos centramos en los universos ficticios y los bloques textuales sobre los que se construye el análisis (apartado 2). En el apartado 3 se presenta el corpus de estudio. En el apartado 4 se detalla la metodología empleada para localizar los ejemplos y se muestran los resultados clasificados según grandes bloques de significado. El apartado 5, que comprende la parte analítica, se divide en dos subapartados. En primer lugar, se realiza una aproximación general a las construcciones identificadas a través de bloques textuales que sirven para configurar la dimensión espacial de las novelas (apartado 5.1), para luego centrar el foco de nuestro análisis en un bloque textual sobre lenguaje corporal (*las manos en los bolsillos*) que servirá para ilustrar cómo un segmento de texto específico es empleado en distintas novelas del corpus para configurar un significado concreto que va más allá de su sentido literal (apartado 5.2). El artículo finaliza con sugerencias sobre futuras vías de análisis basadas en las posibilidades que ofrecen los resultados que aquí se presentan pero que, por cuestiones de espacio, no pueden ser analizadas en profundidad.

2. UNIVERSOS FICTICIOS Y BLOQUES TEXTUALES

Los universos ficticios son la realidad literaria de la que se ocupa la teoría de los mundos textuales (*Text-World Theory*) (Werth, 1999; Gavins, 2007). Un mundo textual o universo ficticio puede definirse como el conjunto de «mental representations constructed in the process of discourse comprehension» (Mahlberg, 2013: 34). Estos universos ficticios definen las circunstancias físicas y el contexto social que hacen posible la existencia de los personajes que pueblan la historia. Desde un punto de vista estilístico, los universos ficticios se construyen en gran medida recurriendo a bloques textuales, que desempeñan funciones literarias determinadas. Aunque pueda ser una representación de algo que sí que existe, conviene incidir en que un universo ficticio no existe fuera del texto literario. En otras palabras, aunque al Madrid de los *Episodios Nacionales* le concedemos verosimilitud y somos capaces de *ver* el Madrid del siglo XIX, como lectores somos conscientes de que se trata de una construcción literaria en el marco de la novela, como todo lo que hay dentro de ella: los lugares, las circunstancias (el tiempo meteorológico, por ejemplo), los objetos o los personajes y sus relaciones. Si todo ese universo ficticio solo existe como dimensión literaria, no es difícil aceptar que su construcción dependerá en gran medida de las elecciones del autor para darle forma en plano textual. Así, los segmentos de texto empleados para definir los antedichos lugares, circunstancias, los objetos o las personas tienen una relevancia estilística clara, aunque solo sean porque son parte de un texto finito y el resultado de una elección concreta por parte del autor.

Tradicionalmente, los estudios de estilística se han ocupado de analizar el proceso de lectura con el fin de «arrive at a detailed account of how readers understand particular texts in the ways they do» (Short y Semino, 2008: 117). Los análisis de estilística de corpus, con una motivación teórica similar, siguen la estela marcada por los análisis de estilística tradicional. Sin embargo, presentan una ventaja fundamental: las metodologías de corpus permiten identificar patrones estilísticamente relevantes que habitualmente han pasado desapercibidos en el análisis de esos mismos textos. Para entender cómo la estilística de corpus se ocupa del análisis de los universos ficticios a través de construcciones tradicionalmente desatendidas, debemos detenernos en dos conceptos: las

funciones textuales locales (*local textual functions*) y los bloques textuales (*textual building blocks*). Las funciones textuales locales (Mahlberg, 2007: 4) se refieren a aquellos segmentos de texto que son propios de un autor o incluso de una obra literaria concreta. Es decir, son locales en tanto que «they do not claim to capture general functions, but functions specific to a (group of) text(s)» (Mahlberg, 2007: 4). Cuando se analizan los universos ficticios, las funciones textuales locales suelen ser el objeto de estudio más habitual, pues sirven para revelar idiosincrasias estilísticas de un autor que pueden haber pasado desapercibidas en su exégesis literaria. De nuevo, el canon anglosajón nos ofrece ejemplos dignos de mención en este sentido, como los estudios llevados a cabo con Joseph Conrad (Stubbs, 2005), Jane Austen (Fischer-Starcke, 2006) o Charles Dickens (Mahlberg, 2007), entre otros. En español también existe un estudio de corpus de funciones textuales locales llevado a cabo con Galdós, si bien su ámbito se ciñe a la configuración del espacio en sus novelas (Nieto Caballero, 2019).

Cuando en lugar de circunscribirse a un texto (o a varios) de un solo autor, el estudio de segmentos de texto pretende ser transversal, nos encontramos, como en el caso de este artículo, con bloques textuales que van más allá de la individualidad de cada autor. Además de ser unidades que contribuyen a dar forma a la realidad literaria que se nos presenta, estos bloques textuales pueden revelar «conventional patterns and norms» (Mahlberg, 2007: 3) dentro del marco en el que aparecen. En otras palabras, mientras que el significado de los segmentos de texto que se analizan en los estudios de funciones textuales locales tienen una dimensión subjetiva de cómo un individuo los utiliza, en el caso de los bloques textuales empleados en un movimiento literario (la novela realista, en nuestro caso), el significado de los ejemplos adquiere una dimensión que podríamos denominar social, en tanto que este significado «is shared within the discourse community» (Mahlberg, 2007: 3). En nuestro caso, el estudio que presentamos pretende arrojar luz sobre mecanismos trasversales y comunes al autor realista en general. Es decir, más allá de la idiosincrasia estilística de cada autor que forma nuestro corpus, este estudio se centra en ejemplos que contribuyen a definir el sustrato sobre el que se construyen los universos ficticios de la novela realista en general.

Es importante mencionar, antes de cerrar este apartado, que los bloques textuales no son necesariamente construcciones llamativas o con particularidades lingüísticas dignas de reseña. Más bien al contrario, pues su frecuencia se debe, en gran medida, a que forman parte de un discurso naturalizado. De hecho, como mencionábamos al comienzo del artículo, ni los autores ni tampoco los lectores tienen por qué ser conscientes de la existencia de este tipo de bloques como unidades estilísticamente relevantes que contribuyen a definir la realidad literaria (Mahlberg, Wiegand, Stockwell y Hennessey, 2019: 349). Esta es la razón por la que un enfoque de estilística de corpus resulta fundamental para poder identificarlos y analizar su valor, pues una metodología de corpus permite descubrir «patterns that we as readers may not be aware of, although such patterns might still contribute to the effects we perceive» (Mahlberg, 2013: 27). En concreto, en este artículo mostraremos cómo diferentes aspectos de la creación de los universos ficticios de novelas de distintos autores tienen aspectos en común y cómo el significado que construimos al leer las novelas se articula sobre segmentos de texto que podríamos considerar característicos de la novela realista como movimiento literario en general.

3. EL CORPUS DE ESTUDIO

El estudio se ha desarrollado con un corpus de novela realista formado por cien novelas de nueve autores diferentes: Pedro Antonio de Alarcón, Vicente Blasco Ibáñez, Leopoldo Alas «Clarín», Luis Coloma, Armando Palacio Valdés, Emilia Pardo Bazán,

José María de Pereda, Benito Pérez Galdós y Juan Valera (véase apéndice). La elección de este periodo literario se explica por varias razones. En primer lugar —y esta es una razón importante en estudios de este tipo— porque toda la producción de estos autores se encuentra digitalizada y libre de derechos, de manera que es más fácil trabajar con ella sin incurrir en problemas legales. En segundo lugar, por la relevancia de estos escritores en la historia literaria española, en general, y en la historia de la novela española, en particular. No cabe duda de que el periodo realista decimonónico es uno de los periodos más sobresalientes de nuestra novela; no en vano, en ella coinciden Benito Pérez Galdós, Clarín y Emilia Pardo Bazán, considerados por una parte amplia de la crítica figuras imprescindibles de la literatura española².

La extensión del corpus asciende a 8 156 998 palabras. Todos los textos que forman el corpus de estudio se han recopilado del repositorio digital de la Biblioteca Virtual Miguel de Cervantes (s.f.)³. Este repositorio cuenta con alrededor de doscientos mil registros bibliográficos, de los cuales sesenta mil son libros. La calidad de la versión digitalizada de los textos, fundamental en los estudios de corpus como el que aquí se realiza, está fuera de toda duda, pues la Biblioteca cuenta con un consejo científico que avala el rigor de los materiales alojados en el repositorio.

Aunque la relación de autores incluidos en el corpus no parece requerir mayor justificación dado su carácter canónico (Ferrerías, 1982; Pedraza y Rodríguez, 1983; Pedraza y Rodríguez, 2012: 236–237), es cierto que la horquilla de tiempo que ocupan (con textos publicados desde 1855 hasta 1930, aproximadamente) puede requerir una explicación. Sirva como ejemplo la inclusión de Blasco Ibáñez, el más tardío de los nueve autores. Se puede considerar que el escritor valenciano es un epígono de la generación realista o grupo del 68 si atendemos a factores natalicios. Por cronología, le correspondería pertenecer a la promoción de fin de siglo, si bien se le suele ubicar a medio camino entre ambas promociones, sobre todo por su tendencia naturalista. Esa tendencia es la que enlaza directamente con otros autores del corpus. Efectivamente, las heterogéneas producciones que conviven a finales del XIX llegan hasta bien entrado el siglo XX, como demuestran «la última etapa de Galdós, las producciones finales de D.^a Emilia y de Palacio Valdés y la ingente obra de Blasco, que culmina y resume, en floración tardía, todo el proceso del realismo y del naturalismo de nuestra novela» (Alborg, 1996: 423). Por estas razones no parece desacertado incluir a Blasco Ibáñez en nuestro corpus.

Por último, también conviene incidir en la presencia desigual de cada autor en el corpus en lo que al número de novelas se refiere (véase apéndice). Estas diferencias se deben fundamentalmente al grado de prolijidad de los distintos novelistas. En otras palabras, autores con una producción más reducida, como Clarín, gozan de menos peso en el corpus de referencia que otros más prolijos, como Pardo Bazán. Además, cabe señalar que, en el caso de aquellos autores con una producción más extensa, se ha hecho una selección de como mucho veinte títulos, de tal suerte que las diferencias entre los distintos autores no sean excesivamente acusadas. La selección de veinte textos pretende ser lo más equitativa posible. Sirva como ejemplo Galdós, el autor con una producción más extensa y del que ha habido que dejar más títulos fuera. Como se puede observar en el apéndice, las veinte novelas seleccionadas ocupan el espectro tanto de las cinco series de los *Episodios Nacionales* como de sus novelas de primera época y de la serie contemporánea.

² Véanse, por ejemplo, las apreciaciones de Iris Zavala sobre Galdós y Clarín. Del primero afirma que para sus coetáneos “fue, sin duda, el novelista por excelencia” (Zavala, 1982: 463). Sobre *La Regenta*, de Clarín, apunta que es “una de las cumbres de la narrativa en lengua española” (Zavala, 1982: 563).

³ *Biblioteca Virtual Miguel de Cervantes*. Universidad de Alicante, s. f. 1 de junio de 2020.

En definitiva, y teniendo en cuenta los matices que hemos detallado sobre los autores que hemos incluido y la selección de textos que hemos realizado, consideramos que este corpus de cien novelas de nueve autores y más de 8 millones de palabras constituye una muestra de literatura realista española lo suficientemente sólida como para poder realizar un estudio cabal con el que extraer conclusiones que puedan ser válidas para la novela realista en general.

4. IDENTIFICACIÓN Y CLASIFICACIÓN DE PATRONES

Para desarrollar el estudio hemos utilizado *WordSmith Tools* (Scott, 2016), la herramienta más empleada en los estudios de estilística de corpus (Archer, 2007: 249). En concreto, hemos utilizado esta *software* para identificar *clusters* en nuestro corpus de estudio.⁴ Desde un punto de vista puramente formal, un *cluster* es una secuencia de dos o más palabras empleada de forma repetida en un corpus de textos (Cheng, 2012: 72). Estas secuencias de palabras repetidas en un corpus de textos también reciben otros nombres. Biber, Johansson, Leech, Conrad y Finegan (1999: 989), por ejemplo, utilizan el término «lexical bundles» para referirse a estos cuando se analizan como parte de un registro de la lengua. En otros estudios han sido denominados «recurrent word-combinations» (Altenberg, 1998), «chains» (Stubbs y Barth, 2003), o «n-grams» (Anthony, 2019). Siguiendo al creador de *WordSmith*, en este estudio hemos optado por la etiqueta *cluster*, por estar definida únicamente por aspectos formales (Scott, 2019). Desde un punto de vista funcional, a los *clusters* identificados en un corpus se les presupone «identifiable discourse functions in texts» (Conrad y Biber, 2005: 58), por lo que pueden ser interesantes como reveladores de bloques textuales (Mahlberg, 2013: 26). Este valor hace necesario extremar las precauciones a la hora de fijar los parámetros que utilizamos para identificarlos en nuestro corpus, pues debemos garantizar una muestra de ejemplos lo más equilibrada posible. Para definir los parámetros de identificación de *clusters* con nuestro *software*, debemos tomar en consideración tres aspectos fundamentales: la longitud de los ejemplos, su distribución en los textos que forman el corpus y su frecuencia de uso.

En cuanto a la longitud de los ejemplos, los análisis suelen girar en un rango de ejemplos formados por entre tres y cinco palabras. Biber, Johansson, Leech, Conrad y Finegan (1999: 992) apuntan que los de tres palabras, aunque más numerosos por su menor longitud suelen ser de tipo gramatical. Los más largos, por el contrario, son «more phrasal in nature and correspondingly less common» (Biber, Johansson, Leech, Conrad y Finegan 1999: 992). Lo cierto es que una unidad más larga puede ser más útil para identificar funciones textuales, si bien se localizarán menos ejemplos. Como señala Mahlberg en su estudio de la novela victoriana en general y de Dickens en particular, «(a) length of five has been shown to be a useful starting point for the analysis of fiction» (Mahlberg, 2013). Esta es, asimismo, la longitud de los ejemplos analizados en Nieto Caballero (2019) en el caso de Galdós. El precedente teórico de estas dos referencias nos ha hecho decantarnos por una longitud de cinco palabras, descartando cualquier ejemplo de una extensión menor.

Sobre la distribución de los ejemplos, es necesario que estos no se localicen en un solo texto o en textos de un solo autor, pues los resultados perderían validez como bloques textuales característicos de la novela realista en general. Sin embargo, es importante mencionar que, en el caso de ejemplos de cinco palabras, no podemos pretender obtener una muestra de ejemplos que aparezca en todos los textos. En el estudio de Stubbs y Barth

⁴ Para una información detallada de qué es un *cluster* y cómo se genera utilizando *WordSmith Tools*, véase Scott (2013).

(2003: 76), por ejemplo, ninguno de sus ejemplos de cinco palabras aparece en todos los textos (teniendo en cuenta, además, que su corpus está formado por únicamente veintitrés textos). En los estudios de *clusters* que pretender generalizar funciones se suele fijar un límite de al menos cinco textos distintos (Biber, Conrad y Cortes, 2004: 376). Cuanto más alta sea esta barrera, más sólida será la muestra de ejemplos localizados. En nuestro caso, la barrera de cinco ejemplos resulta insuficiente, pues hay autores cuya representatividad en el corpus alcanza los veinte títulos. Por ello, hemos optado por establecer una barrera de veinte textos. Este requisito nos permitirá descartar aquellos ejemplos que, como comentábamos, tengan una inclinación hacia la idiosincrasia estilística de un autor o incluso de una obra concreta (imagínese la muletilla «en toda la extensión de la palabra» empleada por Doña Lupe en *Fortunata y Jacinta*, por ejemplo) y asegurar que la muestra de ejemplos es representativa de hábitos lingüísticos ejercidos por los distintos autores que conforman el corpus.

La frecuencia, por último, es el aspecto sobre el que resulta más difícil tomar una decisión. Como señala Kopaczyk, «[t]here is no uniform practice in lexical bundle studies [...]. Every researcher takes an informed but idiosyncratic decision» (Kopaczyk, 2013: 152). Existen casos en los que la frecuencia se mide en términos absolutos y otros que lo miden en términos relativos (por ejemplo, normalizando la frecuencia de uso por cada millón de palabras de texto). El primer caso es el más habitual cuando se trabaja con corpus especializados, mientras que el segundo es más frecuente cuando se trabaja con corpus extensos. Si bien nuestro corpus es especializado (incluye únicamente novelas; todas ellas del realismo decimonónico), lo cierto es que se trata de un corpus extenso (más de 8 millones de palabras). Es por ello que lo más lógico parece ser recurrir a frecuencias normalizadas. En concreto, hemos utilizado la fórmula de Biber, Johansson, Leech, Conrad y Finegan (1999: 993), que consideran susceptibles de análisis solo aquellos ejemplos con una frecuencia de cinco por cada millón de palabras.

En resumen, hemos seleccionado aquellos *clusters* con una longitud de cinco palabras, que aparezcan en al menos veinte textos diferentes y con una frecuencia de al menos cinco apariciones por cada millón de palabras. En total, con estos parámetros se han localizado cuarenta y seis ejemplos. Los resultados se muestran en la tabla 1.

Tabla 1. Ejemplos identificados en el corpus de estudio

N	Ejemplo	Frecuencia absoluta	Núm. Textos	Frecuencia por millón de palabras	Grupo
1	A LA PUERTA DE LA	161	58	19,73765348	espacio
2	AL FIN Y AL CABO	86	41	10,54309441	discurso
3	EN LA CUENTA DE QUE	84	43	10,29790617	discurso
4	AL PIE DE LA LETRA	83	40	10,17531204	otro
5	DE UN LADO A OTRO	81	34	9,930123803	espacio
6	DE UN MOMENTO A OTRO	81	35	9,930123803	tiempo
7	EN EL FONDO DE SU	80	40	9,807529682	otro
8	LA MAYOR PARTE DE LOS	76	46	9,317153198	otro
9	LE DIGO A USTED QUE	74	38	9,071964956	discurso
10	LA PALMA DE LA MANO	71	38	8,704182593	lenguaje corporal
11	LAS MANOS EN LOS BOLSILLOS	69	43	8,458994351	lenguaje corporal
12	EN LA PUERTA DE LA	69	36	8,458994351	espacio
13	DE LA NOCHE A LA	65	33	7,968617867	tiempo

14	EN EL MOMENTO EN QUE	64	42	7,846023746	tiempo
15	LA NOCHE A LA MAÑANA	62	30	7,600835503	tiempo
16	LA PUERTA DE LA CALLE	62	37	7,600835503	espacio
17	DE LA CALLE DE LA	61	26	7,478241382	espacio
18	NI MÁS NI MENOS QUE	59	37	7,23305314	discurso
19	EN EL FONDO DE LA	59	35	7,23305314	otro
20	LA MAYOR PARTE DE LAS	58	37	7,110459019	otro
21	A UN LADO Y A	57	28	6,987864898	espacio
22	A LA LUZ DE LA	57	32	6,987864898	espacio
23	UN LADO Y A OTRO	56	27	6,865270777	espacio
24	LA PUERTA DE SU CASA	55	33	6,742676656	espacio
25	EN EL CENTRO DE LA	55	33	6,742676656	espacio
26	ALTAS HORAS DE LA NOCHE	54	43	6,620082535	tiempo
27	EN MEDIO DE LA CALLE	54	27	6,620082535	espacio
28	LAS TRES DE LA TARDE	53	29	6,497488414	tiempo
29	LA PUERTA DE LA CASA	52	33	6,374894293	espacio
30	A LA HORA EN QUE	49	29	6,00711193	tiempo
31	A LAS TRES DE LA	47	27	5,761923688	tiempo
32	A LA PUERTA DE SU	47	34	5,761923688	espacio
33	LAS MANOS A LA CABEZA	46	22	5,639329567	lenguaje corporal
34	EN EL CASO DE QUE	46	22	5,639329567	discurso
35	EN EL SENO DE LA	46	29	5,639329567	otro
36	POR QUÉ NO HA DE	45	24	5,516735446	discurso
37	A LO LARGO DE LA	45	29	5,516735446	otro
38	EN LA CALLE DE LA	44	25	5,394141325	espacio
39	PERO ES LO CIERTO QUE	44	27	5,394141325	discurso
40	EL FONDO DE SU ALMA	44	25	5,394141325	otro
41	A FIN DE QUE NO	43	26	5,271547204	otro
42	A LOS PIES DE LA	42	27	5,148953083	espacio
43	LO QUE A MÍ ME	42	24	5,148953083	discurso
44	SIN DARSE CUENTA DE ELLO	41	24	5,026358962	otro
45	EL SILENCIO DE LA NOCHE	41	26	5,026358962	otro
46	HASTA EL PUNTO DE QUE	41	31	5,026358962	otro

Para analizar estos ejemplos, sin duda lo más práctico es recurrir a una clasificación. Las clasificaciones de *clusters* suelen realizarse tomando como referencia criterios formales o funcionales (Mahlberg, 2007: 9). Teniendo en cuenta los propósitos de este artículo, lo más adecuado es agrupar los ejemplos en grupos que potencialmente desempeñen funciones en ámbitos parecidos. Estas clasificaciones deben hacerse «with no *a priori* assumptions» (Mahlberg, 2007: 13), es decir, evitando consideraciones previas. Se trata, pues, de mirar los ejemplos e intentar agruparlos en bloques que resuman las funciones que pueden desempeñar. Teniendo en cuenta los ejemplos localizados, hemos recurrido a cinco grupos: tiempo, espacio, discurso, lenguaje corporal y «otros».

Los ejemplos de «tiempo» y «espacio» son tal vez los más elocuentes, pues se refieren a aspectos que tienen que ver con la definición de la dimensión temporal (*las tres de la tarde*) y espacial (*de un lado a otro*) que configuran los universos ficticios. El grupo «discurso» engloba aquellos casos que forman parte de la representación del discurso de los personajes (*le digo a usted que*) y que pueden ser útiles para ver cómo los autores recurren a segmentos de texto similares para configurar funciones similares (principios de cortesía, por ejemplo). El grupo «lenguaje corporal», también relacionado con los personajes, incluye descripciones por parte del narrador que sirven para construir a las figuras que pueblan los universos ficticios mediante comportamientos habituales que pueden desempeñar funciones concretas (*las manos en los bolsillos*). Por último, el grupo «otros» recoge aquellos ejemplos que no encuentran asiento en ninguno de los grupos anteriores. Aquí se incluyen desde ejemplos de fraseología (*al pie de la letra*) hasta segmentos de texto que forman parte del discurso narrativo pero que tienen un origen más idiomático que funcional en lo que a la configuración de los universos ficticios se refiere (*la mayor parte de los*). La distribución de los ejemplos en estos grupos se puede observar en la tabla 1.

5. ANÁLISIS

Al contrario que en la metodología empleada para la localización de ejemplos, en los que la automatización del procedimiento permite identificarlos sin la interferencia del investigador (por ejemplo, de nuestro conocimiento previo del texto), en el análisis debe primar el componente cualitativo. En otras palabras, una vez localizados los ejemplos que cumplen los requisitos descritos en el apartado anterior para garantizar su representatividad (de extensión, distribución en el corpus y frecuencia de uso), los ejemplos localizados deben analizarse en su contexto y en el marco de la literatura científica existente. Aunque por limitaciones de espacio el análisis que presentamos no puede ser demasiado extenso, a continuación nos ocupamos del valor de algunos de los ejemplos localizados, tanto desde un punto de vista más general (utilizando la clasificación que acabamos de explicar como marco punto de partida) como desde una perspectiva más específica (analizando el valor de un ejemplo concreto como bloque textual que contribuye a configurar significados similares en distintos textos del corpus).

5.1. CARACTERIZACIÓN DEL ESPACIO MEDIANTE BLOQUES TEXTUALES

Los grupos que hemos utilizado para clasificar los ejemplos localizados pueden (y de hecho deben) ser analizados con dos enfoques bien diferenciados. En primer lugar, los *clusters* que forman cada grupo deben ser considerados en conjunto, de manera que podamos analizarlos «in terms of tendencias on a continuum» (Mahlberg, 2007: 17). Este aspecto nos permitirá ver cómo elementos de un mismo grupo sirven para configurar ese aspecto en la creación de los universos ficticios. Sirva como ejemplo el grupo «espacio». Con quince ejemplos, se trata del grupo más numeroso: *a la puerta de la, de un lado a otro, en la puerta de la, la puerta de la calle, de la calle de la, a un lado y a, a la luz de la, un lado y a otro, la puerta de su casa, en el centro de la, en medio de la calle, la puerta de la casa, a la puerta de su, en la calle de la y a los pies de la*. Que sea el grupo más numeroso no debe resultar especialmente sorprendente, pues, junto con el tiempo, los personajes y los acontecimientos que tienen lugar en la historia, el espacio es uno de los pilares estructurales de la sintaxis narrativa: toda narración implica una serie de acontecimientos y personajes que se desarrollan en un tiempo y también en un espacio determinados (Zubiaurre, 2000: 20–21). La importancia del espacio en el género

novelesco es tal que hay quienes llegan a distinguir entre tipologías de novelas como «novelas de situaciones» o «de espacio» (Valles Calatrava, 2008: 178). En el caso de la novela realista española, el espacio se erige como un elemento de importancia cardinal en la ilusión de transparencia y el efecto de verosimilitud que domina este movimiento⁵.

Por ello, siendo la reproducción fiel del mundo exterior uno de los pilares del realismo (Martínez Carazo, 2006: 36–37), la actitud del novelista no puede ser sino objetiva e impersonal, tratando de describir los espacios de manera detallada y exacta. Los ejemplos identificados en nuestro corpus nos permiten ver esa tendencia en la definición del espacio, lo que revela bloques textuales compartidos por los nueve autores que desempeñan funciones similares en la configuración de los universos ficticios que nos plantean en sus historias. El ejemplo más claro lo encontramos en los bloques textuales que contienen referencias espaciales concretas, tanto a calles (*la puerta de la calle, de la calle de la, en medio de la calle, en la calle de la*) como a casas (*la puerta de su casa, la puerta de la casa*) e incluso puertas (*a la puerta de la, de un lado a otro, en la puerta de la, la puerta de la calle, la puerta de su casa, la puerta de la casa, a la puerta de su*). La existencia de bloques textuales con estos referentes concretos de tipo situacional revela un hábito lingüístico compartido por los autores realistas en la configuración del espacio de sus novelas, que encaja con la descripción sistemática de los espacios que domina la novela realista que otorga a las escenas, y sobre todo a los personajes que las protagonizan, un entorno físico corpóreo, que hace que la historia parezca real (cf. Matzat, 2007)⁶.

En segundo lugar, sobre esta lectura general que podemos hacer de los ejemplos como parte de una tendencia, los usos de cada ejemplo pueden analizarse de forma individualizada recurriendo a una concordancia. Esto permite una lectura vertical de los ejemplos (Tognini-Bonelli, 2001: 18), que hace posible, a su vez, identificar patrones más específicos. Por continuar con el ejemplo de los bloques textuales de tipo espacial, si tomamos como ejemplo las calles (*la puerta de la calle, de la calle de la, en medio de la calle, en la calle de la*), en general advertimos que el autor realista recurre a este referente para situar la acción presentada en un espacio real, logrando un alto grado de verosimilitud en el marco del paisaje urbano en el que transcurren la gran parte de sus historias. Aunque es bien sabido que las calles son una parte importante del paisaje urbano en la tendencia realista (González, 1998: 32), una concordancia de estos ejemplos nos permite identificar los hábitos más frecuentes compartidos por los distintos autores en el uso de este referente espacial. El uso más habitual que hemos localizado es, además de posicionar lugares en los que tiene lugar la acción mediante su enclave en una vía determinada mencionado más arriba, plantear la referencia como el lugar en el que se detiene el personaje y sobre el que se construye la descripción de la escena. Sirvan como botón de muestra los ejemplos de *de la calle de la* mostrados en (1) y (2), extraídos de novelas de Galdós y Clarín respectivamente. Como se puede observar, en ambos casos los personajes se detienen en un punto definido sobre la ubicación en una calle determinada y, a partir de ahí, se presenta el resto de la escena al lector. Esta función no es exclusiva de este bloque textual, sino de ejemplos que contienen referencias a las calles, lo que refuerza su función estilística. Tal es el caso del ejemplo (3) de *en la calle de la* perteneciente a *La Quimera*, de Pardo Bazán. Aunque el ejemplo se refiere a un coche en lugar de a un personaje, la descripción se construye en términos similares: el

⁵ El realismo, a diferencia de la literatura romántica que le antecede, que rompía con la concepción mimético-realista de la obra en pos de la libertad de creación imaginativa, tiende a considerar la obra como un espejo, una copia de la realidad.

⁶ La función del espacio en Galdós, con especial atención en las calles de sus novelas, ha sido tratado en Nieto Caballero (2019).

coche (del que se baja un personaje) se detiene en un punto definido por su ubicación en una calle concreta, que sirve como eje vertebrador de la escena que se presenta.

- (1) Morton le observó cuando estuvo lejos, le vio detenerse en la esquina *de la calle de la Poterna* donde había un farolillo, y examinar la moneda a la débil claridad de la antorcha municipal; después le vio inclinarse al suelo para sonar la pieza de oro sobre una piedra, y luego el anciano volvió corriendo al lado de Daniel Morton. (*Gloria*, Perez Galdós, capítulo 9)
- (2) Don Pompeyo, que daba diente con diente, de frío con fiebre, se detuvo en lo más alto *de la calle de la Rúa* para contemplar aquella muchedumbre apiñada a los pies de la torre, en tan estrecho recinto, cuando podía extenderse a sus anchas por toda la plazuela. (*La Regenta*, Clarín, capítulo 26)
- (3) No tuvo tiempo de completar la descripción. Ya el coche se paraba ante una joyería, *en la calle de la Paz*. Espina se bajó, ayudada por Silvio, y el joyero, solícito, enseñó modelos; discutieron el arreglo y aumento de los largos hilos de gruesas perlas que Espina poseía y pensaba escalonar sobre el escote, a lo Médicis, y para los cuales deseaba un broche espléndido, un rubí único en tamaño, color y talla. (*La Quimera*, Pardo Bazán, capítulo 3)

Además del hábito de construir la escena sobre el lugar en el que se detiene un personaje definido por su enclave en una vía determinada, en los ejemplos que contienen referencias a calles también se advierten patrones igualmente significativos que dan cuenta de hábitos estilísticos compartidos por la pléyade de autores realistas en lo que al tratamiento del espacio se refiere. Tal es el caso de la identificación de negocios por su ubicación en una calle determinada. De nuevo encontramos ejemplos en los distintos autores que forman el corpus. En el ejemplo (3) que acabamos de mostrar, de hecho, el bloque textual hace referencia a una joyería. Por citar dos autores distintos, en los ejemplos (4) y (5) vemos dos usos de *en la calle de la* en novelas de Palacio Valdés y Pereda en los que nuevamente el bloque textual hace referencia a sendos negocios. La diferente naturaleza de los lugares —una joyería, un almacén de quincalla o un café— no es óbice para que el autor realista los sitúe en una calle concreta, de la que parece servirse para otorgar esa sensación de realismo a la que se suele referir la crítica tradicional y que ya hemos mencionado más arriba.

- (4) Los comerciantes recibían a sus amigos en las tiendas, departían y reían con ellos y apenas se curaban de la venta de sus artículos. Había un tendero llamado Braulio que poseía *en la calle de la Herrería* un bastante bien surtido almacén de quincalla. (*La novela de un novelista*, Palacio Valdés, capítulo 19)
- (5) Diéronme por de pronto minuciosas señas de la calle del Príncipe, porque yo les dije que en ella vivía don Augusto Valenzuela, a quien necesitaba visitar; me explicaron cómo podría yo, recién llegado a Madrid, con algún dinero en el bolsillo, pasarlo regularmente entretenido, de día brujuleando por las calles, de noche con ellos, a primera hora en el café de La Esmeralda, *en la calle de la Montera*, y más tarde en Capellanes o en el paraíso del Teatro Real, etc. (*Pedro Sánchez*, Pereda, capítulo 9)

En definitiva, los bloques textuales con referencias a calles revelan hábitos estilísticos que, además de reforzar la consabida importancia del espacio en la novela realista, sirven para identificar algunos de los criterios de caracterización del espacio del referente y su función. Como se puede observar, en la novela realista, en donde el elevado grado de verosimilitud que caracteriza este movimiento se cifra en gran medida en una reproducción fiel de los emplazamientos en los que se desarrollan las historias, el tratamiento del espacio por parte de los autores es con frecuencia similar desde un punto de vista formal. En último término, será el lector quien decidirá, en función de su propia

concepción de qué es creíble en su entorno, si lo que se representa es verosímil o no. No obstante, estas similitudes revelan una correspondencia nítida entre forma y función que ayuda a configurar la realidad literaria que se nos plantea y que, como se ha podido observar, va más allá de la idiosincrasia estilística de cada autor.

5.2. VALORES ESTILÍSTICOS CONCRETOS: PERSONAJES CON *LAS MANOS EN LOS BOLSILLOS*

En el análisis de bloques textuales también resulta muy útil aislar un ejemplo concreto e investigar su función en nuestro corpus de estudio. En este caso hemos escogido el ejemplo de *las manos en los bolsillos*, perteneciente al grupo de lenguaje corporal. Aunque se trata del grupo menos numeroso con solo tres ejemplos (*las manos en los bolsillos*, *la palma de la mano* y *las manos a la cabeza*), el lenguaje corporal es un aspecto de importancia cardinal en la construcción de los personajes que pueblan los universos ficticios (Korte, 1997). De manera más específica, los *clusters* relacionados con este aspecto «can also be a central part of a description and can highlight habits or behaviour of a character» (Mahlberg, 2007: 25). Los ejemplos localizados en nuestro corpus de novela realista no son una excepción.

Para entender cómo funciona el lenguaje corporal en la definición de los personajes que pueblan un universo ficticio debemos tener en cuenta la relación de continuidad entre el mundo real y los mundos textuales. Para cualquier lector, los personajes cobran sentido sobre modelos basados en nuestra experiencia como seres humanos que *enfrentamos* a la información textual que encontramos en el acto de lectura (Stockwell, 2009). Esta relación entre nuestra experiencia y el plano textual se aplica, con más rigor si cabe, a géneros como la novela realista, caracterizada por el ya comentado alto grado de verosimilitud de la realidad literaria que se nos presenta. La información textual más relevante en el ámbito de la caracterización de los personajes tiene que ver con la forma de relacionarse entre estos, sus reacciones, su manera de hablar y, por supuesto, las descripciones tanto de su apariencia física como de su plano mental. En el caso de la novela realista, un segmento de texto como *las manos en los bolsillos* no solo ofrece información circunstancial sobre la posición del personaje (su apariencia física), sino que sirve para configurar un significado concreto que va más allá de la descripción de apariencia externa y que nos puede ayudar a comprender su estado de ánimo (plano mental). En concreto, en nuestro corpus de novelas *las manos en los bolsillos* actúa como un bloque textual que ayuda a proyectar la tristeza de personajes masculinos. En los ejemplos (6) a (9) se muestran cuatro casos de cuatro novelas diferentes. Como se puede observar, los cuatro ejemplos se asocian a hombres. Este hecho no debe resultar especialmente sorprendente, debido sobre todo a las características de la vestimenta de la época, en la que los bolsillos son propios de prendas mayoritariamente masculinas (pantalones, chalecos, gabardinas, etc.). Sí que resulta más llamativo el contexto que rodea al bloque textual, que refuerza la proyección de la tristeza a la que acabamos de referirnos. Como se puede observar, los cuatro ejemplos son parte de pasajes descriptivos en los que aparece referencias claras a los sentimientos del personaje —«resignados a sufrir el resto de una vida sin esperanza» en (6), «cabizbajo» en (7), «la mirada vagabunda y sin fijeza, como su andar y pensamiento» en (8) o «pálido, ojeroso, cabizbajo» y «con los ojos apagados, tristes» en (9). La sistematicidad con la que *las manos en los bolsillos* forma parte de este tipo de pasajes sugiere un hábito estilístico por parte del autor realista a la hora de proyectar la tristeza de personajes. En otras palabras, *las manos en los bolsillos* actúa como elemento de refuerzo de proyección tristeza mediante la asociación entre un hábito de tipo gestual y un sentimiento muy concreto. Gracias a este uso, en la novela realista este bloque textual adquiere, más allá de la descripción física externa del

personaje, un significado concreto en clave de caracterización también en lo que a la dimensión psicológica del personaje se refiere.

- (6) Desde el fielato se les veía alejarse, *las manos en los bolsillos* y la espalda encorvada, con ademán humilde, resignados a sufrir el resto de una vida sin esperanza y sin sorpresa, conociendo de antemano la fatiga monótona y gris que se extendería hasta el momento de su muerte. (*La horda*, Blasco Ibáñez, capítulo 6)
- (7) Por este arte despotricaba en sus adentros Leto Pérez bajando una mañana hacia el muelle, sin corbata ni chaleco, con una ancha boina en la cabeza y, por todo ropaje exterior, una americanilla y unos pantalones de lienzo. Como arreglaba la marcha al compás de los pensamientos, andaba con relativa lentitud, algo cabizbajo y con *las manos en los bolsillos*. (*Al primer vuelo*, Pereda, capítulo 17)
- (8) Y se marchó despacio, *las manos en los bolsillos*, la gorra encasquetada, la mirada vagabunda y sin fijeza, como su andar y pensamiento. (*La desheredada*, Galdós, capítulo 14)
- (9) Era uno Desperdicios, el segundo era García, el otro era Gabriel... Pero no Gabriel alegre, risueño, con la corbata azul y blanca, colores de la Inmaculada y regalo de su madre; sino Gabriel, pálido, ojeroso, cabizbajo con *las manos en los bolsillos* del pantalón, y en alto el cuello de la levita, porque tenía frío en todo su cuerpo; con los ojos apagados, tristes como las aguas de un lago en que el sol no penetra, porque tenía también frío en toda su alma. (*Colección de lecturas recreativas*, Coloma, capítulo 5)

Este caso es solo un ejemplo de cómo un bloque textual puede analizarse de manera individualizada y descubrir su valor, lo que ofrece una visión más restringida (pero igualmente efectiva) que en el caso de los ejemplos analizados desde una óptica más general en el apartado 5.1.⁷ En cualquier caso, a la luz de los ejemplos que hemos detallado tanto desde una perspectiva más general como con un análisis más localizado, parece claro que los *clusters* identificados en nuestro corpus actúan como bloques textuales que contribuyen a definir los universos ficticios de la novela realista en términos similares, lo que revela hábitos lingüísticos transversales propios de este movimiento literario.

6. CONCLUSIONES

En este artículo hemos empleado una metodología de corpus para identificar la presencia de bloques textuales estilísticamente relevantes en un corpus de novela realista española. Como se ha podido observar, estos bloques textuales, que contribuyen a definir los universos ficticios que los autores nos plantean, van más allá de las idiosincrasias estilísticas de cada novelista. Por desgracia, ni el análisis de bloques textuales ni los enfoques de estilística de corpus en el análisis de textos literarios en español son aspectos que gocen de tradición en términos de literatura científica en el ámbito de la filología hispánica. Esta laguna nos ha obligado a ofrecer un contexto y a detenernos en algunas

⁷ Cabe destacar que las posibilidades que los *clusters* relacionados con el lenguaje corporal ofrecen a los estudiosos del realismo van mucho más allá de las que, por una cuestión de espacio, aquí tratamos a través del ejemplo concreto de *con las manos en los bolsillos*. Para un análisis más completo, convendría localizar *clusters* partiendo de partes del cuerpo escogidas *a priori*, como demuestra Mahlberg (2013: 100-127) en su estudio sobre el lenguaje corporal en la ficción victoriana en general y en las novelas de Charles Dickens en particular. De este modo, sería posible agrupar construcciones en torno a partes del cuerpo concretas (manos, cabeza, hombros, espalda, etc.) y analizar su valor más allá de la mera descripción del lenguaje corporal, como aquí hemos hecho con el ejemplo de *con las manos en los bolsillos*.

cuestiones de tipo metodológico (criterios de localización de ejemplos, por ejemplo) que, sin duda, ha ocupado un espacio de nuestro estudio en detrimento del análisis propiamente dicho. En cualquier caso, tanto la clasificación como el posterior estudio de algunos ejemplos han servido para revelar hábitos de tipo lingüístico que hacen patente una relación entre patrones formales y valores funcionales compartidos por la pléyade de autores que forman nuestro corpus para construir sus universos ficticios.

El objetivo de este artículo no ha sido tanto realizar un análisis sesudo de los ejemplos —inviabile por cuestiones de espacio— como demostrar la existencia de estos patrones y calibrar su relevancia estilística en el marco de la obra de novelistas que pueden englobarse dentro de un mismo movimiento. De manera más específica, los ejemplos localizados sirven para definir un marco lingüístico que puede ayudar a contextualizar el estilo del XIX, lo que puede resultar muy útil para realizar análisis concretos. En este sentido, y en el marco de una investigación de mayor envergadura en la que se sitúa este artículo, el presente estudio define el punto de partida de diferentes vías de análisis a las que pretendemos dar continuidad. En primer lugar, esta clasificación nos puede servir para descubrir qué autores tienen más proclividad a utilizar qué tipo de bloques textuales. En otras palabras, la localización de ejemplos permite realizar un mapa de uso y calibrar, por así decirlo, la distancia estilística entre los autores que forman el corpus. En segundo lugar, aunque en este artículo nos hemos centrado en la novela realista, convendría realizar un estudio similar sobre una muestra de textos de otra época. Aunque resulta difícil compilar un corpus de una extensión parecida al que aquí hemos utilizado con novelas de otra época, una muestra de texto menor podría ser suficiente para identificar diferencias (y similitudes) en la utilización de bloques textuales para construir universos ficticios en distintos períodos de la historia de la novela. Por último, además de trasladarla a otras épocas, esta metodología también podría aplicarse a la producción de autores concretos. Esto es, podría realizarse un análisis similar para identificar funciones textuales locales en la obra de un autor, que podría compararse con el resto de novelistas con los que comparte plana e identificar rasgos idiosincrásicos que quizás hayan pasado desapercibidos para la crítica literaria tradicional —o que al menos no han sido analizados de forma sistemática con anterioridad—.

Estas vías de análisis no son sino tres sugerencias de análisis para el futuro que servirían para dar continuidad al análisis que aquí hemos presentado. Como se ha podido observar, el estudio de bloques textuales utilizando metodologías de corpus atesora un amplio potencial analítico tradicionalmente desatendido del que los estudios literarios se podrían beneficiar en gran medida y que, a la luz de los resultados aquí mostrados, convendría explorar.

REFERENCIAS BIBLIOGRÁFICAS

- Alborg, Juan Luis. 1996. *Historia de la literatura española. Realismo y naturalismo. La novela. Parte primera*. Madrid: Gredos.
- Altenberg, Bengt. 1998. «On the phraseology of spoken English: The evidence of recurrent word-combinations». En *Phraseology. Theory, Analysis, and Applications*, ed. Anthonie Paul Cowie. Oxford: Oxford University Press, 101-122.
- Anthony, Laurence. 2019. *AntConc* (Version 3.5.8) [Software]. Tokyo: Waseda University.
- Archer, Dawm. 2007. «Computer-assisted literary stylistics: The state of the field». En *Contemporary Stylistics*, eds. Marina Lambrou y Peter Stockwell. Londres: Continuum.

- Balossi, Giuseppina. 2014. *A Corpus Linguistic Approach to Literary Language and Characterization: Virginia Woolf's The Waves*. Amsterdam: John Benjamins.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad y Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Biber, Douglas, Susan Conrad, y Viviana Cortes. 2004. «If you look at...: Lexical bundles in university teaching and textbooks». *Applied Linguistics*, 25.3: 371–405.
- Cheng, Winnie. 2012. «Corpus-Based Linguistic Approaches to Critical Discourse Analysis». En *The Encyclopedia of Applied Linguistics*, ed. Carol A. Chapelle. Oxford: Wiley-Blackwell.
- Conrad, Susan, y Douglas Biber. 2005. «The frequency and use of lexical bundles in conversation and academic prose». *The Corpus Approach to Lexicography, Thematischer Teil von Lexicographica. Internationales Jahrbuch für Lexikographie* 20: 56–71.
- Culpeper, Jonathan. 2009. «Keyness: Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's *Romeo and Juliet*». *International Journal of Corpus Linguistics*, 14.1: 29–59.
- Ferreras, Juan Ignacio. 1982. «La generación de 1868». En *Historia y crítica de la literatura española. 5. Romanticismo y realismo*, ed. Iris Zavala. Barcelona: Crítica.
- Fischer-Starcke, Bettina. 2006. «The Phraseology of Jane Austen's *Persuasion*: Phraseological Units as Carriers of Meaning». *ICAME Journal*, 30: 87–104.
- Fischer-Starcke, Bettina. 2010. *Corpus Linguistics in Literary Analysis: Jane Austen and her Contemporaries*. Londres: Continuum.
- Gavins, Joana. 2007. *Text World Theory: An Introduction*. Edimburgo: Edinburgh University Press.
- González, José Ramón. 1998. «Nueva meditación del marco: ciudad y literatura en el fin del siglo». *Ínsula*, 613: 30–33.
- Kopaczyk, Joanna. 2013. *The Legal Language of Scottish Burghs: Standardization and Lexical Bundles (1380–1560)*. Oxford: Oxford University Press.
- Korte, Barbara. 1997. *Body Language in Literature*. Toronto: University of Toronto Press.
- Mahlberg, Michaela. 2007. «Clusters, Key Clusters and Local Textual Functions in Dickens». *Corpora*, 2.1: 1–31.
- Mahlberg, Michaela. 2013. *Corpus Stylistics and Dickens's Fiction*. Londres: Routledge.
- Mahlberg, Michaela. 2014. «Corpus stylistics». En *The Routledge Handbook of Stylistics*, ed. Michael Burke. Londres: Routledge.
- Mahlberg, Michaela, Viola Wiegand, Peter Stockwell, y Anthony Hennessey. 2019. «Speech-bundles in the 19th-century English novel». *Language and Literature*, 28: 326–353.
- Martínez Carazo, Cristina. 2006. *De la visualidad literaria a la visualidad filmica*. La Regenta de Leopoldo Alas Clarín. Gijón: Llibros del Peixe.
- Matzat, Wolfgang. 2007. *Espacios y discursos en la novela española: del realismo a la actualidad*. Frankfurt: Iberoamericana Vervuert.
- McIntyre, Dan y Brian Walker. 2019. *Corpus Stylistics: Theory and Practice*. Edimburgo: Edinburgh University Press.
- Nieto Caballero, Guadalupe. 2019. «El espacio como eje vertebrador en la creación del universo ficticio galdosiano: un estudio de corpus». *SIGNA*, 28: 1203–1238.
- Pedraza Jiménez, Felipe B. y Milagros Rodríguez Cáceres. 1983. *Manual de literatura española. Tomo VII: Época del Realismo*. Madrid: Cénlit Ediciones.
- Pedraza Jiménez, Felipe B. y Milagros Rodríguez Cáceres. 2012. *Las épocas de la*

- literatura española*. Barcelona: Ariel.
- Rodríguez Marín, Rafael. 2005. *Metalengua y variación lingüística en la novela de la Restauración decimonónica*. Madrid: Real Academia Española.
- Ruano San Segundo, Pablo. 2016. «A corpus-stylistic approach to Dickens' use of speech verbs: Beyond mere reporting». *Language and Literature*, 25.2: 1–15.
- Scott, Mike. 2013. *WordSmith Tools Manual. Version 6.0*. Liverpool: Lexical Analysis Software.
- Scott, Mike. 2016. *WordSmith Tools version 7*. Stroud: Lexical Analysis Software.
- Scott, Mike. 2019. «Single words v. clusters». 1 de junio de 2020. <https://lexically.net/downloads/version7/HTML/single_words.html>.
- Short, Mick, y Elena Semino. 2008. «Evaluation and Stylistic Analysis». En *The Quality of Literature: Linguistic Studies in Literary Evaluation*, ed. Willie van Peer. Amsterdam: John Benjamins.
- Stockwell, Peter. 2009. *Texture: A Cognitive Aesthetics of Reading*, Edinburgh, University Press.
- Stubbs, Michael. 2005. «Conrad in the computer: Examples of quantitative stylistics methods». *Language and Literature*, 14.1: 5–24.
- Stubbs, Michael e Isabel Barth. 2003. «Using recurrent phrases as text-type discriminators. A quantitative method and some findings». *Functions of Language*, 10.1: 61–104.
- Tognini-Bonelli, 2001. Elena. *Corpus Linguistics at Work*. Ámsterdam: John Benjamins.
- Werth, Paul. 1999. *Text Worlds: Representing Conceptual Space in Discourse*. Londres: Longman.
- Zavala, Iris (ed.). 1982. *Historia y crítica de la literatura española. 5. Romanticismo y realismo*. Barcelona: Crítica.
- Zubiaurre, María Teresa. *El espacio en la novela realista*. México: Fondo de cultura económica, 2000.

Apéndice

Autor	Novela	Año	Palabras
Alarcón	<i>El final de Norma</i>	1855	41 892
	<i>El sombrero de tres picos</i>	1874	24 909
	<i>El escándalo</i>	1875	78 583
	<i>El Capitán Veneno</i>	1881	40 655
Blasco Ibáñez	<i>Arroz y tartana</i>	1894	91 328
	<i>La barraca</i>	1898	54 575
	<i>Entre naranjos</i>	1900	80 873
	<i>Cañas y barro</i>	1902	4913
	<i>La catedral</i>	1903	98 430
	<i>El intruso</i>	1904	96 969
	<i>La horda</i>	1905	101 600
	<i>Los argonautas</i>	1914	205 215
	<i>Los cuatro jinetes del Apocalipsis</i>	1916	132 439
	<i>Los enemigos de la mujer</i>	1919	163 461
	<i>El paraíso de las mujeres</i>	1922	90 540
<i>El fantasma de las alas de oro</i>	1930	70 788	
Clarín	<i>La Regenta</i>	1884-85	309 796
	<i>Su único hijo</i>	1890	89 969
Coloma	<i>La almohadita del niño Jesús</i>	1887	21 753
	<i>Pequeñeces</i>	1890-91	20 246
	<i>Cuentos para niños</i>	1890	17 095
	<i>La reina mártir</i>	1898	77 077
	<i>Jeromín</i>	1902	146 177
	<i>Colección de lecturas recreativas</i>	1902	216 807
	<i>Por un piojo...: cuadro de costumbres</i>	1912	26 915
Palacio Valdés	<i>El idilio de un enfermo</i>	1884	58 113
	<i>La aldea perdida</i>	1903	107 153

	<i>La novela de un novelista</i>	1921	103 561
	<i>Sinfonía pastoral</i>	1931	107 154
Pardo Bazán	<i>Pascual López: autobiografía de un estudiante de Medicina</i>	1879	62 435
	<i>Un viaje de novios</i>	1881	69 014
	<i>La tribuna</i>	1883	65 160
	<i>El Cisne de Vilamorta</i>	1885	53 123
	<i>Los pazos de Ulloa</i>	1886-87	84 000
	<i>La madre naturaleza</i>	1887	90 463
	<i>Insolación (Historia amorosa)</i>	1889	42 849
	<i>Una cristiana</i>	1890	61 242
	<i>La prueba</i>	1890	65 040
	<i>La piedra angular</i>	1891	65 040
	<i>Doña Milagros</i>	1894	67 826
	<i>Memorias de un solterón</i>	1896	58 184
	<i>El tesoro de Gastón</i>	1897	33 367
	<i>El saludo de las brujas</i>	1899	60 078
	<i>El niño de Guzmán</i>	1900	31 830
	<i>Misterio</i>	1902	107 472
	<i>La quimera</i>	1905	155 430
	<i>La sirena negra</i>	1908	40 553
	<i>Dulce dueño</i>	1911	66 499
<i>La última fada</i>	1916	15 011	
Pereda	<i>Escenas montañosas</i>	1864	89 339
	<i>Tipos y paisajes</i>	1871	100 753
	<i>Don Gonzalo González de la Gonzalera</i>	1878	88 297
	<i>El buey suelto...</i>	1878	72 537
	<i>De tal palo tal astilla</i>	1880	85 974
	<i>El sabor de la tierra</i>	1882	81 756
	<i>Pedro Sánchez</i>	1883	110 193

	<i>Sotileza</i>	1885	117 441
	<i>Esbozos y rasguños</i>	1888	70 323
	<i>La Montálvez</i>	1888	106 115
	<i>La mujer del César</i>	1888	25 179
	<i>Los hombres de pro</i>	1888	45 060
	<i>Tipos trashumantes</i>	1888	29 631
	<i>Nubes de estío</i>	1891	121 738
	<i>Al primer vuelo: Idilio vulgar.</i>	1891	100 914
	<i>Oros son triunfos</i>	1894	27 233
	<i>Peñas arriba</i>	1895	134 976
	<i>Pachín González</i>	1896	60 309
	<i>Escritos de juventud</i>	1954 ⁸	82 726
Perez Galdós	<i>Trafalgar</i>	1873	51 453
	<i>Zumalacárregui</i>	1898	71 013
	<i>España sin Rey</i>	1907-08	78 784
	<i>La Fontana de Oro</i>	1870	125 722
	<i>La desheredada</i>	1881	141 130
	<i>Mendizábal</i>	1898	83 240
	<i>El amigo manso</i>	1882	90 703
	<i>El doctor Centeno</i>	1883	113 949
	<i>Bailén</i>	1873	63 819
	<i>Doña Perfecta</i>	1876	65 486
	<i>O'Donnell</i>	1904	78 986
	<i>Gloria</i>	1876-77	127 697
	<i>La de Bringas</i>	1884	71 862
	<i>Marianela</i>	1878	51 234
	<i>Lo prohibido</i>	1884-85	172 610
<i>Fortunata y Jacinta</i>	1886-87	379 607	

⁸ Aunque su escritura fue anterior (entre 1858 y 1879), la edición que manejamos es de ese año.

	<i>Un voluntario realista</i>	1878	66 867
	<i>Miau</i>	1888	97 628
	<i>Halma</i>	1895	76 162
	<i>Misericordia</i>	1897	84 131
Valera	<i>Pepita Jiménez</i>	1874	56 575
	<i>Las ilusiones del doctor Faustino</i>	1875	119 890
	<i>Pasarse de listo</i>	1878	51 851
	<i>Morsamor</i>	1899	84 574
	<i>Juanita la Larga</i>	1895	73 329
	<i>Genio y figura</i>	1897	53 358
	<i>Elisa la malagueña</i>	1964 ⁹	11 844
	<i>Leyendas del antiguo oriente</i>	1964	35 652
	<i>Mariquita y Antonio</i>	1964	50 824
	<i>Doña Luz</i>	1879	59 517
	<i>El comendador Mendoza</i>	1876	63 355
Total palabras			8 156 998

⁹ *Elisa la malagueña*, *Leyendas del antiguo oriente* y *Mariquita y Antonio* fueron publicadas antes de la fecha indicada: 1896, 1870 y 1861, respectivamente.