

Guadalupe Nieto Caballero y
Pablo Ruano San Segundo

Estilística de corpus:
Nuevos enfoques en
el análisis de textos literarios



PETER LANG

Bern · Berlin · Bruxelles · New York · Oxford · Warszawa

Tabla de contenidos

Índice de tablas, ilustraciones y gráficos	9
Introducción	19
Parte I: Teoría	
Capítulo 1. Estilística de corpus	25
1.1. Hacia una definición de estilística de corpus	25
1.1.1. Qué ofrece la lingüística de corpus a la estilística de corpus	26
1.1.2. Qué ofrece la estilística tradicional a la estilística de corpus	31
1.1.3. Qué es la estilística de corpus	36
1.2. La estilística de corpus en el marco de las Humanidades Digitales	37
1.2.1. Antecedentes teóricos	38
1.2.2. El nacimiento, el desarrollo y la consolidación de la estilística de corpus	42
1.2.3. La estilística de corpus en el ámbito hispanico	47
1.3. Puntos de partida de análisis	51
1.3.1. Universos ficticios	53
1.3.2. Caracterización	55
1.4. Una nota de precaución	58
1.5. Consideraciones finales	63
Capítulo 2. El corpus	67
2.1. Opciones de compilación: de la digitalización manual a los repositorios digitales	68
2.1.1. Biblioteca Virtual Miguel de Cervantes	70
2.1.2. Proyecto Gutenberg	71
2.1.3. Más repositorios y otros sitios web	72

2.2. Principios para la elaboración de un corpus de estudio	75
2.2.1. Nuestro corpus de estudio	78
2.3. Principios para la elaboración de un corpus de referencia	80
2.3.1. Nuestro corpus de referencia	82
2.4. Consideraciones legales	85
2.4.1. Derechos de autor	86
2.4.2. Excepciones en los derechos de autor	88
Capítulo 3. Herramientas para el procesamiento de textos	91
3.1. Herramientas integradas en sitios web	92
3.1.1. CLiC (Corpus Linguistics in Cheshire)	94
3.1.2. Real Academia Española y Biblioteca Virtual Miguel de Cervantes	98
3.2. Herramientas independientes	102
3.2.1. <i>WordSmith Tools</i>	104
3.2.1.1. Listas de palabras	106
3.2.1.2. Concordancias	108
3.2.1.3. Palabras clave	110
3.2.1.4. <i>Clusters</i>	114
3.3. Anotación de un corpus	117
Parte II: Práctica	
Capítulo 4. Análisis estilístico de la producción narrativa de Galdós	127
4.1. La proyección del discurso en la narrativa de Galdós	128
4.1.1. Bloques textuales y construcción del universo ficticio	131
4.1.2. Funciones textuales locales: caracterización	134

4.2. La narración de los silencios	138
4.2.1. La narración del silencio en el género novelesco	138
4.2.2. Localización de ejemplos en CorBPG y CorXIX	139
4.2.3. Análisis de la narración retrospectiva de las pausas en CorBPG	142
4.3. Evolución en el uso de estrategias de representación del discurso	145
4.3.1. La evolución de Galdós	146
4.3.2. Análisis	149
4.3.2.1. Fase inicial y época de madurez	149
4.3.2.2. Estilo directo y construcción del universo ficticio	152
4.4. Consideraciones finales	156
Capítulo 5. Intertextualidad: concomitancias estilísticas en Dickens y Galdós	159
5.1. La influencia de Dickens en Galdós	160
5.2. Localización de ejemplos en CorBPG y CorCD	162
5.2.1. Resultados	165
5.3. Análisis	168
5.3.1. <i>Ojos</i>	169
5.3.2. <i>Manos</i>	172
5.3.3. <i>Cabeza</i>	174
5.3.4. <i>Hombros</i>	176
5.4. Consideraciones finales	177
Capítulo 6. El Galdós traductor	181
6.1. Galdós y su única traducción: <i>Las aventuras de Pickwick</i>	182
6.2. Charles Dickens y los verbos de habla	184
6.3. Localización de ejemplos en <i>Pickwick Papers</i> , de Charles Dickens	186

6.4. Localización de ejemplos en <i>Las aventuras de Pickwick</i> , de Galdós	189
6.5. Análisis de ejemplos	192
6.5.1. Omisión del verbo de habla	193
6.5.2. Utilización sistemática de “decir”	196
6.5.3. Traducciones libres	199
6.5.4. Errores de traducción	202
6.6. Consideraciones finales	204
Conclusiones: dónde estamos y hacia dónde podemos dirigirnos	207
Apéndices	211
Bibliografía	217

Introducción

Estilística de corpus: Nuevos enfoques en el análisis de textos literarios es un libro en el que pretendemos acercar al lector técnicas de análisis relativamente novedades en el marco de los estudios literarios. Los enfoques de estilística de corpus se caracterizan por emplear metodologías propias de la lingüística de corpus en conjunción con acercamientos más propios de la estilística tradicional y apoyándose, como no puede ser de otra manera, en el contexto de la crítica literaria. La dimensión cuantitativa de estos estudios ha abierto nuevas vías de análisis y ha permitido, en el ámbito internacional, explorar el estilo de autores manejando un volumen de texto inasumible para disciplinas más tradicionales. El reconocimiento —y podríamos decir que cierto recorrido, a pesar de su corta vida— de los estudios de estilística de corpus en el ámbito internacional no se corresponde, sin embargo, con una réplica en el caso de los estudios literarios de textos en lengua española, donde la estilística de corpus constituye una disciplina aún por descubrir. Este camino aún por transitar en el mundo hispánico no solo justifica sino que parece demandar un libro como el que aquí se presenta, en donde el lector pueda familiarizarse con algunos de los presupuestos teóricos y metodológicos fundamentales de esta disciplina con los que poder implementar análisis cabales de textos literarios en nuestra lengua. Además de una explicación pormenorizada de los elementos más importantes desde una perspectiva de instrucción teórica (la construcción de un corpus, herramientas para el procesamiento de textos o puntos de partida de análisis, entre otros aspectos), en este libro el lector también encontrará un análisis práctico sobre la figura de Galdós, en un decidido intento de ilustrar el potencial de este tipo de análisis en el estudio de una figura de referencia de nuestro canon literario.

Antes de explicar los contenidos del libro, nos gustaría hacer sendas aclaraciones sobre dos aspectos fundamentales para el lector. En primer lugar, hemos intentado huir de dogmatismos y posturas excluyentes. Dicho de otro modo, no hemos diseñado el libro como una lectura que solo pueda entender gente con formación especializada en lingüística de corpus ni para la que haya que tener experiencia en el campo de las Humanidades Digitales. Esta actitud no responde tanto a un presupuesto eclecticista como al convencimiento de que un libro de estas características puede ser útil a lectores con dos perfiles bien distintos. De hecho, consideramos que

el libro puede ofrecer puntos de interés a lectores diametralmente opuestos. Así, aquellos que se acerquen al libro desde la estilística tradicional podrán encontrar información sobre los estudios de corpus en general y cómo poner en práctica su aplicación, precisamente, en estudios de tipo estilístico, añadiéndolos a su panoplia de técnicas de análisis estilísticos que manejan habitualmente. Por otra parte, quien se acerque al libro con una formación previa en estudios de corpus, encontrará en la aplicación literaria una nueva manera de aproximarse al texto objeto de estudio, donde la obra literaria no esté considerada simplemente como una muestra más de la lengua, sino como un género en sí mismo susceptible de análisis crítico. En el espectro entre estos dos polos, además, es fácil imaginar lectores menos especializados que también puedan encontrar en la estilística de corpus novedosas vías de análisis en el ámbito de las Humanidades Digitales.

En segundo lugar, a lo largo del libro no situaremos en ningún caso a la estilística de corpus como un enfoque que pretenda reemplazar a otros más tradicionales, sino como una forma de análisis que puede complementarse con otras disciplinas como la crítica literaria o la estilística tradicional, del que el análisis de textos literarios en general se puede beneficiar en gran medida. El tono del libro, aunque pueda (y deba) resultar obvio, es conciliador. De hecho, somos conscientes de que, precisamente por su carácter novel, al acercarse a la estilística de corpus uno tiene que tener en cuenta “its lack of a clear definition and its as yet not fully developed analytical methods” (McIntyre y Walker 2019: 1). A pesar de ello, sin embargo, estamos plenamente convencidos del potencial analítico de la estilística de corpus y de que, en ocasiones, estos enfoques pueden resultar el modo más eficaz de investigar aspectos que, bien por su numerosidad bien por su carácter recóndito en el texto, han pasado desapercibidos para la crítica tradicional.

En el transcurso de su lectura, en el libro se podrán encontrar distintos focos de interés. El primero y más importante, desde luego, es la definición de la estilística de corpus. Esta definición se acompañará de una explicación previa de los principales campos de estudio de los que se nutre (la lingüística de corpus y la estilística tradicional), de una contextualización en el marco de la Humanidades Digitales, de sus antecedentes teóricos y de enfoques computacionales previos como la estilometría o el análisis estadístico de textos. También se ofrecerán explicaciones detenidas sobre los puntos de partida sobre los que se suelen construir los análisis, como

la caracterización de personajes o la construcción de los universos ficticios que los autores nos plantean, al tiempo que se detallarán notas de precaución sobre algunos de los errores habituales cometidos por aquellos que ven en la estilística de corpus un enfoque exclusivamente cuantitativo. De todos estos aspectos nos ocuparemos en el primer capítulo.

También abordaremos cuestiones relacionadas con qué es y cómo se debe construir un corpus de estudio en un análisis de estilística de corpus (capítulo 2) y con las herramientas que se emplean para procesarlos (capítulo 3). Las cuestiones relacionadas con la compilación de un corpus serán de tipo general (extensión del corpus, formato de los textos, repositorios digitales, etc.) y también de índole más específica (cuestiones de derechos de autor, por ejemplo). En cuanto a las herramientas de análisis, se ofrecerán ejemplos sobre las distintas opciones a disposición del usuario, estableciendo una distinción entre aquellas integradas en sitios web de bancos de datos, por un lado, y los *softwares* autónomos especializados, por otro. Por su importancia, en el caso de estos últimos explicaremos sus principales opciones de análisis, como las listas de palabras, las concordancias o las palabras clave.

Tras esta explicación de los presupuestos teóricos y metodológicos fundamentales de la estilística de corpus y los aspectos sobre los que se cimienta, en la segunda parte del libro llevamos a cabo un análisis práctico en torno a la figura de Benito Pérez Galdós. La razón para escoger a Galdós como objeto de análisis tiene que ver con que se trata de uno de los novelistas españoles más laureados de siempre. Se trata, además, de un escritor muy prolijo, lo que le convierte en un autor ideal para ser analizado con herramientas de corpus, pues cabe esperar que estas revelen patrones tanto formales como funcionales que hasta ahora han pasado desapercibidos para la crítica y que ayuden a una mejor comprensión de su estilo. Además, pertenece una de las épocas doradas de la novela española: el Realismo. Este hecho hace posible que su obra pueda compararse con la de sus contemporáneos, lo que permitirá calibrar la representatividad de los rasgos que se analizan en relación a la pléyade de autores con los que compartió plana. En el análisis de la obra de Galdós se abordarán distintas dimensiones de su figura: su estilo como novelista (capítulo 4), elementos de intertextualidad relacionados con la influencia que Dickens ejerció sobre él (capítulo 5) y su faceta menos conocida como traductor (capítulo 6). El objetivo de esta segunda parte del libro es ilustrar cómo todo lo explicado en la primera mitad puede ponerse en práctica desde

distintos ángulos de análisis y demostrar que es posible analizar aspectos del estilo de Galdós que, debido a la falta de herramientas adecuadas, no han podido ser identificados con anterioridad —o al menos no han podido ser analizados de forma sistemática.

Esta doble vertiente por la que discurre el libro pretende ofrecer, tanto desde la teoría como desde la praxis, argumentos razonados a favor del uso de este tipo de enfoques en el análisis de textos literarios en español. Desde luego, el hecho de que la estilística de corpus aún no se haya consolidado en el ámbito hispánico y los estudios que adopten este tipo de enfoques sea prácticamente inexistente puede explicarse, al menos parcialmente, por el carácter refractario de una parte de la crítica literaria tradicional, recelosa del componente cuantitativo de las metodologías computacionales en el análisis estilístico de una obra literaria. Estos prejuicios, plenamente superados en el ámbito internacional, parecen estar vigentes entre algunos estudiosos de la literatura en español. Aquí intentaremos combatirlos y demostrar que no hay razones para no comenzar a utilizar este tipo de enfoques con normalidad, pues constituyen un valioso complemento para llevar a cabo análisis estilísticos de textos literarios en nuestra lengua. *Estilística de corpus: Nuevos enfoques en el análisis de textos literarios*, en suma, pretende no solo ser un manual de introducción y una fuente a la que acudir para aquellos que no estén familiarizados con este tipo de enfoques, sino un ejemplo que sirva para demostrar, en el ámbito hispánico, que tenemos a nuestra disposición opciones de análisis de las que podemos beneficiarnos y que merece la pena empezar a explorar.