

ANÁLISIS CUANTITATIVO DE LOS RASGOS LÉXICO-SEMÁNTICOS PROPIOS DE LA NOVELA GÓTICA

JUAN GARCÍA IBORRA
Universidad de Murcia

RESUMEN. El presente trabajo pretende indagar sobre el grado de validez del análisis multidimensional de variación estilística en la creación de un modelo estadístico que nos permita la identificación y discriminación de textos literarios pertenecientes a un mismo subgénero literario o modo narrativo, en este caso la novela gótica. Para este fin estableceremos una comparación de los rasgos o componentes léxico-semánticos de este modo narrativo según el modelo estadístico creado y los descritos por la crítica de la ficción gótica.

La temática de investigación escogida está a caballo entre la literatura y la lingüística del corpus, se trata de un área de investigación interdisciplinaria conocida como *computación literaria (literary computing)*, área en la que comienzan a producirse ciertos estudios de importancia sobre el gótico (Harris 1998).

ABSTRACT. Our present work aims at being a brief investigation about the validity degree of the stylistic variation multidimensional analysis, and its subsequent setting up of a statistical model which allows us to identify and discriminate literary texts of the same genre or narrative mode, being in this paper the gothic novel. To achieve this goal we will establish a comparison, on one hand, between the different characteristics, features, or lexical components of this narrative mode according to the statistical model created and, on the other hand, those features described by the gothic fiction critical experts.

We have chosen an interdisciplinary thematic area of investigation situated between corpus linguistics and literature known as *literary computing* in which several important works have appeared recently (Harris 1998).

1. INTRODUCCIÓN

El presente trabajo pretende indagar sobre el grado de validez del análisis multidimensional de variación estilística en la creación de un modelo estadístico que nos permita la identificación y discriminación de textos literarios pertenecientes a un mismo subgénero literario o modo narrativo, en este caso la novela gótica. Para este fin estableceremos una comparación de los rasgos o componentes léxico-semánticos de este modo narrativo según el modelo estadístico creado y los descritos por la crítica de la ficción gótica.

La temática de investigación escogida está a caballo entre la literatura y la lingüística del corpus, se trata de un área de investigación interdisciplinaria conocida como *computación literaria (literary computing)*, área en la que comienzan a producirse ciertos estudios de importancia sobre el gótico (Harris 1998).

Al supuesto de utilidad del análisis estadístico multivariante para modelar formalmente los componentes léxico-semánticos de un subgénero literario o modo narrativo, subyace la hipótesis de partida de que estos rasgos lingüísticos son susceptibles de ser modelados estadísticamente mediante técnicas de reducción de variabilidad entre variables, conocidas también como multivariantes. Así pues, nuestra hipótesis inicial queda

formulada como que los componentes léxico-semánticos del subgénero gótico inglés pueden ser modelados con técnicas estadísticas multivariantes. A diferencia de los métodos univariantes y bivariantes, las técnicas multivariantes estudian la incidencia conjunta o interacción de más de una (univariante) o dos (bivariante) variables.

La metodología empleada parte de la base teórica propuesta por Biber (1988), Biber y Finegan (1994) y Conrad y Biber (2001) en la que se fundamentan los rasgos de variación estilística, por un lado, y, por otro, los rasgos descritos en la bibliografía gótica (Aguirre 1990, Howard 1994 y Punter 1996).

La muestra empírica utilizada en este trabajo son cinco novelas digitalizadas pertenecientes al mismo modo narrativo y al denominado periodo clásico de la novela gótica inglesa. Los textos son representativos de este subgénero, pues pertenecen a los autores que marcaron la evolución del género durante este periodo de la literatura inglesa desde su aparición en 1764 con *The Castle of Otranto* de Horace Walpole hasta su declive en 1820 con la obra *Melmoth the Wanderer* de Charles R. Maturin.

El procedimiento empleado para la consecución de los resultados pasa primeramente por analizar e identificar los rasgos estilísticos, en particular los léxico-semánticos propios y genuinos de la novela gótica. Para este fin hemos creído necesario tener en cuenta la evolución diacrónica del gótico y variaciones de las categorías léxicas del gótico propuestas. Asimismo determinamos las categorías léxico-semánticas del gótico que utilizamos en este trabajo.

Posteriormente realizaremos el estudio cuantitativo de los rasgos léxico-semánticos establecidos en las obras seleccionadas mediante el programa de gestión textual *MonoConc* (Barlow 1996), el cual nos mostrará una tabulación cuantitativa de los datos. Dicha tabulación, en un posterior análisis con el programa SPSS v.11 (Social Package for Social Sciences), identificará mediante el análisis factorial aquellos rasgos léxico-semánticos más pertinentes para nuestro estudio. De este modo, conoceremos la semejanza o divergencia que exista entre los textos góticos estudiados.

Finalmente evaluaremos los resultados y validaremos el método. Observaremos la evolución de la ficción gótica desde una visión histórica del género con el fin de vislumbrar las distintas etapas en las que se diversificó el género. Esto nos proporcionará datos contrastables con el resultado del análisis factorial del corpus analizado en tanto en cuanto el estudio cuantitativo verifica en cierta medida las teorías literarias de la evolución del gótico.

La novela gótica ocupa un papel único en la historia de la novela inglesa. Sirvió de nexo importante entre los grandes escritores del dieciocho y los grandes maestros de la novela de principios del diecinueve como Scott, Austen o Dickens. Influyó positivamente a escritores de la talla de Austen, Scott, Byron, Keats, Shelley y las hermanas Brontë en el XIX (Briggs 1977: 47; Punter 1996: 187) y también en el XX (Punter 1980: 373-401; Botting 1996: 155-180). En este espacio temporal sobresalen una serie de escritores como Walpole, Lewis, Radcliffe, Godwin, Beckford, Inchbald, Mackenzie, Mary Shelley, Maturin y Charlotte Smith.

El tiempo durante el cual transcurre la novela gótica es otra dificultad añadida a la hora de definir el desarrollo del género. Entre la primera novela gótica, *The Castle of Otranto* (1764) y las últimas de Mary Shelley, *Frankenstein* (1818) y de C.Maturin, *Melmoth the Wanderer* (1820) transcurre medio siglo. Es importante apuntar también que este modo narrativo sufrió una considerable evolución durante el medio siglo que pervivió como género de ficción. Se distinguen principalmente tres tipos o canales paralelos del desarrollo de la ficción gótica (Summers 1938; Varma 1957; Botting 1996) que divergen de la novela de Walpole:

- gótico sobrenatural
- gótico terrorífico
- gótico horrorífico.

Estos tres tipos o canales no se originaron al unísono, sino que, partiendo de la novela Walpole, evolucionaron hasta que, a mediados de 1790, observamos la coexistencia de los tres tipos hasta el declive del género, que según la crítica se produce con *Melmoth the Wanderer* (1820).

A pesar de que la bibliografía especializada del gótico como modo narrativo (Punter 1996) o subgénero literario (Kilgour 1995) es muy amplia, no existen, sin embargo, estudios filológicos específicos del gótico, con lo cual pensamos que este trabajo, quizás, podría contribuir en parte a paliar esta deficiencia de los estudios sobre el gótico. Existen referencias generales sobre los aspectos más predominantes del gótico partiendo de la temática o las convenciones (Sedgewick 1986, Sage 1991) y, últimamente, partiendo de aspectos ideológicos, sociales y políticos (Punter 2000 y Norton 2000).

Robert Harris, profesor de la Universidad de Riverside en California, especialista en el uso de la informática y *software* en el lenguaje y estudios literarios, desarrolló en 1998 un proyecto en Internet¹ titulado “Elements of the Gothic Novel”, en el que analiza los elementos de la novela gótica inglesa clásica y distingue diez apartados, siendo el décimo, las categorías léxico-semánticas, las que describirían totalmente o parcialmente los restantes elementos o convenciones góticas.

Harris establece seis categorías léxico-semánticas de las que partimos en nuestro estudio. Estas categorías abarcan el léxico específico del gótico en la mayoría de la obras, aunque como observaremos, no todas las categorías establecidas se muestran con la misma intensidad en todas las obras del gótico. También es cierto que se podrían haber establecido más categorías, pues en algunas obras hay elementos como el sexo, que sobresalen especialmente, vease el caso de *The Monk*, pero esta misma categoría no aparece en absoluto en obras como *Vathek* o *Melmoth the Wnderer*.

Así, y siguiendo las categorías delimitadas por Harris, nos centraremos en:

- a. Lo sobrenatural, misterio (Supernatural/Mystery).
- b. El miedo, terror, horror (Fear/Terror/Horror).
- c. La sorpresa (Surprise)
- d. La prisa (Haste)
- e. La ira (Wrath)
- f. La magnitud (Largeness)

2. MÉTODO

Compilamos por medio de Internet todas las bases de datos de textos literarios existentes *on-line*² para compilar un corpus de textos digitalizados.

Para adecuar el estudio a la amplitud de esta comunicación redujimos el corpus de 11 a 5 textos (Ver Cuadro 1), centrándonos en los más representativos del gótico clásico y ateniéndonos al factor cronológico en la evolución del gótico (Varma 1957). La primera obra representa el comienzo del gótico (*The Castle of Otranto*); la segunda (*Vathek*); y la tercera son obras que simbolizan su auge (*The Monk*); la cuarta (*Frankenstein*) y la quinta (*Melmoth the Wanderer*) se enmarcan dentro de las obras escritas en la decadencia de este modo narrativo. Así, las obras seleccionadas son en sí mismas hitos importantes en la historia de este modo narrativo.

Una vez seleccionadas las obras, procedimos a la identificación de los rasgos léxico-semánticos, antes apuntados. Debido al ingente trabajo que nos hubiera exigido marcar cada una de las palabras relacionadas con las categorías léxicas de nuestro estudio, decidimos optar por un método de catas al azar. Para ello, marcamos manualmente todas las voces pertenecientes a las categorías léxicas en una de las obras: *The Castle of Otranto* (Ver Cuadro 2). Seguidamente seleccionamos un total de catas al azar de la misma obra, cuyo número total de palabras estuviese en torno al 10% del total de palabras de la obra. La comparación de las catas reales frente a las catas nos permitiría, mediante una proyección (regresión lineal), conocer el grado de ajuste de las catas a la realidad de la novela y, a su vez, validar o rechazar el método empleado.

NOVELA	NÚMERO DE PALABRAS
<i>The Castle of Otranto</i> (1764) = Otranto	36.488
<i>The Monk</i> (1796) = Monk	138.551
<i>Melmoth the Wanderer</i> (1820) = Melmoth	14.080
TOTAL PALABRAS	294.959

Cuadro 1. Relación cronológica de novelas góticas

CATEGORIA	ITEMS
Miedo	afraid, diabolical, doom, enchantment, ghost, goblins, haunted, infernal, magic, magician, miracle, necromancer, omens, ominous, portent, preternatural, prodigy, prophecy, secret, sorcerer, spectre, spirits, strangeness, tomb, vision
Terror	afflicted, affliction, agony, anguish, apprehensions, apprehensive, blood, bloody, commiseration, despair, dismal, dismay, dread, dreaded, dreading, fatal, fearing, frantic, fright, frightened, grief, hopeless, horrid, horror, lamentable, melancholy, miserable, mournfully, outrage, panic, sadly, scared, shrieks, sorrow, sympathy, tears, terrible, terrified, terror, unhappy, wretched
Sorpresa	alarm, amazement, astonished, astonishment, shocking, staring, surprise, surprised, thunderstruck, wonder
Prisa	anxious, breathless, flight, frantic, hastened, hastily, impatience, impatient, impatiently, impetuosity, precipitately, running, sudden, suddenly
Ira	abhor, anger, angrily, cholera, enraged, furious, fury, incense, incensed, provoked, rage, raving, resentment, temper, wrath, wrathful, wrathfully
Magnitud	Enormous, gigantic, giant, large, tremendous, vast

Cuadro 2. *Items léxicos de las categorías de la novela gótica*

mediante el programa de gestión textual *MonoConc* (Barlow 1996), nos mostrará una tabulación cuantitativa de los datos. Mediante dicha tabulación identificaremos los rasgos léxico-semánticos más pertinentes para nuestro estudio. Así podremos establecer la semejanza o divergencia que exista entre los textos góticos analizados.

3. RESULTADOS

Comprobamos que ahora, efectivamente, los componentes seleccionados sí aportan a la composición del factor léxico-semántico. Este modelo final tiene un factor con cuatro componentes que resumen y explican las seis variables o categorías léxico-semánticas de la novela gótica. Sobre este factor tiene especial incidencia el misterio (que explica el 93,5 % de toda la varianza), seguido del miedo (5,2 %); mientras que la aportación de la sorpresa y la prisa son claramente menores (inferior al 1 %). Esto debe entenderse como que la selección léxica en la novela gótica está en primer lugar, relacionada con el misterio que, a su vez, conlleva el miedo y estos dos, en menor medida, contribuyen a la aparición de rasgos léxico-semánticos referidos a la sorpresa y la prisa.

Según nuestro análisis factorial, (ver cuadro 3) no parece que la ira, a pesar de su notable presencia, ni la magnitud, categoría que en los datos descriptivos siempre ha aparecido como claramente menor, sean componentes claros de los rasgos más genuinos de las novelas góticas. Además, su aparición no va necesariamente ligada a los otros cuatro componentes del único factor, ya que no “aportan” a la explicación del factor.

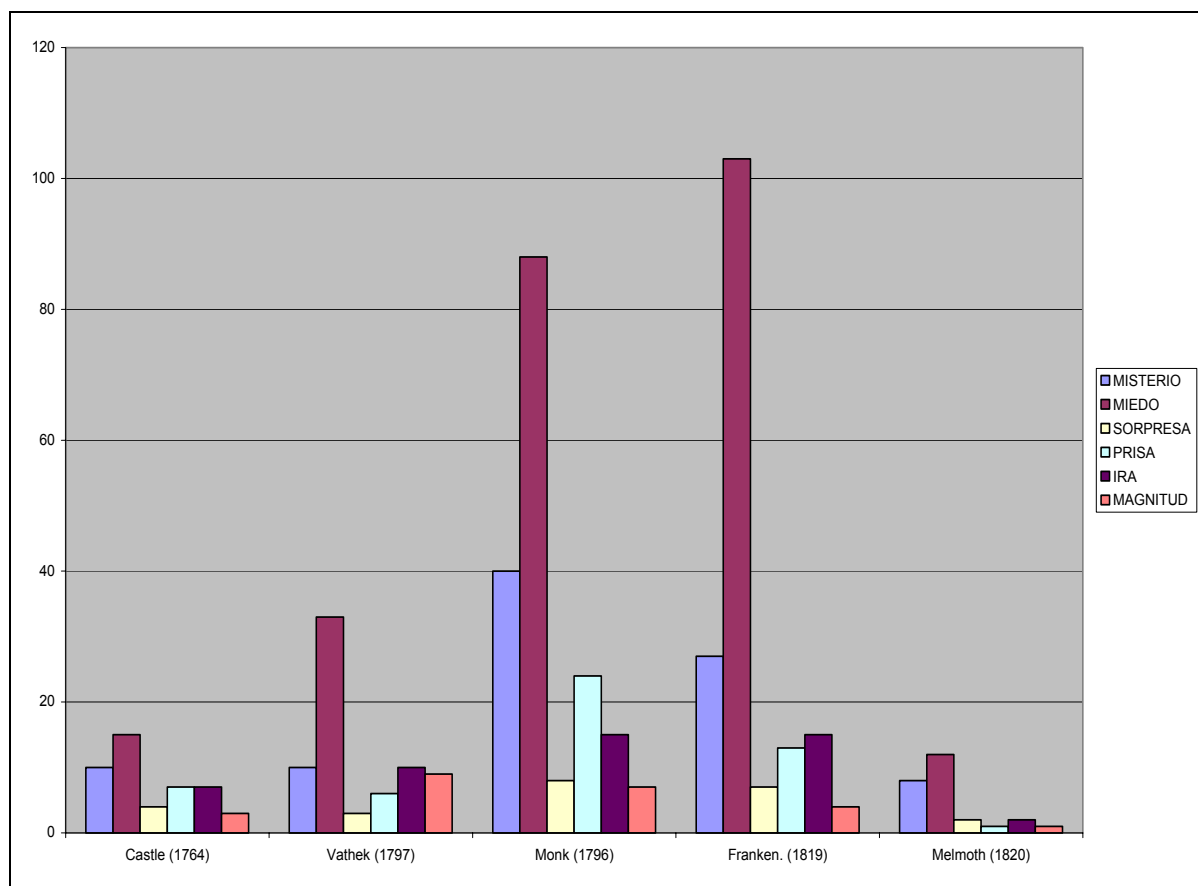
	Componente
	1
MISTERIO	,957
MIEDO	,932
SORPRESA	,968
PRISA	,958
IRA	,755
MAGNITUD	,506

Cuadro 3. *Análisis factorial. Matriz de componentes*

Así pues, sabemos que las categorías léxico-semánticas en la novela gótica pueden reducirse a un único componente (ver Cuadro 3).

Nuestra evaluación de los resultados y conclusiones tiene dos vertientes: 1) por un lado las implicaciones del modelo propuesto respecto a los resultados del análisis descriptivo en la novela gótica y por otro 2) las de los resultados del análisis multivariante en cuanto a las categorías propuestas descriptoras del gótico.

Si observamos el Cuadro 4 advertimos a primera vista una evolución ascendente cronológica de las ocurrencias en las distintas categorías que hemos propuesto para la definición del subgénero gótico. Esta progresión ascendente corroboraría nuestra hipótesis inicial de que un modelado estadístico pudiera servir en la identificación y discriminación de textos góticos, que por sus particulares características al definirse por determinadas convenciones resulta más plausible de aplicación de un modelo estadístico de literatura computacional.



Cuadro 4

4. DISCUSIÓN

El análisis descriptivo nos indica que en un primer estadio la novela gótica contiene un uso atemperado de las categorías léxico-semánticas establecidas para su definición, lo que entronca con el establecimiento por parte de la crítica (Botting 1998) de tres periodos en la evolución del gótico: los comienzos, la explosión de fin de siglo y la decadencia.

En *The Castle of Otranto* de Walpole se encuentran, al menos potencialmente, muchos de los ingredientes que más adelante se desarrollaron en novela gótica. *The Monk* coincide con el periodo de mayor auge de la novela gótica, la mitad de la década de 1790. Al observar la progresión ascendente de los datos léxico-semánticos observamos que coinciden en este hecho, situándose en algunas categorías en el punto más álgido del análisis. Lewis, el autor con el que horror-gótico se hizo popular entre el público lector de la última década del siglo dieciocho y la primera del diecinueve, incorpora el elemento latino que anteriormente Radcliffe ya había consolidado como parte esencial en sus obras. Maturin en *Melmoth the Wanderer* mezcla el horror con un sorprendente realismo. La historia consta a su vez de una serie de historias intrincadas formando una estructura interior que las une. En cada una de las seis historias el Vagabundo o Errabundo, como se ha traducido al español, quien ha vendido su alma en aras de la inmortalidad y poder sin límites, busca la mortalidad e ir al infierno. Todas las convenciones góticas aparecen en las

distintas historias de forma más menos uniforme, hecho que se puede vislumbrar en el análisis descriptivo. Se produce como una vuelta a los comienzos. Si comparamos los datos descriptivos de *Otranto* y *Melmoth* observamos que existe una gran similitud en los resultados de ambos textos. En la decadencia del gótico se produce una vuelta a los comienzos, buscando el gótico sobrenatural, no desmesurado.

Es notable para nuestro estudio que a mayor presencia del escenario sur en el gótico, mayor impacto terrorífico sobre la audiencia. Así, en el caso de *The Monk* el catolicismo sería el elemento que catalizaría de forma destacada el terror gótico en la novela. Estableciendo este parámetro podríamos cotejar todas las obras analizadas, que estén digitalizadas, sobre la proyección de *The Monk*. De esta forma podríamos averiguar la importancia de este elemento en cada una de las obras analizadas.

El análisis realizado bajo la disciplina de la literatura computacional utilizando la metodología de la lingüística del corpus, aunque no es usual en los análisis literarios de los géneros o modos narrativos supone un acercamiento cuantitativo a la ficción. Como dice Stubbs (1996: 81) nuestro modesto objetivo en el fondo, con la hipótesis de trabajo propuesta, es la cambiar las percepciones del estudio de la narrativa y crear nuevos medios de trabajo desde el punto de vista multidisciplinar. Este breve análisis en sí no es más que un esbozo de otras posibles investigaciones en las que se podría realizar un proyecto en el que se podría compilar y etiquetar un corpus completo de novela gótica e intentar establecer un modelado estadístico integral que al mismo tiempo incluyera algoritmos capaces de establecer discriminaciones dentro de la ficción narrativa.

NOTAS

1. <http://www.virtualsalt.com/evalu8it.htm> (1998)
2. Ver recursos de Internet en la Bibliografía

BIBLIOGRAFÍA

- Aguirre, M. 1990. *The Closed Space: Horror Literature and Western Symbolism*. Manchester: Manchester University Press.
- Biber, D. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. & Finegan, E. 1994. "Multi-dimensional analyses of author's style: Some studies from the eighteenth century". D. Ross & D. Brink (eds.) *research in Humanities Computing*, vol. III. Oxford: Oxford University Press, 3-17.
- Botting, F. 1996. *Gothic*. London & New York: Routledge.
- Briggs, Julia. 1977. *Night Visitors. The Rise and Fall of the English Ghost Story*. London: Faber.
- Conrad, S. & Biber, D. 2001. *Variation in English: Multi-Dimensional Studies*. Harlow: Pearson Education Limited.
- Frank, S.F. 1987. *The First Gothics: A Critical Guide to the English Gothic Novel*. New York: Garland Publishing.

- Harris, R. 1997. "Evaluating Internet Research Sources." *VirtualSalt*. 17 Nov. 1997. 17 Oct 2000. <<http://www.virtualsalt.com/evalu8it.htm>>.
- Howard, J. 1994. *Reading Gothic Fiction. A Bakhtinian Approach*. Oxford: Clarendon Press.
- Kilgour, Maggie. 1995. *The Rise of the Gothic Novel*. London & New York: Routledge.
- Punter, D. (ed.) 2000. *A Gothic Companion*. Oxford: Blackwell.
- Sage, Victor & Smith, Allan Lloyd. 1994. *Gothick Origins and Innovations*. Amsterdam; Atlanta, Georgia: Rodopi.
- Stubbs, Michael. 1996. *Text and Corpus Analysis*. Oxford: Blackwell.
- Summers, M. 1938. *The Gothic Quest; a History of the Gothic Novel, by Montague Summers*. London: The Fortune Press.
- Varma, Devendra. 1966. *The Gothic Flame*. New York: Russell & Russell.

Recursos de Internet

<http://etext.lib.virginia.edu/modeng0/browse.html>
<http://www.gutenberg.net>
<http://www.litgothic.com/Texts/spaniard.html>
<http://www.unipr.it/~dsaglia/almeyda/almprosp.html>
<http://www.virtualsalt.com/evalu8it.htm>

Software

MonoConc. 1996. Barlow.
SPSS. 2000. *Statistical Package for Social Science*. v. 11. Hull & Nie.