

Notas para un análisis estadístico del léxico del Laberinto de Fortuna *

Francisco RUIZ NOGUERA **

1. La aplicación de determinadas técnicas lógicas y matemáticas a los estudios del lenguaje han contribuido de forma decisiva a un rápido desarrollo de la lingüística moderna (1). Un claro ejemplo de ello tenemos en las relaciones que se establecen entre lingüística e informática, cuyas bases están en que «la cibernética en última instancia intenta reproducir sintéticamente la base física de la inteligencia humana y en ella se encuentran multitud de actividades de las que, en este momento, nos interesan a los lingüistas» (2). En fin, la importancia de la colaboración entre lingüística e informática es algo que hoy nadie pone en duda, sobre todo si se tiene en cuenta que los frutos son ya una realidad: ahí están los casos de la traducción automática, la glotocronología, los estudios sobre análisis automático del discurso (ADD) de Michel Pêcheux (3), los estudios sobre estadística lingüística (4) o las aplicaciones que de estos principios se hace a la moderna lexicografía para el despojo de los textos y la obtención de unas concordancias (5).

(*) El presente artículo es una muestra del trabajo que bajo la dirección del Dr. Alvar Ezquerro se está llevando a cabo en el Cento de Cálculo de la Universidad de Málaga.

(**) Profesor agregado del I.B. «Cánovas del Castillo» de Málaga.

(1) Vid. J. Roca Pons, *El lenguaje*, Teide, Barcelona, 1973, págs. 438-441.

(2) M. Alvar, «Informática y lingüística», *Revista Española de Lingüística*, 7, 1, 1977, pág. 193.

(3) M. Pêcheux, *Hacia el análisis automático del discurso*, Gredos, Madrid, 1978.

(4) Ch. Muller, *Estadística lingüística*, Gredos, Madrid, 1973.

(5) M. Alinei, *Gli spogli elettronici delle opere di Dante in volgare*, Bolonia, 1975; M. Alvar Ezquerro, *Concordancias e índices léxicos de la «Vida de San Ildefonso»*, Universidad de Málaga, 1980; R. y L. S. de Gorog, *Concordancias del «Arcipreste de Talavera»*, Gredos, Madrid, 1978; Kasten y Anderson, *Concordance to the «Celestina» (1499)*, The Hispanic Seminary of Medieval Studies, Madison, 1977; R. Pellen, *Poema de Mio Cid*. *Dictionnaire lemmatisé des formes et références*, Paris, 1979; A. M. Pollin, *A concordance to the Plays and Poems of Federico Garcia Lorca*, Cornell University Press, Ithaca and London, 1975.

2. El trabajo del que aquí se da muestra ha sido realizado teniendo en cuenta las pautas de colaboración entre las dos disciplinas aludidas.

El primer problema se planteó a la hora de elegir una de entre las distintas ediciones del texto. Se pensó en la edición de J. M. Blecua, e incluso se consultó con este investigador, que aconsejó se utilizara la edición que Foulché-Delbosc incluye en su *Cancionero castellano del siglo XV* (6), texto que presenta la ventaja, para nuestro estudio, de no modernizar la grafía; éste es, finalmente, el texto que hemos utilizado, a pesar de las más recientes ediciones de M. A. Pérez Priego, J. G. Cummins y L. Vasvari Fainberg (7).

El sistema de soporte utilizado para la introducción del material en el terminal UNIVAC UTS 700 del Centro de Cálculo de la Universidad de Málaga, ha sido el de perforación de fichas de doce líneas y ochenta columnas. En cada una de estas fichas figura un verso. Al frente del texto del verso correspondiente aparecen los siguientes elementos: por una parte una referencia a la obra en que se trabaja, en este caso JM LF (Juan de Mena *Laberinto de Fortuna*); a continuación un número que indica la estrofa a que pertenece el verso en cuestión, este número irá, lógicamente, del 1 al 297, total de estrofas de que consta el *Laberinto*; y finalmente una letra (A, B, C, D, E, F, G, H) que indicará la posición del verso dentro de su estrofa.

La preparación previa del material se hizo de acuerdo con las directrices seguidas por M. Alvar Ezquerro (8); es decir, puesto que el teclado de la máquina carece de algunos signos se establecieron las siguientes equivalencias: para la «ç» se empleó el signo @; para la «ñ» una «n» seguida de un guión (n-); las mayúsculas van precedidas de una cruz (+) (9); la tilde se marcó con un apóstrofo colocado detrás de la vocal correspondiente; la diéresis con el signo %, también pospuesto a la vocal.

La perforación del texto arrojó un total de 2.376 fichas, que una vez sometidas a proceso dieron un total de 14.921 palabras-texto —por palabra-texto entendemos las distintas apariciones que puede ofrecer un término o voz en razón de su realización en el habla— y 4.132 palabras distintas. En definitiva, un promedio de 3,61 apariciones por palabra.

Los datos que ofrece la máquina, tras el proceso a que han sido sometidas las fichas, carecen de elaboración; por ejemplo, si nos fijamos en las palabras-texto que tienen una frecuencia absoluta mayor que cien encontramos éstas: *de, que, la, e, los, el, en, non, por, a, con, las, su, más, del, se* (10); algunas de

(6) *Cancionero castellano del siglo XV*, I, Madrid, 1912, págs. 152-183; esta edición es reimpresión de la realizada por el propio Foulché-Delbosc en Mâcon, 1904.

(7) Otras ediciones del *Laberinto de Fortuna*: ed. J. M. Blecua, Espasa Calpe, Madrid, 5.ª ed. 1973; ed. L. Vasvari Fainberg, Alhambra, Madrid, 1976; ed. M. A. Pérez Priego, Editora Nacional, Madrid, 1976; ed. John G. Cummins, Cátedra, Madrid, 1979.

(8) M. Alvar Ezquerro, *Concordancias e índices léxicos*, op. cit., pág. 8.

(9) No se tiene aquí en cuenta las palabras que se escriben con mayúsculas en razón de su posición en el texto, sino sólo los nombres propios.

(10) En este punto las confluencias con otras obras medievales es bien clara; así Alvar Ezquerro en su artículo «Algunos rasgos léxicos de Berceo y su cotejo en otros poemas hagiográficos», *Anuario de Letras*, XVI, 1978, págs. 251-260, constata que las diez formas de mayor frecuencia en las obras de que se ocupa —*Vida de Santa María Egipcíaca, Vida de San Ildefonso, Libro de la Infancia y Muerte de Jesús y Martirio de San Lorenzo*— son prácticamente las mismas, por lo que concluye diciendo que «el abultado número de sus apariciones indica que se trata de unidades gramaticales más que léxicas, con un débil contenido semántico».

ellas no presentan problemas a la hora de ser caracterizadas gramaticalmente, operación que incluso puede hacerse prescindiendo del contexto en que la forma se halle; este es el caso de las preposiciones: *de, en, por, con*; los adverbios: *non, más* (que se introdujo como MA'S, quedando así eliminada la confluencia con la forma de la conjunción adversativa MAS); el artículo: *el* (por la misma razón que acaba de exponerse queda eliminada la posibilidad de confluencia de esta forma con la del pronombre personal, que se consignó como E'L); el posesivo: *su*; y la contracción de preposición y artículo *del* (que a la hora de proceder al siguiente paso de análisis —la lematización— será tratada como *de + el*). Todas las demás formas, de estas dieciséis en que hemos fijado nuestra atención, plantean problemas que deben ser resueltos teniendo en cuenta las referencias que las concordancias nos ofrecen sobre el contexto en que la palabra se halla; no sabemos, en principio, si las formas *la, los, las* corresponden al artículo determinado o al pronombre personal átono de tercera persona; las formas *a, e*, pueden corresponder a la preposición y la conjunción copulativa, respectivamente, o también a formas del verbo *haber*; la forma *que* presentará casos en que sea un pronombre relativo, pero en otras ocasiones estaremos ante una conjunción (sólo quedan eliminadas las formas interrogativas y exclamativas, consignadas como QUE'); finalmente, con respecto a *se*, ignoramos, en principio, si se trata de un pronombre personal, si tiene matiz reflexivo o recíproco, o si es una marca de impersonalidad o pasividad refleja. Pero estas operaciones entran, como ya se ha apuntado, dentro del siguiente proceso, la lematización, que aunque está contemplada en los fines de este trabajo, no corresponde a los límites que hemos establecido para el presente artículo.

3. Observaciones sobre el léxico

Al no incluirse en esta muestra el proceso de lematización, las observaciones que a continuación se hacen no tienen un carácter definitivo, mas no por ello dejan de tener un considerable valor indicativo.

Los trabajos de estadística lingüística son un buen punto de partida para llevar a cabo diversos tipos de estudios entre los que se halla el de la determinación de los campos semánticos de una obra determinada. Lo que sigue a continuación constituye un esbozo sobre este punto. Hemos fijado la atención en algunos de los núcleos temáticos que se observan en el *Laberinto*, y se han recogido los términos pertenecientes a cada uno de ellos, así como sus frecuencias de aparición; se verá que existe una relación directa entre el interés manifestado por el poeta ante un determinado tema y el número de vocablos que en relación con él aparecen.

3.1. Campo semántico de la «guerra»

Aunque el vocablo *guerra* (y otros formados a partir de él como *guerrero*) sólo aparece en diecisiete ocasiones, hay en el texto un total de 257 términos que pertenecen a este campo semántico, con lo que se convierte en el de más abundante aparición (su frecuencia relativa se eleva a 1,7219%). Debemos tener en cuenta que una de las pretensiones de Mena es la de mover a Juan II para que impulse la empresa de la Reconquista, por ello, al hacer referencia a los

distintos reyes de Castilla se detendrá en aquellos cuya dedicación a este menester pueda ser considerado como modelo. Junto a la lucha contra los árabes se hallan también presentes las contiendas internas en el reino castellano; es lógico, por tanto, que el campo semántico de «lo guerrero» esté ampliamente representado.

Los términos que consideramos aquí pueden agruparse en varios apartados:

- a) La acción propia de «guerrear», donde además de los ya señalados tenemos:

<i>Término</i>	<i>Frec. absoluta</i>	<i>Frec. relativa</i>
atacar	1	0,0067
aventura	1	0,0067
batalla (y variantes)	10	0,0670
combate (y variantes)	3	0,0201
fazaña	4	0,0268
fecho	39	0,2613
lucha	1	0,0067
pelea	3	0,0201
querella	2	0,0134

- b) Acciones que aparecen como consecuencia del combate:

conquistar (y variantes)	12	0,0804
ferir (y variantes)	19	0,1273
ganar (y variantes)	23	0,1541
triumfo	3	0,0201
vençer (y variantes)	15	0,1005
vitoria	6	0,0402

- c) Sobre las cualidades del buen guerrero:

brauo (y variantes)	7	0,0469
brio	1	0,0067
corage	2	0,0134
feroçe	6	0,0402
fuerça	18	0,1206
valor (y variantes)	13	0,0871
vigor	1	0,0067

- d) Sobre la fortaleza del combate:

sangre (y variantes)	10	0,0670
saña (y variantes)	10	0,0670

- e) Sobre los enemigos:

adversario	2	0,0134
enemigo	7	0,0469
moro (y variantes)	9	0,0603

- f) Sobre las armas:

armas	10	0,0670
arnés	2	0,0134

3.2. Campo semántico de la «realeza, los títulos nobiliarios, el tratamiento»

Los términos incluidos en este campo alcanzan una frecuencia absoluta de 180 ocurrencias, lo que supone una frecuencia relativa de 1,2060%.

De los apartados que observamos en este grupo merece especial atención el primero, ya que al estar dirigida la obra a Juan II de Castilla es lógico que el campo léxico de los vocablos que hacen alusión al monarca sea el que cuente con mayor número de términos:

a) Referencia directa al rey o a personas de la familia real:		
alteza	1	0,0067
emperador	2	0,0134
monarca (y variantes)	3	0,0201
príncipe	5	0,0335
rey (y variantes)	74	0,4958
soberano	1	0,0067
b) Atributos reales:		
cepro	3	0,0201
corona	8	0,0536
trono	2	0,0134
c) La nobleza:		
conde	14	0,0934
condestable	8	0,0536
duque	1	0,0067
noble (y variantes)	5	0,0335
d) Otros títulos y formas de tratamiento:		
adelantado	3	0,0201
cauallero	3	0,0201
don (y variantes)	15	0,1005
señor (y variantes)	31	0,2077

3.3. Muy relacionado con éste tenemos el que podíamos denominar *campo semántico de los términos que indican ejercicio de poder*. Téngase en cuenta que el *Laberinto* es una obra en la que Mena se muestra partidario de una España cuya fortaleza radique en su buen gobierno, por ello abundarán los términos que hagan referencia a la cuestión, bien para propiciar la alabanza, bien para ejercer la censura. La frecuencia absoluta que alcanzan los vocablos aquí considerados es de 122, la relativa de 0,8174%. Estableceremos los siguientes grupos:

a) Sobre el poder:		
auctoridad	1	0,0067
gouernar (y variantes)	6	0,0402
mandar (y variantes)	12	0,0804
ordenar (y variantes)	16	0,1072
poder (y variantes)	10	0,0670
regir	9	0,0603
tiranía (y variantes)	6	0,0402

b) Sobre formas de ordenación social:		
estado	11	0,0737
nação	1	0,0067
política	1	0,0067

Es curioso comprobar que, aun cuando el *Laberinto* es concebido como un poema nacional, el término *nação* sólo aparezca en una ocasión.

c) Cualidades que debe tener el buen gobernante y fines que debe perseguir:		
justiça (y variantes)	30	0,2010
onor (y variantes)	7	0,0469
prudencia	9	0,0603
d) Si hay gobernantes es lógico que haya gobernados, sin embargo es muy breve la referencia que se hace a ellos, téngase en cuenta que Mena declara que la pretensión del <i>Laberinto</i> está en «virtudes e vícios narrar de potentes», de ahí la escasez de términos en este apartado:		
servidor (y variantes)	3	0,0201

3.4. Relacionado con el primer campo semántico de que nos ocupábamos —el de la «guerra»— está el de *la muerte y el dolor*, aunque a veces estos dos términos no aparezcan como consecuencia de una acción bélica, sino —sobre todo el segundo de ellos: *dolor*— como muestra ejemplar, cuando se hace referencia a aquellos personajes que por haber obrado mal se hallan en la parte baja de las ruedas que ve el poeta.

La frecuencia absoluta de estos términos es de 120 apariciones, relativa de 0,8040%:

a) El dolor:		
aziago	1	0,0067
cruel	16	0,1072
dolor (y variantes)	11	0,0737
padeçer (y variantes)	8	0,0536
pena	1	0,0067
pesar	3	0,0201
plañir (y variantes)	3	0,0201
sufrir	12	0,0804
b) La muerte:		
morir (y variantes)	65	0,4355

3.5. El del *bien y el mal* es otro de los campos semánticos cuyos términos tienen una más alta frecuencia, 204 absoluta y 1,3668%, no en vano el *Laberinto* tiene un considerable componente de obra moralizante. Dos grupos fundamentales estableceremos:

a) El bien:		
bien (y variantes)	52	0,3484
bueno (y variantes)	37	0,2479
virtud (y variantes)	37	0,2479

b) El mal:

mal (y variantes)	45	0,3015
perfidia (y variantes)	3	0,0201
viçio (y variantes)	20	0,1340
vil (y variantes)	10	0,0670

Atendiendo al carácter ejemplar y al tono moral a que hemos aludido puede incluirse aquí la contraposición que se establece entre riqueza y pobreza (sobre todo en la copla 227, en que Amiclas se alza como modelo):

riqueza (y variantes)	19	0,1273
miseria	5	0,0335
pobreza	8	0,0537

En relación con ello podemos considerar las alusiones de tipo religioso que hallamos en el poema, cuya frecuencia absoluta de 53 y relativa de 0,3551% es notablemente inferior a la de los campos anteriores:

católico	2	0,0134
cielo	14	0,0938
diuino	12	0,0804
religión (y variantes)	2	0,0134
sacro (y variantes)	6	0,0402
santo	16	0,1072

Contrastando con ello encontramos 20 términos (frecuencia relativa 0,1340%) que aluden a la magia o a la predicción del futuro:

agüero	2	0,0134
astrólogo	1	0,0067
astronomía	1	0,0067
mago (y variantes)	7	0,0469
nigromantessa	1	0,0067
profetizar (y variantes)	8	0,0536

Entre los consejos que en el *Laberinto* se encuentran, relativos a cuestiones diversas, debemos destacar los que se refieren al amor, sentimiento censurado por Mena, a no ser que se trate del «linpio católico amor virtuoso» que se consigue «por el sacramento matrimonial». La frecuencia absoluta de estos términos es de 39 y la relativa de 0,2613%.

amor (y variantes)	28	0,1876
castidad (y variantes)	9	0,0603
matrimonial (y variantes)	2	0,0134

Digamos, en fin, que una de las finalidades del *Laberinto* está en exponer los hechos virtuosos para que la Fama se encargue de difundirlos. Los términos relativos a este campo no tienen, sin embargo, una frecuencia muy elevada; hay un total de 38 ocurrencias, lo que supone una frecuencia relativa de 0,2546%.

fama (y variantes)	21	0,1407
gloria (y variantes)	17	0,1139

El tratamiento de este tema no constituye, como se sabe, una innovación de Mena sino que está inserto en una tradición (11).

3.6. Los nombres propios

En la obra aparecen un total de 461 nombres propios, lo que supone una frecuencia relativa de 3,0887%. Estos nombres pueden agruparse según el siguiente esquema:

	Frec. absoluta	Frec. relativa	Porcentaje en relación con el número total de nombres propios
A. Referentes a España	153	1,0251	33,2
A.1. Personas	51	0,3417	11,05
A.1.1. Reyes	29	0,1943	6,3
A.1.2. Otros nobles y guerreros	22	0,1474	4,75
A.2. Lugares geográficos	102	0,6834	22,15
B. No referentes a España	308	2,0636	66,8
B.1. Personajes de la Antigüedad	49	0,3283	10,6
B.2. De la Mitología	126	0,8442	27,3
B.3. Lugares geográficos	117	0,7839	25,45
B.4. Relacionados con lo religioso	16	0,1072	3,45

Como vemos, a pesar del carácter nacional del *Laberinto*, las referencias a los nombres no relacionados de forma directa con España es superior al de aquellos que sí lo están. Es destacable el elevado número de términos mitológicos, que en la mayoría de los casos aparecen como modelos (para la imitación o para la censura). Con todo hay que aclarar que el nombre de lugar que tiene mayor frecuencia es el de *España*, con 12 ocurrencias, y a continuación el de *Castilla*, 6 ocurrencias; además, la mayoría de los nombres de los personajes de la nobleza española, así como de los lugares geográficos, tienen una relación directa con la Reconquista, lo que viene a reforzar la idea, ya expuesta, de la importancia que da el poeta a esta empresa guerrera.

Digamos, en fin, que la explicación a muchos de los términos utilizados por Mena tiene una estrecha relación con la idea que defiende M.^a Rosa Lida al considerar al autor como un poeta del prerrenacimiento español (12).

(11) Vid. M.^a Rosa Lida, *La idea de la fama en la Edad Media*, México, 1952.

(12) Vid. M.^a Rosa Lida, *Juan de Mena, poeta del prerrenacimiento español*, México, 1950. Para una explicación del léxico utilizado por Mena vid. F. Lázaro Carreter, «La poética de arte mayor castellano», *Estudios de poética*, Taurus, Madrid, 1976, págs. 75-111.