

## ¿ Ruptura poética y nuevo repertorio ? El repertorio léxico como posible síntoma crítico en la poesía castellana de Pere Gimferrer

**ELOI GRASSET**

*Université Paris – Sorbonne (Paris IV)*

*La aparición del poemario de Pedro Gimferrer Arde el mar en 1966, supuso una ruptura respecto a los cánones que imperaban por aquellos tiempos en la poesía española. Llamó la atención el descaro de aquel poeta de diecinueve años que rechazaba la realidad en la que estaba inmersa España, ausentando de su obra la memoria histórica y social más reciente del país e incluyéndose en una tradición poética esencialmente liberadora importada de otras latitudes. Pasados ya cuarenta años de la fecha de su publicación, si nos acercamos hoy al libro es porque sigue manteniéndose como el impulso próspero que fue y que supo desencadenar una decisiva poética de ruptura. En este sentido, la continuidad de su obra —y su deliberada novedad— se manifiesta, en gran medida, a través del repertorio léxico.*

Este artículo analiza el léxico de situación y su evolución en la obra escrita en español de Gimferrer, para definir de manera precisa el desplazamiento que significó respecto al léxico en el que se había asentado la poesía inmediatamente anterior y, por consecuencia, la ampliación que también supuso respecto a un imaginario que había permanecido hasta entonces sin cuestionarse. La aparición en 1966 del poemario de Pere Gimferrer *Arde el mar*<sup>1</sup>, confirmaba el empecinamiento de un poeta de diecinueve años en intentar alejarse del rumbo que llevaba marcando la poesía española desde hacía ya demasiados años. Antes de este profético libro, Gimferrer había publicado un volumen, *Mensaje del Tetrarca*<sup>2</sup> y escrito otro, *Malienus*<sup>3</sup>, no editado — parcialmente — hasta mucho tiempo después<sup>4</sup>. En todos ellos se prefiguraba ya toda su obra posterior y uno percibe aún hoy — como acertó a ver Vicente Aleixandre en su día — la decidida voluntad de cambio respecto a « una escritura gastada, y sin capacidad de reacción »<sup>5</sup>.

<sup>1</sup> P. Gimferrer, *Arde el mar*, Barcelona, El Bardo, 1966

<sup>2</sup> P. Gimferrer, *Mensaje del Tetrarca*, Barcelona, Trimer, 1963

<sup>3</sup> J.M. Castellet, (ed.), *Nueve novísimos poetas españoles*, Barcelona, Península, 2001 [1970], p. 152: « [...] estaba escrito en alejandrinos blancos, sin puntuación, presidido por la obsesión del ritmo, las imágenes, el poder sugestivo de un lenguaje extremadamente barroco y muy influido por Saint-John Perse. Era, por lo demás, escritura automática »

<sup>4</sup> Gimferrer dio a conocer una parte de *Malienus* en 1986, en la *Revista de Occidente* (nº61) y luego en el libro *Poemas (1962-1969)*, Madrid, Visor, 1988

<sup>5</sup> P. Gimferrer, *Poemas (1962-1969)*, Madrid, Visor Libros, 2000 [1969], Carta de Vicente Aleixandre, p. 94: « [...] Disuena este poema, ya se lo he dicho, de lo inmediato; pero creo que puede disonar menos dentro de unos años, y en todo caso es una disonancia saludable y advertidora. [...] yo estimo que un poema escrito así por un joven de

Aún así, fue la concesión en 1966 del Premio Nacional de Poesía por *Arde el mar*, lo que supuso la definitiva consagración pública del autor y la aceptación por parte de la crítica de una nueva poética que representaba una ruptura respecto a los cánones que imperaban por aquellos tiempos en la poesía española. Como escribió C. Barral hablando de la estética que en esa época gravitaba en la lírica del país : « los poetas acumularon a un léxico escogido un recitativo gris y monótono, parco en metáforas o con recurso, un sistema metafórico con mucha tradición literaria. »<sup>6</sup> Estos modelos, bajo los que se escribía poesía en la España de esos días, fueron descartados por P. Gimferrer activando en su lugar una serie de lecturas e influencias artísticas que no eran nada comunes. Como él mismo explica :

Mi desinterés por la poesía imperante entonces en España era completo, así como mi falta de respeto por las escalas de valor establecidas : a mí me gustaba la poesía modernista, la novela erótica de 1900, los folletines, etc. Mis poetas preferidos eran en España los del 27, a quienes debo mucho, y fuera de ella, además de Perse, T. S. Eliot y sobre todo Pound. En prosa Faulkner, Proust y Henry James. Siempre me sentí deudor del surrealismo. Todo ello no era un aspecto de mi vida, sino toda mi vida; no había otra cosa en mi vida que esto.<sup>7</sup>

Llamó rápidamente la atención<sup>8</sup> el descaro de aquel joven poeta que escapaba de la realidad en la que estaba inmersa España, ausentando de su obra la memoria histórica y social más reciente del país e incluyéndose en una tradición poética completamente distinta importada de otras latitudes. Así podemos entender la rotunda acusación de Gimferrer : « Es obvio que la mayoría de los poetas españoles no escriben nada que valga la pena ser leído »<sup>9</sup>. El principal motivo de la denuncia contra esa poesía giraba en torno a la escasa calidad del lenguaje poético, y la obstinada pretensión de comunicabilidad que pasaba por delante de la preocupación propiamente estético-literaria.

Pasados ya cuarenta años de la fecha de su publicación, si nos acercamos hoy al libro es porque sigue manteniéndose como el impulso próspero que fue y que supo desencadenar : una decisiva poética que podríamos considerar *de ruptura*. Creemos que resulta

---

diecisiete años, y ahora, en estos años, no significa retraso (como pensaba usted lo podrían estimar), sino una forma de adelanto : es, a parte de otras cosas, una protesta, una ruptura contra una escritura ya gastada, sin capacidad de reacción.»

<sup>6</sup> Barral, C., « Consideraciones acerca de las aventuras del estilo en la penúltima literatura española », *Cuadernos para el diálogo*, mayo 1969. Añade el autor en el mismo artículo : « la llamada etapa de la poesía social es, desde el punto de vista de los procedimientos literarios, una etapa lamentable »

<sup>7</sup> « Poética de Pere Gimferrer » Castellet (ed.), *Nueve novísimos poetas españoles*, p. 152.

<sup>8</sup> Ver artículos de José-Carlos Mainer [reseña de *Arde el mar*], *Ínsula*, 233 (IV/1966) Ó: p. 9-10 « [...] en Gimferrer hay un atractivo neo-exotismo cuyo rastreo de fuentes en cierta literatura europea se debe tener en cuenta por cuanto tiene de fenómeno artístico en alza. » y de Antonio Fernández Molina, *Papeles de Son Armadans*, núm. CXXV, 1966, p. 110-112 « [...] es prueba de la inquietud que hay entre estos poetas (jóvenes) que se esfuerzan, por distintos caminos, en lograr una expresión original y en salirse de los diversos dogmas que gravitan sobre nuestra poesía ».

<sup>9</sup> Poética incluida en *Nueva poesía española*, de Enrique Martín Pardo, Madrid, 1970.

hoy especialmente oportuno afrontar este estudio ya que nos asegura cierta distancia el hecho de que la lírica en España haya tomado muy otros caminos, y queda la obra en castellano de Gimferrer y particularmente *Arde el mar*, como el representante simbólico de un episodio en gran medida clausurado de nuestra historia literaria.

Partiendo de este revelador asentamiento histórico de la poesía de P. Gimferrer — del que se ha pretendido justificar aquí, sucintamente, su importancia —, se intentará observar cómo se construye esta nueva voz y cómo se distancia de la poesía vigente en la España del momento, consiguiendo enraizar unos nuevos hábitos líricos muy minoritarios hasta el momento.

Se aspirará en este artículo a dar algunas razones que justifiquen esta alteración del paradigma poético, más allá de la simple apreciación impresionista que se contenta con estimar cambios sin poder precisarlos. Intentaremos aquí hacer un estudio detallado del repertorio léxico que Gimferrer va extendiendo a lo largo de su obra poética en castellano, y lo pondremos en relación con el repertorio léxico de la generación poética inmediatamente anterior a la suya. Vamos, pues, a llevar a cabo un ejercicio de comparativismo que deberá darnos algunas razones que justifiquen lingüísticamente un cambio de gran calado que se produce en la historia de la estética poética del siglo XX en España. Evidentemente, no consideramos que el cambio del que se está hablando se agote aquí ni que el estudio del léxico sea la única manera posible de entenderlo. Pero sí que nos parece que a partir del estudio del repertorio léxico, podremos incidir en la transformación que va construyendo el poeta. Se deberá entender pues que el léxico opera como un evidente síntoma crítico — de ahí el subtítulo del artículo —, y apuntamos al papel de éste como señal, como el indicio de algo que está sucediendo y del que queda alguna traza. El léxico considerado, por tanto, particular repertorio y marca de registro que deberá dar cuenta — en cierta medida y según lo que de aquí en adelante se intentará plantear —, de una ruptura poética.

## Selección del corpus

Para plantear el ejercicio de comparativismo léxico que aquí se propone debemos antes que nada elegir el corpus con el que vamos a trabajar. En el caso de Pere Gimferrer consideraremos el corpus cerrado que conforma el libro *Poemas (1962–1969)*,<sup>10</sup> que es la reunión de los distintos poemarios en castellano del autor: *Malienus* (1962), *Mensaje del Tetrarca* (1963), *Arde el mar* (1966), *La muerte en Beverly Hills* (1967) y *De Extraña fruta y otros poemas* (1969), a los que hay que añadir algunos poemas inéditos u otros que fueron publicados en revistas y ediciones mínimas<sup>11</sup>. Por lo que respecta al corpus

<sup>10</sup> Gimferrer, *Poemas (1962-1969)*. Edición a cargo de Julia Barella.

<sup>11</sup> Hacemos referencia a los poemas *Larra*, publicado por Juan Goytisolo en una pequeña edición francesa, *Medicis*, poema inédito, y *Himno a enero*, publicado en la revista *Ínsula*, n.º 458-459, enero-febrero, 1985, p.3

de la poesía inmediatamente anterior a la de Pere Gimferrer, la elección no ha resultado tan evidente. Dado que resultaba imposible contar con un corpus que abarcara toda la poesía publicada en los años 40 y 50, se ha juzgado necesario buscar algún tipo de antología que intentara mostrar esencialmente el carácter dinámico de la evolución de la poesía española durante ese preciso período de tiempo y dentro de una muy determinada sociedad. Considerando todo lo dicho, nos ha parecido que la antología que cumplía mejor los requisitos que reclamábamos era el libro *Un cuarto de poesía española (1939-1964)*<sup>12</sup> de J.M. Castellet. De esta antología hemos escogido<sup>13</sup> el período que comprende desde el año 1958 hasta 1964, es decir la poesía inmediatamente anterior a la publicación del primer libro del P. Gimferrer en 1966. Hemos valorado especialmente, el hecho de que abarca un período lato de tiempo y no se centra en un tipo particular de poesía. El criterio de clasificación consiste únicamente en una ordenación cronológica de poemas<sup>14</sup>, dejando de lado la selección por prestigio adquirido. También nos ha parecido importante que no consistiera en una antología de *autores* sino que procurase reflejar la complejidad<sup>15</sup> y riqueza de la poesía española de esos años.

### Del vocabulario al léxico

Como cabe suponer, la lista de los vocablos que conforman tanto el corpus de las obras de Gimferrer como el de la antología de J.M. Castellet que aquí consideramos y con la que trabajaremos, no incluye evidentemente todo el léxico de sus autores<sup>16</sup>. Así, debemos entender que los vocablos que conforman los corpus que vamos a tratar, son un subconjunto del léxico que queremos comparar. Si aceptamos la clásica distinción entre *langue* y *discours*<sup>17</sup>, deberemos distinguir entre léxico — referido a la *langue* — y el vocabulario — referido al *discours* —. Habrá identidad entre léxico y vocabulario en el supuesto caso de que el autor no conozca ningún otro vocablo además de los que ha utilizado en el texto, en otras palabras, que haya actualizado en su *discours* todo

<sup>12</sup> J.M. Castellet, (ed.), *Un cuarto de siglo de poesía española (1939-1964)*, Barcelona, Seix Barral, 1973 [1965]

<sup>13</sup> Creemos hubiera sido interesante poder comparar el vocabulario de la antología entera con el vocabulario de Gimferrer. El hecho de que para el estudio de la riqueza lexical debamos trabajar con dos corpus de extensión más o menos similar, nos obliga a considerar tan sólo la poesía del período que va desde 1958 hasta 1964. La *Antología* de J.M. Castellet nos permite tomar esta sección, al estar organizado cronológicamente.

<sup>14</sup> Castellet (ed.), *Un cuarto de siglo de poesía española (1939-1964)* p.27 « [...] No habría, pues, agrupación de poemas seleccionados únicamente por un siempre peligroso (si es el único que se tiene en cuenta) criterio de calidad literaria, alrededor de unos nombres escogidos, como es habitual en estos casos, a través de un casi obligado criterio de «prestigio adquirido». Habría, solamente, una ordenación cronológica de poemas, cuya selección, empezada por tanteo, iría adquiriendo sus razones de ser, tanto en función de sus implicaciones históricas, como en la de sus peculiaridades estrictamente literarias [...] »

<sup>15</sup> *Ibid.*(ed.), p.28

<sup>16</sup> Por otro lado, el intento de construir el inventario exhaustivo del léxico individual de un autor es una utopía, y su cuantificación sólo podrá ser evaluativa y aproximativa.

<sup>17</sup> G. Guillaume, « Observation et explication », *Langage et science du langage*, Paris, Québec, 1964

su léxico virtual<sup>18</sup>. El léxico queda, por tanto, en el orden de lo virtual y el vocabulario es la actualización de este léxico en el *discours*. En este sentido, cualquier trabajo de comparativismo léxico — como el que se realizará aquí —, se efectuará a partir de un texto (*discours*) y consecuentemente cada vez que intentemos sacar alguna conclusión sobre la *langue* se tratará de una inferencia.

Analizaremos el vocabulario que conforma cada corpus y cómo éste se organiza. Lo haremos a partir del estudio comparativo de la riqueza léxica, la distribución de las categorías gramaticales o la distribución de las frecuencias lexicas de cada corpus para ponerlo seguidamente en relación con el *léxico de situación*. A través de este análisis, que nos deberá aportar interesantes resultados, el texto se convierte en la caracterización misma de una determinada actualización. Evidentemente, tanto el orden estilístico como el orden temático de la transformación léxica han de considerarse conjuntamente. Su estudio deberá servirnos para definir el desplazamiento que se habrá producido en el vocabulario y al mismo tiempo para observar como se amplía el imaginario que en ésta se había consolidado.

Empezamos con el recuento de las palabras y los vocablos que conforman cada corpus sin contabilizar ni los signos de puntuación ni los números que aparecen en los distintos textos. El número de *palabras* hará referencia al número total de ocurrencias y por tanto a la extensión del corpus; en cambio, cuando hablamos de *vocablos*, hacemos referencia a las palabras diferentes que tiene como mínimo una ocurrencia en el corpus. El corpus cerrado de los poemas de Gimferrer (A) que consideramos aquí tiene 15633 palabras ( $N_A$ ) y 4439 vocablos ( $V_A$ ). Por lo que respecta al corpus que conforma la *Antología* de Castellet (B), tenemos 16640 palabras ( $N_B$ ) y 4544 vocablos ( $V_B$ ). La extensión del vocabulario — la cantidad de vocablos de cualquier corpus = ( $V_x$ ) — es dependiente de la extensión del corpus, es decir que si consideramos dos fragmentos suficientemente homogéneos y de longitud desigual — como es el caso — cabe prever que el más largo tendrá un vocabulario más extenso que el más corto. Así pues, el hecho de que el corpus de Gimferrer tenga 4439 vocablos y el de la *Antología* de Castellet tenga 4544, no implica de ninguna manera que el léxico del corpus antológico sea más rico que el gimferreriano, ya que el resultado está condicionado por la longitud de los textos con los que se trabaja. Debemos, en consecuencia, intentar verlo de otra manera y eso pasa por hacer algunos cálculos estadísticos.

## Riqueza léxica

Vamos, ahora, a intentar comparar la *riqueza léxica* de los dos corpus considerados. Cuando hablamos de *riqueza léxica* hacemos referencia solamente a un dato técnico descriptivo y en ningún caso deberá entenderse que la riqueza léxica califica subje-

<sup>18</sup> Condición irrealizable en el ejercicio común del lenguaje.

tivamente el texto tratado. Nos dirá qué corpus de los dos utiliza mayor número de vocablos. De hecho, cuando hablamos de riqueza léxica, estamos hablando de riqueza de vocabulario por las razones comentadas anteriormente: el léxico es imposible de determinar positivamente, lo único que podemos hacer es calcular la riqueza de vocabulario y a partir de la cual obtener una estimación del léxico. Cabe no confundir, pues, esta noción con la apreciación subjetiva del lector que juzga rico o pobre un vocabulario traduciendo en esta impresión la presencia más o menos habitual de palabras juzgadas raras o poco frecuentes y la ausencia de las palabras que intervienen en mayor medida en el lenguaje habitual.

Aplicado a un texto, hablar de riqueza léxica tal y como lo hacemos nosotros aquí, consiste tan sólo en clasificar los distintos corpus considerando el número de vocablos y el número total de ocurrencias. El resultado de esta relación nos la debería dar en principio la *frecuencia teórica*<sup>19</sup>, pero entendiendo que ésta varía sustancialmente a medida que va creciendo el texto, el resultado que podemos obtener con su cálculo es inutilizable<sup>20</sup>. El único caso en el que podrá sernos útil, sería si  $N_A$  fuera igual a  $N_B$ .

Para comparar la *riqueza léxica* de los dos corpus que tenemos y a sabiendas de que tienen una extensión desigual deberemos llevar a cabo el cálculo del vocabulario teórico, fundado en el modelo binomial<sup>21</sup>. Lo que pretenderemos con la aplicación de esta ley es calcular, siguiendo principios estadísticos, cómo quedaría alterado el vocabulario — cantidad de vocablos =  $(V_B)$  — de la antología de Castellet si en lugar de tener un corpus de 16640 palabras, tuviéramos un corpus de igual extensión que el de Gimferrer, es decir de 15633 palabras. Es decir, la estadística nos permite calcular cuántos vocablos perdería el Corpus Castellet si por selección aleatoria lo redujeráramos a la medida del de Gimferrer. Gráficamente lo podemos representar de la manera siguiente :

<sup>19</sup>  $N/V$  = número teórico de ocurrencias por vocablo, considerando que todas las palabras aparecen las mismas veces en el texto. Frecuencia teórica Corpus de Gimferrer (A) = 3,52 y Frecuencia teórica Corpus Antología Castellet (B) = 3,66.

<sup>20</sup> Al principio del texto, cuando no contamos todavía ninguna repetición, sabemos que  $N = V$ . A medida que avanza el texto las repeticiones se van sucediendo y entonces  $N$  aumenta mucho más rápido que  $V$ . Cada vez resultará más difícil de encontrar un vocablo nuevo que no se haya ya encontrado en el texto. Por lo tanto la frecuencia teórica se desplaza cada vez más de la frecuencia real de cada vocablo. Trabajando con un número tan grande de palabras, resulta imposible considerar el valor de la frecuencia teórica como un dato significativo para establecer diferencia de riqueza léxica entre los dos textos.

<sup>21</sup> C. Muller, *Principes et méthodes de statistique lexicale*, Paris, Champion, 1992 [1977], p. 101

A = Corpus Gimferrer	_____ $N_A = (15633)$
B = Corpus Castellet	_____ - - - - - $N_B = (16640)$
B' = Corpus reducido Castellet	_____ $N_{B'} = N_B - (N_B - N_A) =$ $16640 - (16640 - 15633) = 15633$

Para poder aplicar la ley binomial, necesitamos saber en primer lugar la probabilidad ( $P_B$ )<sup>22</sup> de que una ocurrencia de B figure también en la reducción de B, es decir B'. Y tenemos que  $P_B = 15633 / 16640 = 0.9395$ . Lo que nos indica que es muy probable que una ocurrencia escogida al azar forme parte también del Corpus Castellet reducido. Necesariamente su probabilidad complementaria  $Q_B$ , — la probabilidad de que una ocurrencia escogida al azar no forme parte del corpus reducido — es muy baja.  $Q_B = 0,0605$ . Necesitamos también saber el número exacto de frecuencias de los vocablos del corpus que queremos reducir: esto es, saber cuántos vocablos aparecen una sola vez, cuántos dos veces, etcétera. Para el corpus de *Castellet* tenemos los resultados siguientes:

Vocablos con frecuencia 1	$V_1 = 3145$ i.e. 69,21%
Vocablos con frecuencia 2	$V_2 = 604$ i.e. 13,29%
Vocablos con frecuencia 3	$V_3 = 222$ i.e. 4,89%
Vocablos con frecuencia 4	$V_4 = 143$ i.e. 3,15%
Vocablos con frecuencia 5	$V_5 = 73$ i.e. 1,61 %
Vocablos con frecuencia 6	$V_6 = 53$ i.e. 1,12%
Vocablos con frecuencia 7	$V_7 = 40$ i.e. 0,88%
Vocablos con frecuencia > 7	$V_{>7} = 264$ i.e. 5,81%

Con los datos que hemos recopilado hasta ahora podemos calcular la ley binomial y saber cuál será la estimación del número de vocablos de B' = selección corpus de

<sup>22</sup> Índice que toma valores entre 0 y 1. Cuanto más cerca de 1, más probabilidades hay de que cualquier ocurrencia de B tomada al azar forme parte también de la reducción.

Castellet. Siguiendo la fórmula siguiente:  $E(V')$  (estimación de vocablos de B') =  $V_B - \sum q^i \cdot Vi$

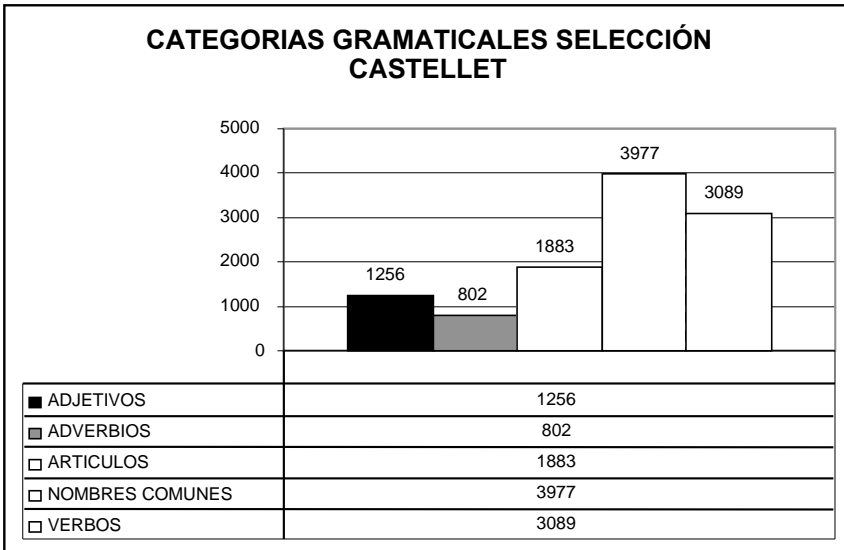
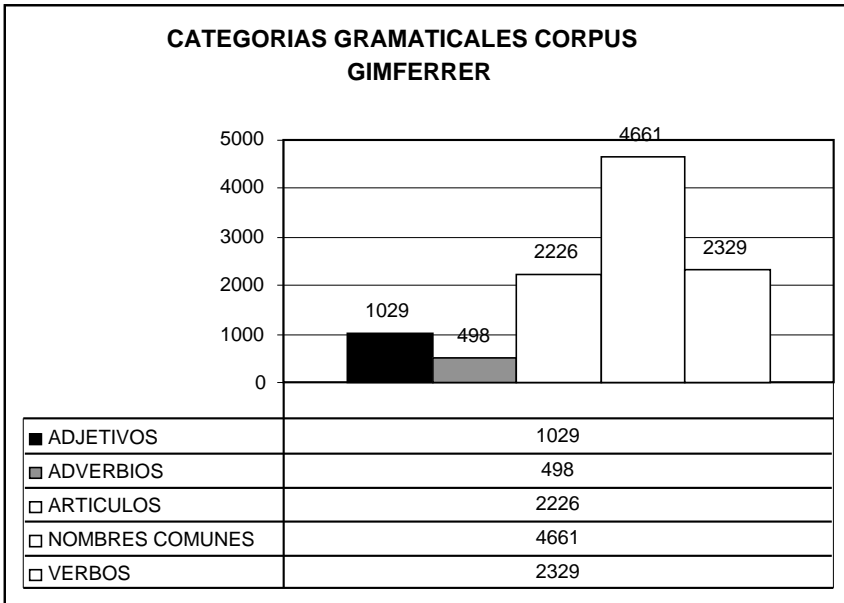
i	$V_i$	$q^i$	$q^i \cdot V_i$
1	3145	0,06	188,7
2	650	0,0036	2,34
3	222	0,00022	0,048
4	143		
5	73		
6	53		
7	40		

Podemos parar los cálculos porque  $q^3 \cdot V_3$  es inferior a 0,05. Así pues, tenemos que  $E(V') = 4544 - (188,7 + 2,34 + 0,048) = 4544 - 191,1 = 4352,9$ , que es el número de vocablos existentes en la estimación de la reducción del corpus de Castellet (B). Nos falta calcular la desviación de los resultados obtenidos. De momento tenemos que, si elegimos al azar 15633 palabras de las 16640 del corpus de Castellet, escapan de la selección 191,1 vocablos y encontraremos  $4544 - 191,1 = 4352,9$ . Así pues, la probabilidad de que un vocablo no esté es de  $191,1 / 4544 = 0,0420$ , y la probabilidad de encontrarlo es de 0,9579. Así pues tenemos que  $4439 > 4352,9 \Leftrightarrow V_A > V_B$ . Ahora ya podemos afirmar que el corpus de Gimferrer es lexicalmente más rico que el corpus de Castellet. Es decir, que la riqueza léxica de A es superior a la de B, lo que nos advierte que en el corpus de Gimferrer se utiliza más cantidad de vocablos que en el corpus de la antología de Castellet que hemos definido anteriormente. Aún así, cabe decir que la diferencia no es altamente significativa como para considerarla una prueba del cambio estilístico que se produce. La riqueza léxica es superior pero no significativamente.

### Categorización gramatical

Siguiendo con el intento de establecer el cambio del que hemos hablado anteriormente, intentaremos observar ahora cómo se organizan las ocurrencias respecto a las categorías gramaticales, para poder determinar algún cambio en el uso dentro de los dos corpus. Nos centraremos tan sólo en cinco categorías que se presentan a priori como las más significativas: *adjetivos*, *adverbios*, *artículos*, *nombres comunes* y *verbos*. Si hacemos una gráfica de ambos corpus considerando solamente estas cinco categorías, tenemos lo siguiente:





Cabe decir antes que nada que seguimos trabajando con los corpus iniciales, es decir que  $N_A$  = Corpus de Gimferrer consta de 15633 palabras y  $N_B$  = Corpus de Castellet consta de 16640 palabras. Observamos que ambos corpus se distribuyen de la misma manera. De mayor a menor cantidad de ocurrencias tenemos: *nombres comunes*, *verbos*, *artículos*, *adjetivos* y *adverbios*. Aún así, destaca en el corpus de Gimferrer

el porcentaje de *nombres* (4661 ocurrencias - 29,8% del corpus) que se eleva muy por encima del resto de categorías y también la similitud en la cantidad de *adjetivos* (1029 ocurrencias - 14,2%) y *verbos* (2329 ocurrencias - 14,9%). En lo que concierne el corpus de Castellet, encontramos tan sólo 3977 ocurrencias de *nombres comunes* (23,9% del corpus), muy por debajo de las ocurrencias del texto de Gimferrer aún contando con 1007 palabras más. Es decir, que el corpus de Gimferrer cuenta proporcionalmente (y también en términos absolutos) con muchos más *nombres comunes* que el de Castellet y de esta manera podemos asegurar que en la categorización del repertorio léxico desde luego se produce un cambio importante. También destaca que la cantidad de verbos de  $N_B$  (3089 ocurrencias - 18,6%) es mucho mayor que el de  $N_A$ . Dado que los *adverbios* son los caracterizadores más habituales del *verbo*, es lógico que aparezcan más adverbios que en el corpus de Gimferrer, así que no resulta significativo. El conjunto de estos resultados nos da un primer indicio, que haría falta explorar a través del análisis de los campos semánticos asociados a los verbos y sustantivos, que podríamos proyectar a través de la caracterización siguiente : el corpus Castellet está vinculado a la acción/movimiento y el de Gimferrer a la descripción.

### Diferencia y repetición

Siguiendo la traza ofrecida por estos resultados, vamos a centrarnos ahora en el análisis de los vocablos más frecuentes de cada corpus. La experiencia nos demuestra que, en todo idioma, los vocablos que tienen como mínimo una ocurrencia cada 100 palabras, es decir que tienen una frecuencia relativa mayor o igual a 0.01, son muy pocos. Se trata de palabras relacionales, preposiciones, conjunciones, determinantes, pronombres... Estos vocablos son estables estadísticamente en el sentido que no hay variaciones importantes en corpus distintos. Así pues, un pequeño número de unidades léxicas conforman la mayor parte de todo discurso, y se estima que las 50 unidades utilizadas más frecuentemente cubren un 50% del texto. A parte de este grupo de vocablos repetidos, existe otro de vocablos dispersos que incluye vocablos raros y específicos. Se construye a partir de esta distinción la dialéctica entre diferencia y repetición : el vocablo anómalo que aparece tan sólo una vez precisamente por el hecho de aparecer tan sólo una vez, puede adquirir más importancia que la palabra más frecuente. Veamos qué es lo que sucede con nuestros corpus. Si miramos los 20 vocablos más utilizados nos encontramos con lo siguiente :

**Lista Corpus Gimferrer<sup>23</sup>**

**Lista Corpus Castellet<sup>24</sup>**

<i>Nº Vocablo</i>	<i>Frec. Abs</i>	<i>Nº Vocablo</i>	<i>Frec. Abs</i>
1 DE	879	1 DE	741
2 EL	608	2 Y	628
3 EN	581	3 LA	600
4 LA	580	4 EL	516
5 Y	455	5 QUE	441
6 LOS	350	6 EN	391
7 QUE	288	7 A	274
8 LAS	270	8 LOS	262
9 UN	248	9 UN	198
10 DEL	200	10 LAS	174

**Lista Corpus Gimferrer (cont.)**

**Lista Corpus Castellet (cont.)**

<i>Nº Vocablo</i>	<i>Frec. Abs</i>	<i>Nº Vocablo</i>	<i>Frec. Abs</i>
11 A	188	11 NO	173
12 COMO	156	12 CON	150
13 UNA	144	13 ES	134
14 MI	131	14 DEL	126
15 NO	131	15 SE	124
16 SE	105	16 SU	117
17 QUÉ	99	17 ME	108
18 AL	90	18 MI	98
19 SU	90	19 COMO	97
20 O	88	20 AL	96

Donde *Nº* indica la posición dentro de la lista de vocablos ordenados según la cantidad de ocurrencias y *Frec. Abs* (i.e. frecuencia absoluta), nos indica el número absoluto de ocurrencias dentro del corpus. Como podemos observar, no hay diferencias destacables entre las dos listas de vocablos utilizados más frecuentemente. Indiquemos sólo que ambas listas están formadas íntegramente por palabras relacionales: preposiciones, determinantes, conjunciones, relativos... y que no encontramos ningún vocablo que nos pudiera dar una

<sup>23</sup> Lista de los veinte vocablos con frecuencias más altas — palabras utilizadas más veces — que aparecen en el corpus de Gimferrer.

<sup>24</sup> Lista de los veinte vocablos con frecuencias más altas — palabras utilizadas más veces — que aparecen en el corpus de la *Antología* de Castellet.

idea de las diferencias de vocabulario entre los dos corpus. Si lo que queremos es buscar los vocablos que puedan ser significativos semánticamente, y que por tanto nos permitan diferenciar los dos corpus, deberemos llevar a cabo la lematización de cada corpus, es decir conseguir para cada lexema de cada corpus el número de ocurrencias seguido del lema que lo representa. Esto se podrá obtener gracias a la reunión de las diferentes formas de flexión del lema. Antes de comentar los resultados, debemos explicar lo indispensable que resulta realizar este ejercicio comparativo de los sustantivos semánticamente significantes a partir de la lematización. Si tomamos por ejemplo el lema OJO, consideramos al contar el número total de ocurrencias las variantes de flexión del lexema, en este caso el uso del plural: OJOS. Cabe considerarlo así dado que estamos haciendo una aproximación comparativa al léxico y resultaría caer en un error contar por separado las ocurrencias de OJO y OJOS, ya que no darían buena cuenta de lo que realmente aparece en el texto.

**Corpus Gimferrer<sup>25</sup>**

**Corpus Castellet<sup>26</sup>**

*Nº Lema Frec. Abs Nº Lema Frec. Abs*

24 LUZ	77	31 VIDA	60
25 NOCHE	71	32 HOMBRE	60
26 OJO	70	39 TIERRA	48
32 MAR	52	40 MANO	48
36 AMOR	45	41 DÍA	46
38 TIEMPO	42	44 TIEMPO	41
39 HOMBRE	40	45 VEZ	40
40 VIENTO	33	48 LUZ	37
41 MANO	32	50 PALABRA	32
45 MUERTE	31	52 OJO	30
48 SANGRE	29	53 NOCHE	30
49 SILENCIO	29	54 AÑO	30
50 AIRE	28	55 AIRE	29
51 VIDA	28	56 AMOR	28
52 SOMBRA	28	57 COSA	28
53 AGUA	27	69 SOL	24
56 JARDÍN	26	79 CASA	21
68 VOZ	26	86 MUNDO	20
64 LLUVIA	22	91 MUERTE	19
65 CIELO	22	92 SANGRE	19

<sup>25</sup> Lista de los veinte sustantivos utilizados más frecuentemente en el corpus de Gimferrer.

<sup>26</sup> Lista de los veinte sustantivos utilizados más frecuentemente en el corpus de Castellet.

Para este ejercicio tendremos sólo en cuenta los sustantivos de cada corpus, ya que a nuestro parecer puede servirnos como un indicador fidedigno del cambio de léxico del que pretendemos dar parte.

Para poner otro ejemplo de la importancia de la lematización, si efectuáramos el mismo ejercicio considerando separadamente el conjunto de ocurrencias que se incluyen bajo el lema HOMBRE (i.e. HOMBRE y HOMBRES) del corpus de la *Antología* de Castellet, veríamos que HOMBRE aparece en la posición nº43 con 42 ocurrencias donde  $N^o$  indica la posición en una lista de vocablos ordenados según la cantidad de ocurrencias. El vocablo HOMBRES no aparecería en la lista que acabamos de dar, pues contaría tan sólo con 18 ocurrencias y no formaría parte de esta lista de los 20 vocablos más frecuentes. Por el contrario, considerando conjuntamente las ocurrencias de HOMBRES y HOMBRE, contamos con 60 ocurrencias y por lo tanto pasa a ser junto con VIDA el vocablo más usado de los sustantivos que hemos considerado significativos del corpus de Castellet. Vemos, pues, que resulta indispensable considerar la lematización de los dos corpus. Pasamos ahora al examen de los resultados obtenidos.

Hemos considerado exclusivamente sustantivos, dado que siendo unidades lexicales que designan autónomamente una cosa o noción, nos permitirá hacer hincapié en la carga significativa de estos vocablos y en los campos semánticos dentro de los cuales los podríamos considerar. Para poder hacer un análisis comparativo de la diferencia de corpus, hemos elegido los veinte primeros lemas de cada corpus. Por lo que respecta al Corpus de Gimferrer (A), el primer lema sustantivo — LUZ — aparece en vigésimocuarta posición y posee una frecuencia lematizada<sup>27</sup> de 0,0049 (es decir que cada 100 palabras del corpus, el lema LUZ bajo cualquiera de sus formas aparece 0,49 veces). En cuanto al corpus de Castellet (B), no encontramos un vocablo significativo hasta la posición 31, y tiene una frecuencia lematizada de 0,0036 (es decir que para cada 100 palabras del corpus, el lema VIDA bajo cualquiera de sus formas aparece 0,36 veces).

Lo que llama primero la atención es que de la lista de los veinte lemas sustantivos que más aparecen en A y B, once de ellos se repiten en ambos (LUZ, NOCHE, OJO, AMOR, TIEMPO, HOMBRE, MANO, MUERTE, SANGRE, AIRE, VIDA). Así pues, lo que habíamos dado por supuesto al empezar el artículo, es decir que el léxico de los dos corpus que queríamos comparar era muy distinto, no es del todo cierto y se tambalean así las suposiciones — hecho que justifica sin lugar a dudas la metodología empleada y el artículo que aquí se presenta — al menos por lo que respecta a los veinte primeros lemas sustantivos.

Aún así, cabe decir que los lemas repetidos ocupan distintas posiciones dentro de la lista: por ejemplo VIDA es el lema sustantivo que más encontramos en el corpus de Castellet y en cambio en el corpus de Gimferrer ocupa la posición 14. Al contrario, LUZ y NOCHE ocupan el primer y segundo lugar de la lista de Gimferrer y en cambio ocupan las posiciones 8 y 11

<sup>27</sup> Frecuencia lematizada = frecuencia lematizada absoluta / Número total de vocablos ( $V_A$ )

de la selección de Castellet. De los primeros 10 lemas de la lista de Gimferrer, los únicos que no se encuentran en la lista de Castellet son los lemas MAR y VIENTO.

Hemos estudiado los sustantivos de frecuencia máxima, y no aparecen diferencias significativas que nos permitan efectuar la caracterización isotópica de los sustantivos del corpus. Observaremos ahora, siguiendo con el mismo ejercicio de comparativismo que acabamos de hacer, cómo se comportan los vocablos de frecuencia 1. La cantidad de estos vocablos puede ser tratada como un signo estilístico diferenciador y deberá ser puesta en conexión con la riqueza léxica que hemos analizado anteriormente y con el léxico de situación. Si contamos los vocablos de frecuencia 1 del corpus de Gimferrer, tenemos que ocupan el 68,69% (3049 vocablos) del total (4439 vocablos). Si del conjunto, prestamos atención tan sólo a los *nombres comunes* – categoría gramatical que nos puede dar la información más significativa –, tenemos que hay 1554 *nombres comunes* de frecuencia 1, – un 35,01% de los vocablos utilizados en el corpus de Gimferrer, sólo se utilizan una vez –. Cabe destacar de estos cálculos que el lematizador no es capaz de determinar el lema de 980 vocablos por no contemplarlos en el lexicón. Éstos representan el 63% del total de sustantivos de frecuencia 1 del corpus de Gimferrer, hecho significativo pero en ningún caso problemático para la consecución del análisis. Entendemos que estos vocablos que no aparecen en el lexicón son vocablos anómalos<sup>28</sup>, y por lo tanto son significativos estilísticamente.

Por lo que respecta a los vocablos de frecuencia 1 del corpus de la *Antología* de Castellet y considerando todas las categorías gramaticales, contamos 3145 vocablos, es decir el 69,21% del total (4544 vocablos), tan sólo un poco superior al porcentaje que encontrábamos en el corpus de Gimferrer. Si prestamos atención a los *nombres comunes* de frecuencia 1 encontramos 1076 unidades, un 23,68% de los vocablos totales que contempla el Corpus de Castellet. De estas unidades, el lematizador no es capaz de determinar el lema de 506 vocablos por no contemplarlos en el lexicón – 47% de los nombres comunes de frecuencia 1 del corpus de Castellet –. Así pues, hay una diferencia cuantitativa importante entre los *nombres comunes* de frecuencia 1 de los dos corpus – un 35,01% de todos los vocablos y en el caso de Castellet, un 23,68% –, una diferencia muy significativa del orden del 11,33%. Dentro de los *nombres comunes* de frecuencia 1 de cada corpus, tenemos que en el caso del Corpus de Gimferrer 980 (63,06%) de ellos no están contemplados en el lexicón. Como hemos comentado, se trata de vocablos anómalos que suponen una marca evidente de estilo. En el caso del corpus Castellet encontramos que 506 (47,02%) *nombres comunes* de frecuencia 1 no están contemplados en el lexicón. Tenemos, pues, una diferencia del orden del 16,04%, también significativa estilísticamente.

---

<sup>28</sup> Ver p. 10.

## Conclusiones

La intención del artículo era detectar efectivamente un cambio importante en la evolución de la poesía española para, en segundo lugar, intentar establecer los elementos que lo habían determinado. Un trabajo, que partía de la detección impresionista de un síntoma e intentaba establecer de forma precisa aquello que lingüísticamente estaba en juego.

A parte de algunos resultados ya esperados, también se han obtenido otros que contradicen sustancialmente la idea presupuesta y que por esa misma razón resultan interesantes y justifican el hecho de que les prestemos atención. Hemos empezado por comparar la riqueza léxica de cada corpus y aunque esperando una diferencia importante, hemos visto que el corpus de Gimferrer era más rico léxicamente pero que de ninguna manera esta diferencia era significativa. En segundo lugar, hemos observado la distribución de la categorización gramatical y hemos empezado a precisar qué elementos intervenían significativamente en este cambio. Lo más importante ha sido observar la muy distinta cantidad de verbos y sustantivos que tienen ambos corpus: El corpus de Gimferrer posee muchos más sustantivos y el de Castellet muchos más verbos. Por último, empujados por estos resultados, nos hemos centrado en el análisis de las frecuencias, para ver cómo, aunque los campos semánticos que ocupan los sustantivos más utilizados parecen ser los mismos, sí que hay significativas diferencias entre los dos corpus por lo que respecta al número de sustantivos de frecuencia 1 y al número de sustantivos no contemplados en el lexicón con el que trabajamos. El corpus de Gimferrer utiliza muchos más *nombres comunes* de frecuencia 1 que el corpus de Castellet y además el tipo de vocabulario utilizado por Gimferrer es un vocabulario que podríamos considerar anómalo respecto al vocabulario incluido en el corpus de Castellet y, por lo tanto, distanciado de la lengua común. En este sentido y tal como esperábamos a partir de la impresión que nos produjo su lectura, el léxico del corpus de Castellet se acerca más a la lengua hablada, cumpliendo así su deliberada misión comunicativa, mientras que el léxico de Gimferrer tiende a alejarse de esta lengua y construye su literatura a partir de criterios puramente estéticos.

Tal como hemos dicho al principio, el cambio del que se está hablando no se agota aquí, ni creemos que el estudio comparativo del léxico sea la única manera de afrontarlo. Lo que sí nos parece, en cambio, es que el ejercicio minucioso de lexicometría nos ha dado fértiles resultados que no podríamos haber obtenido de ninguna otra manera. Deberíamos continuar ahora con la comparación de los distintos campos semánticos que ocupa el léxico de los dos corpus y las frecuencias de uso de los distintos vocablos o con el estudio de la conexión léxica, pero estos son aspectos que escapan del objetivo primero de este artículo. Por lo que respecta al tema tratado ha quedado pues de manifiesto cómo el cambio estético que se produce desencadena una serie de transformaciones en el repertorio léxico, pero quizá no de la manera que la intuición nos sugería. Y esto era precisamente lo que aquí se pretendía establecer, lo que aquí se quería determinar.

## Bibliografía

- J. M. CASTELLET, (ed.), *Un cuarto de siglo de poesía española (1939-1964)*, Barcelona, Seix Barral, 1973 [1965]
- , (ed.), *Nueve novísimos poetas españoles* [1970], Barcelona, Península, 2001
- C. BARRAL, « Consideraciones acerca de las aventuras del estilo en la penúltima literatura española », *Cuadernos para el diálogo*, mayo 1969.
- GIMFERRER, Pere, *Poemas (1962-1969)* [1969], Madrid, Visor Libros, 2000
- , *Arde el mar* [1966], Madrid, Cátedra, 1997
- J. MOLINO, et J. GARDES-TAMINE, *Introduction à l'analyse de la poésie I - Vers et figure*, I [1982], Paris, PUF, 1992,
- G. MOLINIÉ, *Éléments de stylistique française*, Paris, PUF, 1986
- CH. MULLER, *Principes et méthodes de statistique lexicale* [1977], Paris, Champion, 1992
- , *Initiation à la statistique linguistique*, Paris, Larousse, 1968.