

APORTACIONES INFORMÁTICAS Y ESTADÍSTICAS AL ANÁLISIS LÉXICO DE UNA OBRA LITERARIA

ÁNGELES SÁNCHEZ HERNÁNDEZ
Universidad de las Palmas de Gran Canaria

0. INTRODUCCIÓN

En esta comunicación, queremos mostrar cómo un análisis estadístico, complementado por el empleo de programas informáticos para la obtención de datos, ayuda a establecer unas bases objetivas y sólidas para los estudios literarios, tachados con frecuencia de intuicionistas.

La crítica contemporánea ha puesto de relieve la consideración de la obra literaria como un todo orgánico que no podría reducirse a la suma de sus partes, significado y significante, ni olvidar tampoco al lector y su manera de percibir los mensajes literarios. La importancia de los distintos elementos que conforman la obra radica en las interrelaciones entre ellos.

La esencia de la obra literaria se centra en su lengua, y no podemos negar que en cada obra artística existen unas peculiaridades que la diferencian de las demás. Estas particularidades concuerdan con el concepto de desviación que es importante para el estudio del estilo de un autor. Para que una figura del lenguaje resulte significativa, desde el punto de vista estilístico, tiene que desviarse de una norma. La estilística ha tomando como base los estudios lingüísticos que consideran que toda expresión verbal, sin excepción, está unida a un presente y a unas circunstancias precisas que la lengua enuncia por medio de una morfología específica (Benveniste, 1974: 226). Este tipo de investigación literaria ha intentado llevar a cabo estudios que analicen las especificidades lingüísticas de los textos literarios. Los estudios recientes de estilística se preocupan de fundamentar científicamente sus investigaciones con técnicas tomadas de la lingüística, de la estadística o de la teoría de la recepción empírica en lo que algunos, como Van Buuren (1997: 52), denominan *micro-stylistique* que se ocupa de las unidades inferiores a la frase.

Una de las teorías estilísticas que pueden incluirse dentro de esta microestilística es la que se ocupa de las palabras como unidades que sirven para establecer los campos estilísticos, fundamentados en los campos semánticos. La profesora Gardes-Tamine (1988: 102) define los campos semánticos como “l’association d’un champ notionnel et d’un champ lexical” y considera que el campo conceptual debe ser un conjunto cerrado y el campo léxico una lista cerrada. Opone la noción de *champs sémantiques* a las *séries étymologiques* o familias de palabras cuyos componentes no están unidos por un concepto sino que tienen una relación formal, una misma raíz. También la considera contraria a *champs associatifs* que agruparían a varias palabras entorno a un concepto dado porque estos últimos admiten términos de distintas partes del discurso (adjetivos, nombres, verbos) mientras que los campos semánticos sólo incluyen palabras de una misma parte del discurso. En España, la profesora Bobes Naves (1989: 23) explica este mismo concepto en términos muy similares considerando el campo semántico como una organización de significaciones de estructura cerrada cuyas unidades se relacionan por coincidencias u oposiciones entre ellas. El sentido de una palabra en la obra depende, pues, de sus relaciones con las demás.

El vocabulario es una de las principales fuentes de expresividad de la lengua. Hablar supone elegir unas formas concretas de esa lengua y combinarlas de una manera personal en el discurso. La riqueza léxica es un criterio que se ha usado habitualmente para clasificar los textos (Delcourt, 1987: 137); no obstante, hay que tener en cuenta la distinción entre léxico y vocabulario, y reservar el primer término para la lengua cuyas unidades serán los *lexemas* y el segundo término para su actualización en el discurso con unidades llamadas *vocablos*; la *palabra* designará toda ocurrencia de un vocablo cualquiera como señala Muller (1973: 230).

Si las palabras son un elemento constituyente de ese estilo que manifiesta el pensamiento, individual o colectivo, que marcará la desviación de la norma, parece evidente que medir y definir esa desviación por medios matemáticos propiciará un punto de partida, lo más objetivo posible, para una interpretación posterior. En la base de este tipo de trabajos, está la idea de *frecuencia* como característica de las palabras, a partir de las cuales se pueden deducir determinados rasgos formales o de contenido relativos al texto en el que están inmersas. La cantidad de información de un signo es definida por su probabilidad. La significación de las palabras está íntimamente ligada a la distribución y a la frecuencia de uso de éstas (Guiraud, 1955: 99; Escobedo Rodríguez, 1994: 16).

Es difícil sustraerse a la idea del cálculo matemático como medio más útil y fiable de medición de esos fenómenos lingüísticos. La cita que sigue explica este proceso:

El vocabulario empleado puede ser tomado como índice revelador del modo en que los sujetos conciben un determinado hecho, fenómeno o realidad, y el estudio de las frecuencias de este vocabulario, las frecuencias relativas dentro de unos textos y de otros, y las estructuras relacionales que pueden llegar a establecerse entre vocablos y determinadas características de los emisores nos permiten aproximarnos al contenido de la información textual y llegar a interpretarla (Etxeberría et al., 1995: 140).

La estadística es la ciencia que se ocupa de las alteraciones aleatorias. Recurrir a las técnicas estadísticas supone dar mayor rigor a los datos obtenidos por otras vías más intuitivas o subjetivas (Muller, 1977: 77). En los últimos años, la aplicación de la informática para realizar los recuentos necesarios para la matemática ha hecho que estos estudios se desarrollen y simplifiquen su tarea vertiginosamente, pudiéndose almacenar, ordenar y clasificar los datos lingüísticos de manera electrónica (Ávila Muñoz, 1998: 1; Morras, 1999: 196).

Nuestro trabajo se fundamenta en la selección de los verbos-clave de *Citadelle*, obra póstuma de A. de Saint-Exupéry. La obtención de esos datos es útil para la constitución de los campos estilísticos. Utilizamos la Estadística Textual para realizar el recuento de unidades verbales y operar, posteriormente, algún tipo de análisis matemático. Antes de proceder al recuento de unidades hay que llevar a cabo un paso previo que es la reducción de datos, basado en la categorización y codificación (Muller, 1977: 35). Hay que desambiguar el texto y realizar la segmentación del corpus textual.¹ Es necesario dividir el corpus en varias partes para poder establecer más tarde, mediante los cálculos estadísticos, las coincidencias o las divergencias léxicas que se observan entre las distintas subpartes de la obra literaria, porque el objetivo del análisis cuantitativo no es la cantidad en sí, sino las distorsiones que podemos constatar de un texto a otro en ciertos empleos (Bernard, 1999: 39; Muller, 1992: 16; Tournier, 1989: 21).

Para proceder a la segmentación debemos diferenciar las unidades elementales. Entre ellas las más habituales son: la forma gráfica, el lema, los cuasisegmentos y los segmentos repetidos.

1. En general, se denomina *texto* a todo enunciado o sucesión de enunciados; por *corpus* se entiende una colección definida de textos.

Se habla de *lema* identificándolo con *lexema* y de *lematización* como la operación que consiste en unir cada término de un texto a un lema o, si seguimos la propuesta de Sassi (1999: 174): “es la asignación de cada palabra a su entrada en el diccionario”. Algunos investigadores, como Charles Muller, consideran más útil trabajar el vocabulario del texto por medio de la lematización; otras posturas, como la de Lebart y Salem, optan por la forma gráfica por creerla más completa. En nuestro estudio, hemos optado por las *formas gráficas* porque nos permiten una segmentación del texto en unidades inequívocamente diferenciadas, fácilmente reconocibles por el ordenador, y propicias para el tratamiento automático. Salem (1989: 31-32) las define como una secuencia de caracteres no delimitadores (en general, letras) comprendidas entre dos caracteres delimitadores (espacios o signos de puntuación).

1. TRASLADO DEL LIBRO A SOPORTE INFORMÁTICO Y ELABORACIÓN DE LOS INVENTARIOS VERBALES

La edición de *Citadelle* empleada para poner la obra en soporte informático es la publicada por Gallimard en 1971, en su colección *Le Livre de Poche*, nº 1.532, 1.533, 1.534. El manuscrito, que su autor no pudo corregir, llegó a la editorial lleno de tachaduras, incorrecciones y líneas en gran parte ilegibles. Esta edición presenta el texto que es considerado como definitivo en la actualidad, gracias a la puesta a punto de Simone Lamblin, Pierre Chevrier y Léon Wencelius.

El traslado del libro a soporte informático conlleva la creación de archivos de texto en ASCII y, posteriormente, su pase a archivos en formato DBF, para así poder trabajar con el programa FRECON.UIB.² El texto de *Citadelle* ha sido dividido en cinco partes con un número de registros o líneas igual en cada una de ellas (4.461 registros en cada una de esas divisiones efectuadas, lo que nos da un número total de 22.305 registros de la base de datos).

Hemos partido, como ya ha quedado dicho, del recuento de formas gráficas porque creemos que perdemos menos información que con la lematización. Para solucionar los problemas de homonimia de ciertas palabras se ha procedido a suprimir su ambigüedad, por medio de un marcaje adicional a la palabra, en los casos que se han revelado como necesarios al realizar los inventarios. Para ello, el programa nos proporcionaba la *forma-polo* que nos interesaba en su *concordancia* correspondiente para poder aclarar el sentido de la palabra en el discurso, contando con la posibilidad de ver el número de líneas del texto que se deseaban; o bien, un resumen de sus apariciones. Con el citado programa informático se generaron los ficheros de frecuencias de todas las palabras que el texto incluye, sobre los que se marcaban las formas verbales.

Así se obtuvieron los inventarios lexicográficos y lexicométricos que nos han permitido acceder a los términos de mayor presencia en el corpus, y nos han dado una primera información sobre la importancia cuantitativa de cada uno de los significados.

2. OBTENCIÓN DE LAS PALABRAS ESPECÍFICAS DE *CITADELLE*

Sobre estos glosarios de los verbos, se procedió a la reescritura en documento ASCII de todos los verbos según la frecuencia de aparición, empleando el programa REPITE.³ Sobre este texto base, se aplicó la ley hipergeométrica desarrollada por Lafont y considerada como la más apropiada para el vocabulario característico. Los cálculos fueron llevados a cabo por medio del programa SPAD.T.⁴ Estas palabras específicas resultan ser aquellas en las que se

2. Programa desarrollado por Gabriel M^a Jordà (Filología Francesa), José L. Guerrero (Informática), Ángel Igelmo (Análisis Matemático) de la Universitat de Les Illes Balears.

3. Programa informático realizado, igualmente, en la Universidad balear por el Dr. Igelmo.

4. SPAD.T, programa de ordenador distribuido por CISIA que funciona en ordenadores PC o MacIntosh.

alcanzan los índices más altos para el valor-test medio, calculado a partir del valor-test de cada una de las formas que la componen. Así hemos obtenido las palabras infrautilizadas y aquellas que gozaban de una sobreutilización en cada una de las partes en las que hemos dividido el texto original de Saint-Exupéry.

Los valores-test obtenidos se presentan en cada una de las cinco partes en que hemos dividido la obra. Nos han servido para establecer las conclusiones convenientes, en cuanto que lo interesante de estos resultados es la comparación de las subpartes del texto para intentar constatar las fluctuaciones o la persistencia de algunas ideas que las palabras de Saint-Exupéry expresan, hecho que resulta imprescindible en una obra como *Citadelle*, escrita a lo largo de varios años.

Los resultados de estos análisis fueron los siguientes:

Figura 1

Primera parte. Los verbos de *Citadelle*. Páginas 15-133

<i>Formas sobreutilizadas</i>	<i>Valor test</i>	<i>Formas infrautilizadas</i>	<i>Valor test</i>
venaient	4.106	est	-10.040
vide	3.481	es	-7.247
faites	3.476	soit	-5.542
disait	3.398	as	-4.491
sommes	3.247	sont	-4.340
enterre	3.195	faire	-4.268
chantait	3.195	dire	-3.768
enseignerez	3.195	sois	-3.683
édifie	3.195	fait	-3.659
forme	3.181	juger	-3.544

Figura 2.

Segunda parte. Los verbos de *Citadelle*. Páginas 133-253

<i>Formas sobreutilizadas</i>	<i>Valor test</i>	<i>Formas infrautilizadas</i>	<i>Valor test</i>
renie	3.737	ont	-5.253
laisse	3.179	parle	-3.795
réponds	2.894	prétends	-3.066
abordent	2.850	irai	-2.917
collabore	2.850	es	-2.478
fane	2.850	sert	-2.433
connaissais	2.830	voir	-2.412
donné	2.740	vide	-2.337
donnes	2.738	importe	-2.292
pouvoir	2.687	changer	-2.277

Figura 3.
Tercera parte . Los verbos de *Citadelle*. Páginas 253-373

<i>Formas sobreutilizadas</i>	<i>Valor test</i>	<i>Formas infrautilizadas</i>	<i>Valor test</i>
est	7.937	suis	-3.409
es	6.729	change	-3.065
voir	4.882	été	-3.045
noue	4.282	saurais	-2.316
trouve	4.185	seront	-2.257
soit	4.069	porte	-2.147
communiquer	3.862	cherche	-2.040
rit	3.822	était	-1.983
pouvoir	3.716	venir	-1.767
sourire	3.437	enferme	-1.967

Figura 4.
Cuarta parte. Los verbos de *Citadelle*. Páginas 374-496.

<i>Formas sobreutilizadas</i>	<i>Valor test</i>	<i>Formas infrautilizadas</i>	<i>Valor test</i>
ait	6.204	vient	-6.472
es	5.045	vois	-5.095
agir	4.855	vu	-5.095
étant	4.175	vivre	-4.164
ira	4.006	voir	-4.021
allé	3.952	vint	-3.444
dévore	3.658	pouvoir	-3.332
ressemble	3.287	refuse	-2.820
est	3.272	vont	-2.782
dirai	3.220	croire	-2.433

Figura 5
Quinta parte. Los verbos de *Citadelle*. Páginas 496-617

<i>Formas sobreutilizadas</i>	<i>Valor test</i>	<i>Formas infrautilizadas</i>	<i>Valor test</i>
été	5.477	es	-5.939
plaît	4.971	a	-5.735
marche	4.708	trouve	-4.281
suis	4.676	étant	-3.350
viens	4.422	étaient	-3.347
vois	4.053	saisir	-2.886
prétends	3.851	créer	-2.375
sauverai	3.813	sourire	-2.343
irai	3.699	laisse	-2.127
chantes	3.630	sachant	-1.995

Con estas cinco tablas de los diez verbos sobreutilizados y los diez infrautilizados en cada una de las partes, podemos ya hacer algunas apreciaciones. El verbo *venir*, bajo la forma del imperfecto *venaient*, (4.108) aparece como la más utilizada en la primera parte, y como infrautilizado bajo las formas de singular del presente *vient* (-6.472) y del pretérito indefinido

vint (-3.444) en la cuarta. Las formas *est* (7.968) y *es* (6.735) son las dos más utilizadas en la tercera parte; sin embargo, resultan ser las dos menos características (-10.040, -7.247) en la primera parte, siguiendo el mismo orden.

Comprobamos cómo otras formas aparecen con sobreutilización en unas partes y en otras las encontramos con una ocurrencia mínima como podemos apreciar en la siguiente tabla – resumen:

<i>Verbo</i>	<i>Sobreutilización</i>	<i>Infrautilización</i>
<i>vide</i>	1 ^a	2 ^a
<i>laisse</i>	2 ^a	5 ^a
<i>est</i>	4 ^a , 3 ^o	1 ^a
<i>es</i>	3 ^a , 4 ^a	1 ^a , 2 ^a , 5 ^a
<i>soit</i>	3 ^a	1 ^a
<i>pouvoir</i>	3 ^a	4 ^a
<i>sourire</i>	3 ^a	5 ^a
<i>été</i>	5 ^a	3 ^a
<i>vois</i>	5 ^a	4 ^a
<i>voir</i>	3 ^a	2 ^a , 4 ^a
<i>suis</i>	5 ^a	3 ^a
<i>prétends</i>	5 ^a	2 ^a
<i>trouve</i>	3 ^a	5 ^a
<i>irai</i>	5 ^a	2 ^a

De las cincuenta formas verbales que resultaron con sobreutilización, encontramos ocho pertenecientes al verbo *être*, tres al verbo *aller*, dos al verbo *venir*, dos al verbo *voir*, dos al verbo *donner* y dos al verbo *chanter*. Hay verbos que sólo aparecen con sobreutilización como: *forme, fane, noue, enterre, édifie, renie, communiquer, rit, sauverai, dévore, agir, plaît, marche*. Otros los encontramos, únicamente, entre los infrautilizados, como por ejemplo: *parle, change, changer, juger, croire, importe, sert, cherche, refuse, vivre, saisir, créer, enferme*.

Se observa, igualmente, que hay verbos que podemos ver tanto entre los más característicos como entre los de menor incidencia, pero hay cambios en cuanto al tiempo, al modo o a la persona a la que se refiere y, sobre ellos, deberemos detenernos porque posiblemente estas variaciones lleven consigo diferencias de sentido. Ejemplo de este caso lo tenemos en el verbo *être* o en el verbo *faire* con su gran variedad de tiempos, modos y personas, pudiéndose inferir de este hecho datos significativos. Por esta razón, aunque los verbos-clave con los que trabajemos se eligen a partir de los más característicos y los de mayor homogeneidad, nunca se ignorarán ciertas relaciones en el léxico que nacen justamente de estas oposiciones.

En cuanto a su contenido semántico vemos algunos que nos indican un movimiento o acercamiento como *marche, venaient, irai*, o una acción *agir, faites*, que se agudiza ciertamente con términos como *dévore* y *sauverai*. Esa acción se acompaña de palabras que nos muestran una cierta alegría: *chantait, rit, sourire, chantes* y una solidaridad en las relaciones: *collabore, donne, donné, plaît, noue*. Algunas pertenecen a campos semánticos que representan los trabajos de la tierra: *enterre* y, a la vez, dan sensación de afianzamiento en lo más tangible, por lo que contrasta con una construcción del hombre que se yergue hacia las alturas en la palabra *édifie*. Esas acepciones referidas a la naturaleza se repiten con otros verbos como *fane*, cuya primera significación está unida al ciclo evolutivo de las flores aunque, por extensión, pueda incluir otros sentidos que indiquen un deterioro siguiendo el contexto en el que aparezca. Llama la atención también la sobreutilización de *connaisais* y la infrautilización, en otra de las partes,

de *sachant*. Estos dos verbos muestran el conocimiento pero con matices diferentes. Leemos verbos, entre los sobreutilizados, que muestran una intención de transmisión de ideas u opiniones: *réponds*, *communiquer*, *dirai*, *disait*, *enseignerez* y que chocan, un poco, con la infrautilización de *parle*, posiblemente porque existen diferencias importantes entre ellas. O bien, dentro del mismo campo conceptual de la expresión de ideas, la negación rotunda de *renie* cuyo contexto es digno de un examen exhaustivo.

Podemos apreciar que de las tablas obtenidas cuantitativamente ya se desprenden ciertas relevancias significativas que merece la pena estudiar con mayor detenimiento y cotejar con otras informaciones que van más allá de análisis lingüístico.

3. FORMAS HOMOGÉNEAS DEL TEXTO

En la fase posterior, hemos obtenido las palabras más homogéneas aplicando el método Chi-cuadrado. La estadística se ha llevado a cabo por medio del programa ESTADI. Diremos que han resultado homogéneas aquellas palabras cuya frecuencia de aparición ha sido similar en las cinco partes en que hemos dividido la obra. Por tanto, no presentaban diferencias estadísticamente significativas.

Previamente, se había elaborado un listado con los verbos cuyo umbral de frecuencia era superior a 99; es decir, las formas que en todo el texto de *Citadelle* tenían una ocurrencia igual o superior a 100. La lista es la que sigue:

Verbos con umbral de frecuencia superior a 99

	FREC.					TOTAL
	1ª parte	2ª parte	3ª parte	4ª parte	5ª parte	
1 a	30	25	34	30	1	120
2 agit	13	24	28	29	19	113
3 ai	14	14	26	26	20	100
4 dire	13	31	29	30	26	129
5 dis	27	37	45	47	40	196
6 dit	43	31	29	29	32	164
7 es	9	13	53	43	72	190
8 est	538	681	806	664	546	3235
9 faire	25	39	58	58	34	214
10 fait	44	55	43	46	49	237
11 faut	23	31	21	35	22	132
12 peut	30	34	40	33	22	159
13 sais	26	16	22	19	24	107
14 sera	22	30	25	18	22	117
15 soit	16	39	63	38	41	197
16 sont	74	87	113	101	107	482
17 suis	23	40	17	39	55	174
18 veut	16	22	31	35	34	138
19 vient	22	33	27	0	34	116
20 était ⁵	80	69	45	63	43	300
21 être	51	81	67	68	42	303

5. Estas dos últimas palabras están fuera del orden alfabético por los acentos que contienen. Se han colocado al final del listado.

Sobre estas palabras el programa ha aplicado el estadístico Chi-cuadrado con un nivel de significancia $\alpha = 0.05$ y 4 grados de libertad (9.488), y nos ha ido dando el estadístico numérico. Comprobamos que se acepta la hipótesis nula en nueve de estas palabras: *sais* (2.93), *sera* (3.51), *dit* (4.39), *peut* (5.74), *faut* (5.97), *dis* (6.15), *ai* (7.08), *dire* (8.39), *veux* (9.48).

Hay un primer dato, con respecto a los verbos más homogéneos, que nos parece curioso, es el hecho de que el primero se refiere al ámbito cognitivo: *sais*. Este verbo *savoir* estaba entre las formas estudiadas en el apartado anterior como menos características bajo su forma de gerundio *sachant*.

Además resaltaremos la similitud de campos semánticos entre las homogéneas y las características, con repetición de algunos de los verbos como *être* y el verbo *dire* que se repite tres veces, en infinitivo *dire* y dos de sus formas de presente correspondientes al *je* y al *il*. Este mismo rasgo, singular y presente, se lee en las restantes formas que marcan una necesidad y voluntad personal, excepto el futuro *sera* que parece abrir una puerta al porvenir.

4. SELECCIÓN DE VERBOS PARA EL ESTUDIO DE LA ESTRUCTURA LÉXICA

Hemos considerado conveniente elegir los tres verbos característicos de cada una de las cinco partes, siempre que su valor test sea superior a 3.000 (por esta razón queda excluida la tercera palabra de la segunda parte), y los nueve que presentan homogeneidad, quedándonos pues con un total de veintitrés para adentrarnos en el estudio de los campos estilísticos y en la estructura léxica de *Citadelle*. Los verbos seleccionados son los que siguen:

CARACTERÍSTICAS	Formas-clave				
	1° parte	2° parte	3° parte	4° parte	5° parte
	venaient	renie	est	ait	été
	vide	laisse	es	es	plaît
	faites		voir	agir	marche
HOMOGÉNEAS:	sais, sera, dit, peut, faut, dis, ai, dire, veux.				

Debemos señalar, primeramente, que hay un verbo que se repite en la 3ª y en la 4ª parte por lo que aparecerán en los análisis posteriores 22 formas y no 23. Hay que subrayar ciertas curiosidades ante el cuadro anterior, como por ejemplo *renie* y *vide* que dan la sensación de pasar página sobre lo conocido. También observar una constatación de gentes que se acercan hacia algo que no podemos saber aún qué es con el verbo *venaient*. Las diferentes formas del verbo *être*, con un número de ocurrencias muy elevado, obligan a verlas en sus contextos, de la misma forma que las dos del verbo *avoir*, para poder clarificar así su sentido. Y otro tanto podemos decir de ese más explícito *faites* que se abre a un público plural, cara a cara, pero del que ignoramos las consignas dadas tras el verbo.

5. CONCLUSIONES

Creemos que el uso del léxico y de la estadística es un instrumento muy válido y objetivo para acercarse al pensamiento de un autor. Esta primera fase, nos ha servido para poder obtener, por un lado las frecuencias y el grado de homogeneidad de los verbos y, por otro, las formas específicas; disponiendo así de un listado de verbos representativos para iniciar la investigación cualitativa. Estos datos nos permiten un acercamiento a la realidad de la obra y podemos establecer los campos estilísticos que necesitan de una interpretación reflexiva para desvelarnos el universo simbólico de Saint-Exupéry.

En esa fase posterior, hemos necesitado la ayuda de las teorías de G. Bachelard o de G. Durand. Sin embargo, la fundamentación lingüística en los verbos ha servido para clarificar, muy ampliamente, el pensamiento del escritor. Las formas verbales, como *je dis* o *je sais*, han señalado claramente su opinión sobre el devenir del hombre, y su necesidad de compromiso con la sociedad si quiere realizarse individualmente. Otras como *je renie*, nos muestran su intención de no excluir a nadie, opinión importante si tenemos en cuenta el momento en el que se escribió *Citadelle* y las constantes controversias en las que su autor participó, involuntariamente, en la etapa de la Segunda Guerra Mundial. Hay imperativos, como *laisse* o *faites*, que expresan deseos de paz. No vamos a seguir explicando las manifestaciones posteriores de estas formas verbales porque es demasiado extenso para esta comunicación, pero sí reafirmar el resultado tan fructífero al que nos ha llevado el estudio del vocabulario.

BIBLIOGRAFÍA

- Ávila Muñoz, A. M (1998) “Elaboración de índices de frecuencia léxica sobre *corpora* de lengua hablada: una aplicación a la lingüística del *corpus*”, en Luque Durán, J. de D. y Manjón Pozas, F.J. (eds.), *Teoría y práctica de la lexicología*, Granada, Granada lingüística y Método ediciones, pp.1-14.
- Bernard, M. (1999) *Introduction aux études littéraires assistées par ordinateur*, Paris, PUF.
- Benveniste, E. (1974) *Problèmes de linguistique générale II*, Paris, Gallimard.
- Bobes Naves, M. C. (1989), *La semiología*, Madrid, Síntesis.
- Delcourt, C. (1987) “La statistique littéraire” en Delcroix, M., Hallyn, F. (eds.), *Méthodes du texte. Introduction aux études littéraires*, Paris - Louvain-la-Neuve, Duculot, pp.132-147.
- Escobedo Rodríguez, A. (1994) *Estudios de lexicología y lexicografía*, Almería, Universidad de Almería, Servicio de Publicaciones.
- Etxeberría, J., García, E., Gil, J., Rodríguez, G. (1995) *Análisis de datos y textos*, Madrid, RA-MA.
- Gardes-Tamine, J. (1988) *La grammaire*, Paris, Armand Colin.
- Guiraud; P: (1955) *La sémantique*, Paris, PUF.
- Morras, M. (1999) “Informática y crítica textual: realidades y deseos”, en Blecua, J-M., Clavería, G., Sánchez, C., Torruella, J. (eds.) *Filología e informática*, Barcelona, Editorial Milenio i UAB, pp.189-210.
- Muller, Ch. (1973) *Estadística lingüística*, Madrid, Gredos.
- Muller, Ch. (1977) *Principes et méthodes de la statistique lexicale*, Paris, Hachette.
- Muller, Ch. (1992) *Initiation aux méthodes de la statistique linguistique*, Paris, Champion.
- Salem, A. (1989) “Segments répétés et analyse statistique des données textuelles”, en Bozzo i Duran, M. (ed.) *Jornades de lexicometria*, Barcelona, PPU, pp.29-58.
- Sassi, M. (1999): “Concordancias para filólogos: en pos de la simplicidad”, en Blecua, J-M., Clavería, G., Sánchez, C., Torruella, J. (eds.) *Filología e informática*, Barcelona, Editorial Milenio i UAB, pp.165-181.
- SPAD.T (1993). Programa de ordenador distribuido por la sociedad CISIA (1, avenue Herbillon, 94160 Saint-Mandé, Paris).
- Tournier, M. (1989): “La réduction lexicométrique”, en Bozzo i Duran, M. (ed.) *Jornades de Lexicometria*. Barcelona: P.P.U., pp. 15-28.
- Van Buuren, M.B. (1997): “Pour une approche stylistique du courant littéraire”, en Van Buuren, M.B. (ed.) *Actualité de la stylistique*. Amsterdam-Atlanta: Rodopi, pp.43-53.

