

EL ESTUDIO DEL LÉXICO LATINO: ENFOQUES ESTADÍSTICOS Y LEXEMÁTICOS

0. INTRODUCCIÓN

Cuando centramos la atención sobre los materiales desarrollados para aprender latín desde la misma Antigüedad, llama nuestra atención el esfuerzo que se ha dedicado al estudio de las *Artes gramaticales* frente a la casi dejadez en que ha quedado la investigación de una de las herramientas más populares de la enseñanza, el vocabulario. *Artes* como la de Donato o Prisciano son verdaderos manuales de enseñanza de la lengua latina, no siempre comprensibles para los que se inician, demostrado por el hecho de que algunos se editaran en versión escolar. En cambio, las distintas obras lexicográficas conservadas parecen subordinadas a la gramática, mucho más unidas a niveles inferiores de enseñanza de la lengua. De hecho, en la Edad Media empezarán a aparecer glosarios como apéndice a obras gramaticales y algunas pueden considerarse prácticamente cuadernos de estudio y aprendizaje (Rico 1978).

Sin embargo, el estudio de todo este material lexicográfico suministra datos sobre temas tan importantes como la relación entre lenguas vernáculas y latín, el grado de conocimiento de latín de los usuarios de estos útiles o los recursos léxicos de los que podían disponer los traductores. La elaboración de estas obras varía desde las simples glosas interlineales a los grandes diccionarios organizados alfabéticamente y depende en gran medida de la utilidad para la que hayan sido creados, unas veces para glosar una palabra latina desconocida por el lector, otras para la simple consulta o como fuente de sinónimos para variar la expresión, según opina Díaz y Díaz (1978).

En el presente artículo, tras un breve repaso por la finalidad pedagógica de algunos de estos vocabularios, exponemos dos de los criterios que han guiado la confección de algunos del siglo XX, la *fre-*

cuencia y la *dispersión*, así como una descripción de los mismos y de algunos diccionarios de este tipo disponibles para el latín. Si bien es cierto que pueden suponer una gran ayuda para la didáctica de la lengua, no lo es menos que tal como se presentan en forma de listas exigen un trabajo posterior para insertar las palabras frecuentes en una programación. Esta dificultad se ve en los propios libros de texto, que, por un lado, apenas siguen las conclusiones de estudios tan generales como los de Guiraud sobre la frecuencia de las palabras en un texto y, por otro, generalmente no aclaran los criterios de introducción del léxico en el manual.

No obstante, este tipo de obras estadísticas ofrece una base empírica para seleccionar el léxico didáctico y ha contribuido junto con otras investigaciones, como las de Meara (1980), Lucas (1986) o Laufer (1992), a revalorizar el papel del vocabulario en el aprendizaje de la lengua. Frente al aprendizaje mnemotécnico de listas de palabras, creemos que la *lexemática*, tal como la han abordado García-Hernández y algunos de sus discípulos, ayuda no sólo a memorizar la palabra, sino a algo más importante: comprender las relaciones de significados que se producen en la lengua.

1. FINALIDAD DE LOS LÉXICOS LATINOS

Ya desde la Antigüedad queda clara la importancia que tuvieron las listas de palabras en el aprendizaje de las lenguas, en concreto, los llamados *Nominalia*, como los *Hermeneuta Pseudo-Dositheana* del siglo III, que se configuran como un complemento de los *Colloquia*. Éstos eran verdaderos modelos de conversaciones muy estereotipadas, utilizados en la Edad Media y todavía en el Renacimiento, como prueban las seis reediciones que tuvieron entre 1582 y 1594 los *Colloquia sive Exercitatio Latinae Linguae Johannis Thomae Fregii notis illustrate* de Juan Luis Vives.

A partir del siglo VIII empiezan a proliferar con una doble finalidad las listas de palabras (Buridant 1986:12): por un lado para ayudar a la comprensión de textos antiguos, cuyo sentido había sido oscurecido por el tiempo y por otro, como obra de consulta para mejorar el estilo al facilitar la búsqueda de sinónimos. Sobre la primera utilidad se tiene claramente constancia de glosas interlineales incluso en textos destinados a enseñar gramática (*Oratio de utensilibus ad domum gerendam* de Adam du Petit Pont y el *Dictionarius* de

Jean de Garlande del siglo XII), lo que indica claramente que los propios maestros conocían mal la lengua que enseñaban (Owen 1929:9). Estas glosas adoptarán unas veces orden alfabético, otras se agruparán en campos nocionales (*Dictionarius* de Jean de Garlande) a fin de facilitar la memorización, método que hay que poner en relación con la composición en verso de obras de enseñanza, como el *Graecismus* de Evrard de Béthune o el *Doctrinale puerorum* de Alejandro de Villadei. En España, Americo Castro editó tres glosarios bilingües alfabéticos del siglo XIV, que serían a su juicio «cuadernos de estudio y aprendizaje» (Castro 1991:7), donde se aprecia claramente que la finalidad práctica del glosógrafo es aprender latín.

También en el Renacimiento tendrá lugar esta doble finalidad, de forma que se elaboran gramáticas con apéndice de palabras, como la de Gutiérrez de Cerezo, o listas independientes, cuya finalidad es más bien el cuidado del estilo, como la obra de Valla (1540). En el caso de Cerezo el *didactismo* es mayor si cabe, pues aunque compone un diccionario monolingüe, aproximadamente el 13% de las definiciones se hace con palabras españolas, como en el ejemplo siguiente de su *Ars Grammatica* (Gutiérrez Galindo 1996) de 1497:

Aggredi: est invadere, 'acometer'.

Alapa: id est, colaphus, 'bofetada'.

Este tipo de composición de gramáticas con listas de palabras ajenas podría considerarse como un antecedente de nuestros libros de texto de latín. En todos ellos la presencia de una lista de vocabulario al final de libro es casi una obligación. Si bien en los de Nebrija y Cerezo servían para aclarar palabras oscuras, en los actuales, concebidos desde otra práctica pedagógica, hay que entenderlos como ayuda para la realización de los distintos ejercicios (traducción, declinación, conjugación). Así pues, la simple aclaración de palabras latinas para ayudar en la traducción de los textos ha pervivido como única finalidad.

2. LA SELECCIÓN DE UN VOCABULARIO DIDÁCTICO

Ahora bien, ¿cuáles han sido los criterios con los que se han elaborado estos vocabularios? Evidentemente, los criterios de selección dependerán en gran medida del fin al que estén destinados. En las

listas que acompañan a las gramáticas antiguas se incluyen las palabras dependiendo de la motivación gramatical, muchas de ellas transmitidas por la tradición pedagógica, y aparecen bastantes términos que se podrían considerar de *realia*, sobre todo nombres propios (*Helena, Hector, Erato, Erebus*).

Comenio (1657, 1637) introduce una novedad al decidir adecuar la enseñanza al nivel del alumno, de manera que se aprende la lengua latina a partir de las cosas conocidas, lo que supone ya un primer criterio de introducción del léxico en un libro de texto, aunque sea de tipo psicológico. Los criterios que utilizan tanto él como el jesuita Bathe están basados en el *uso* y la *frecuencia*, pero no hay que entender éstos como en la actualidad. Con *uso* se refieren a la palabra que más se utiliza al hablar sobre un tema, lo que Michéa (1953) llama *disponible* (cf. Gougenheim 1970), mientras que frecuentes serían las palabras que más se usan en la vida diaria o en la conversación de quienes se estima. En cierto modo, este concepto de frecuencia ya considera la estadística como criterio de selección del vocabulario didáctico, aunque en lugar de mirar hacia los textos latinos, mira hacia las personas que hablan latín.

2.1. *La frecuencia*

Puesto que la elaboración de listas está muy en relación con la finalidad práctica, la existencia de un canon de autores y la consideración de la traducción como ejercicio fundamental supuso desde principios del siglo XX reducir la frecuencia de una palabra en latín a la frecuencia de esta palabra en determinados autores. Así, en la lista de Lodge (1907) se hace el despojo de palabras frecuentes basándose en textos de César (*Gall.*I-IV), Cicerón (*Cat.*, *Pomp.*) y Virgilio (*Aen.* I-IV). El vocabulario de Mathy (1952), por poner un ejemplo de mediados de este siglo, se basa en César (*Gall.*), Cicerón (*Cat.*, *Mur.*, *Verr.*, *Senect.* y fragmentos de las cartas), Salustio (*Cat.*), Nepote (*Paus.* y *Them.*), Virgilio (fragmentos de *En.* y *Ecl.*), Horacio (*Od.*, *Sat* y *Ep.*), Ovidio (*Met.*), Livio (fragmentos), Séneca (*Ep.*) y Tácito (*Ann.*). Con estas obras elabora una lista con las 2.313 palabras más frecuentes agrupadas alfabéticamente y por orden decreciente hasta una frecuencia mínima de 4. Además, añade 280 palabras que considera gramaticales, tomadas del *Index verborum Ciceronis* de Olfather, Canter y Abbot. Este vocabulario supone un salto cualita-

tivo importante respecto al de Lodge, pues la propia selección incluye una crítica, ya que es inútil confeccionar una lista de palabras frecuentes de una lengua si no existe una muestra suficientemente representativa de la población, y la selección de Lodge apenas lo es. En cambio, en cuanto a volumen de ocurrencias el vocabulario de Mathy sí es representativo, aunque pueden plantearse dudas acerca del latín representado, pues prácticamente está basado en obras literarias específicamente históricas. Por esto la palabra *bellum* figura con el rango 19 por orden decreciente de frecuencia, por delante de *pars*, *corpus* y *manus*.

El Laboratoire d'Analyse Statistique de Langues Anciennes dirigido por Delatte comenzó su andadura en los años sesenta muy unido al avance de la informática, con el fin de hacer un despojo completo de todos los textos clásicos; así, además de estudios estadísticos de cada autor, en el año 1974 dos colaboradores, Govaerts y Denooz, publicaron el *Dictionnaire de base latin*, basado en el análisis de 323.085 ocurrencias de obras de Catulo, César (*Gall.*), Cicerón (extractos), Salustio (*Cat.*, *Iug.*), Virgilio (*Aen. I-IV*), Horacio (*Od.*), Fedro (extractos), Ovidio (extractos), Tito Livio (extractos), Séneca (*Ep.*) y Tácito (extractos). Salvando el problema de la representatividad del latín que reflejan, este léxico tiene la originalidad de ofrecer una lista con las palabras que aparecen en todos los autores, a la que llama *léxico de base* (144 en total), y una serie de microléxicos en que figuran las palabras más utilizadas por cada autor además de las del léxico de base. También se ofrecen las curvas de distribución del vocabulario, en las que puede apreciarse que con un vocabulario de unas 300 palabras así seleccionadas se puede alcanzar una tasa de comprensión de un texto de aproximadamente el 80%, algo exagerado si se tiene en cuenta que las 1.000 primeras palabras más frecuentes forman el 85% de un texto (Guiraud 1960:93), lo que no quiere decir que se comprenda el mismo en este porcentaje, pues la tasa de comprensión para esta misma cantidad sería de un 50% de la información del texto.

Sin embargo, el equipo de LASLA no se detuvo aquí, pues logró publicar un nuevo diccionario de frecuencia en 1981, el *Dictionnaire fréquentiel et index inverse de la langue latine*, en el que hay un incremento cuantitativo de las ocurrencias analizadas y un equilibrio mayor entre las obras seleccionadas. En efecto, la muestra está formada por 794.662 ocurrencias, de las que un 73% corresponden

a prosa y un 27% a verso, y entre los escritores se incluyen extractos de Vitrubio, además de Catulo, César (*Gall.*), Cicerón (*Cat.*, *Nat.*, *Pomp.*, *Rab.*, *Tusc.*), Horacio (*Od.*), Juvenal, Ovidio (extractos de *Met.*), Persio, Propercio, Quinto Curcio, Salustio, Séneca (obra filosófica y tragedias), Tácito (*Agr.*, *Ann.*, *Dial.*, *Germ.*), Tibulo, extractos de Tito Livio y Virgilio (*Ecl.*, *Ge.*, *Aen. I-VI*). Aunque sigue habiendo una sobrerrepresentación de la historiografía, se empiezan a introducir obras que pueden reflejar otro latín menos literario. Además, tengo constancia de que la lista de autores despojados estadísticamente sigue aumentando, con lo que es previsible una reedición aumentada de este diccionario (Purnelle 1995). Hasta la fecha, es el último gran diccionario de frecuencia que poseemos los latinistas con la ventaja, además, de incluir también un índice inverso de todas las palabras que contiene, aproximadamente unas 13.000.

Así pues, parece constatar una tendencia en estos cuatro vocabularios a la ampliación del canon de autores seleccionados, motivado en parte por el despojo exhaustivo de los clásicos, en parte por un acercamiento más comprensivo hacia los autores posclásicos y técnicos. No obstante, hay un criterio que ha guiado siempre este tipo de obras, como es su funcionalidad en la práctica docente, lo que les ha hecho centrarse en los autores que en cada país se exigían en el programa oficial de estudios.

Esta motivación, obviando cualquier otra más científica, es la que han tenido otros vocabularios. Entre nosotros tenemos el *Vocabulario básico* de Echave-Sustaeta (1953), basado en el vocabulario de Wormald *et alii* (1949), que contiene un vocabulario general con 2.605 palabras, una segunda parte con una lista de palabras gramaticales clasificada por partes de la oración y otra en que se dan alfabéticamente los sustantivos, adjetivos y verbos más frecuentes; finalmente, hay una tercera parte con agrupamientos de las palabras según esferas léxicas y una lista con algunos sinónimos básicos latinos que se refieren a una palabra española. Su finalidad es absolutamente pedagógica, por lo que no contiene ninguna indicación sobre el rango de frecuencia de las palabras incluidas. Parecido al anterior, aunque con mayor información gramatical en cada entrada, es el libro de Podvin (1981), *Les mots latins*. Éste se sirve de los diccionarios de Lodge, Wormald *et alii* y de Mathy, aunque tampoco incluye ninguna referencia a la frecuencia de cada palabra. En dicha obra aparecen alfabetizados 2.500 términos muy frecuentes en el programa

oficial junto con sus principales estructuras, pero no consta ninguna aclaración de en qué medida se utilizan los materiales de los autores anteriormente citados. Estas dos obras son, pues, tributarias de la finalidad para la que han sido creadas: la enseñanza.¹

2.2. La dispersión

Aunque la frecuencia es el criterio estadístico más importante a la hora de seleccionar un vocabulario, existe otro que también han tenido en cuenta algunos diccionarios: la *dispersión*. Mediante este criterio, se seleccionan las palabras según el número de contextos en que figuran. Al añadir este criterio junto con la frecuencia, se evita que aparezcan palabras muy frecuentes que quizás están sólo en autores muy determinados y se prima aquellas que con una frecuencia menor se encuentran repartidas en más autores. Se empezó a aplicar en diccionarios para inglés y francés,² pero para el latín el primero en aplicar este criterio es el *Frequency Dictionary of Latin Words* de Gardner, publicado en 1971 dentro de una serie más amplia de diccionarios de frecuencia de lenguas románicas llevados a cabo durante los años sesenta en la Universidad de Stanford. Básicamente, es un diccionario de frecuencia como los comentados anteriormente, pero incluye como criterio corrector de la frecuencia la dispersión, representada en un coeficiente³ que cuantifica el número de contextos⁴ en que figura la palabra. De este modo, *ager*, que tiene

¹ El *Lessico latino fondamentale* de Riganti (1989) parece tener fines pedagógicos, a juzgar por la propuesta de la introducción (obtención de un bagaje adecuado, fácil de memorizar y ordenado según algún criterio), pero no creemos que esté dirigido a alumnos que comienzan a aprender esta lengua; así lo demuestra el que incluya más de 2.000 palabras tomadas del *Frequency dictionary of latin words* de Gardner (1971), que figure el índice de uso de éste sin mayor aclaración de su utilidad y que aparezca la etimología indoeuropea, latina, griega o germánica y los derivados italianos.

² Para inglés Thorndike 1921, West 1953, West-Endicott 1941; para francés Henmon 1924, Vander Beke 1935, Aristizabal 1938; para español Buchanan 1927, *Recuento* de la Universidad de Puerto Rico 1952, García Hoz 1953 (cf. Gougenheim *et alii* 1964).

³ Para determinar este coeficiente se utiliza la fórmula $D = 1 - s/2m$, donde D es la dispersión de la palabra y expresa en tantos por ciento el número de contextos en que aparece el término, s es la *desviación típica*, que representa la media de las desviaciones de cada valor en cada contexto respecto a la media y m es la frecuencia media de aparición de la palabra en todos los contextos considerados.

⁴ En concreto son los siguientes: A) poesía épica, B) satíricos, Lucrecio, obras menores de Virgilio y verso popular no hexamétrico, C) historiografía y D) prosa epistolar y ensayística.

una frecuencia menor que *mons* en la muestra de Gardner, aparece antes en la lista final por orden decreciente de importancia, puesto que tiene mayor dispersión (61'5% de los contextos) que esta última (58'9% de los contextos). Así pues, con la dispersión se corrige algo el desequilibrio que puede suponer la sobrerrepresentación de algún tipo de obras literarias.

En concreto, la muestra seleccionada por Gardner trata de representar a una población que abarca textos no fragmentarios que van de Cicerón a Tácito, divididos en los cuatro ámbitos reseñados (cf. nota 4). De cada grupo toma una muestra de 62.500 ocurrencias, lo que supone una muestra total de 250.000. Según las tablas estadísticas,⁵ estas 250.000 ocurrencias son suficientemente representativas de la población que se quiere describir en un 99,7%, con un margen de error de ± 1 , ya que para una población infinita se estimaría como muestra representativa 22.500 ocurrencias. Desde luego es criticable el hecho de que la división en cuatro ámbitos no sea proporcional al número de textos conservados de cada género, dándose casos de sobrerrepresentación (lírica) y de subrepresentación (historiografía).⁶ Sin embargo, sea cual fuere la cantidad de ocurrencias conservadas de cada género o ámbito, las 62.500 ocurrencias escogidas para cada ámbito cumplen suficientemente con los márgenes de confianza que impone la estadística.

Mayores problemas puede plantear el modo en que se ha hecho la selección de cada autor, de cada obra y de cada texto del que se han tomado las ocurrencias. La primera objeción viene de la selección por géneros, pues falta la poesía didáctica y la prosa técnica.⁷ El mismo autor basa la exclusión en su intento de centrarse en el latín literario, pero fuerza un tanto el término *literary* (1971:8) al

⁵ En concreto, estos índices están tomados de G. Tagliacarne, *Técnica y práctica de las investigaciones de mercado*, Ariel, Barcelona, 1968, pág. 156.

⁶ Bustos Tovar (1967) hace unas críticas parecidas al recuento para el español de la misma serie (Juillard y Chang-Rodríguez 1964).

⁷ La ausencia de prosa didáctica no es absoluta, pues podrían considerarse como representativas la obra de Quintiliano, *De institutione oratoria* o la de Plinio el Viejo, *Naturalis Historia*. Sí, en cambio, puede considerarse absoluta la ausencia de obras técnicas, que justifica con estas palabras: «Didactic poetry has generally been excluded on the ground that it is both too obscure. Those scientific works and technical treatises which were neither widely read nor influential on subsequent literary styles have also been omitted. Hence, Columela, Frontinus, Vitruvius and Manlius are not included» (1971:8)..

incluir obras como las de Plinio el Viejo y Quintiliano. Por otra parte, la elección de los textos no ha sido al azar, ni sopesando la importancia literaria, sino que los ha elegido guiándose de los principios en que se basan Palmer, Rose y Meillet, sin mayor aclaración. Ha extraído oraciones completas de los textos seleccionados para cada ámbito, que tienen como característica general no ser fragmentarios.

El diccionario en sí se organiza en dos partes bien diferenciadas. En la primera figura una lista con unas 4.000 palabras por orden alfabético, en cada una de cuyas entradas se ofrecen las ocurrencias de cada contexto junto con su número de frecuencia y de dispersión. En la segunda, se presentan a tres columnas estas mismas palabras colocadas por orden decreciente: en la primera columna según el uso,⁸ en la segunda según la frecuencia y en la tercera según la dispersión.

A pesar de las objeciones que se le puedan hacer al diccionario de Gardner, queda como un modelo válido y científico de cómo se puede seleccionar con criterios estadísticos un vocabulario didáctico.

2.3. *Libros de texto y diccionarios de frecuencia*

La existencia de este tipo de diccionarios no ha tenido reflejo en los libros de texto a la hora de incluir el vocabulario, pero esto es especialmente llamativo en los textos propuestos para traducir. Según Guiraud (1960:93), las 100 primeras palabras de una lista de frecuencia suministran el 60% de las palabras que aparecen en un texto y las 1.000 primeras el 85%. Así pues, deberíamos esperar al menos que se mantuvieran estos tantos por ciento en textos propuestos para primer año de latín. En el cuadro siguiente ofrecemos el resultado de 3 calas hechas en dos libros de texto (Zulueta 1988:46,144 y Jones y Sidwell 1986:54), teniendo en cuenta el diccionario de Gardner.

⁸ El valor con el que figuran en esta columna es la síntesis de combinar los dos criterios, frecuencia (f) y dispersión (d), mediante la fórmula $Uso = (d * f) / 100$.

rango	Zulueta-1	Zulueta-2	Jones-Sidwell
Gramaticales + 1-100	37,5%	48,98%	49,09%
1-1.000	55,35%	73,46%	78,17%

Representación de las 1.000 palabras más frecuentes en libros de texto

Junto con las 100 primeras palabras más usuales, hemos incluido 96 que consideramos gramaticales, entre las que están los posesivos y demostrativos, el artículo, las conjunciones y preposiciones principales, el verbo *sum*, etc. Como puede observarse, todos los tantos por ciento quedan lejos de la cifra del 60% y 85% que da Guiraud, especialmente las calas de Zulueta 1 y 2.

Por otra parte, puesto que son manuales de enseñanza, esperaríamos que la introducción de palabras siguiera un criterio de presentación de éstas progresivo según su frecuencia. De este modo, deberían figurar más veces las palabras con frecuencia mayor y reducir el número de las palabras con menor frecuencia. Para comprobarlo, hemos distribuido en cinco intervalos según rango decreciente de frecuencia las palabras que aparecen en los textos examinados, obteniendo los siguientes resultados:

rango	Zulueta-1	Zulueta-2	Jones-Sidwell
Gramaticales + 1-100	37,5%	48,98%	49,9%
101-250	3,57%	8,16%	7,27%
251-500	3,57%	6,12%	7,27%
501-1.000	10,71%	10,20%	14,54%
1.000-	32,14%	22,45%	9,09%

Distribución en 5 intervalos de las palabras más frecuentes en libros de texto

En lugar de haber más palabras de los rangos menores (donde están las palabras más frecuentes) e irse reduciendo, ocurre a la inversa: tras un grupo bastante numeroso de palabras muy frecuentes (gramaticales y las del rango 1 a 100), en el intervalo siguiente (101-250) hay una disminución del número de palabras, aumentando después el número en intervalos inferiores, sobre todo en las calas de Zulueta. Esto supone una contradicción con los estudios sobre la frecuencia de las palabras, pues la tendencia lógica es a que los términos que se hayan en rangos más bajos, es decir, los más frecuentes, estén más representados en los textos que aquellos de intervalos más altos.

Así pues, aunque no se puede generalizar a partir de estas muestras, permítasenos al menos sospechar del grado en que otros libros de enseñanza básica de latín han tenido en cuenta la frecuencia de las palabras a la hora de elegir los textos para traducir.

3. EL LÉXICO COMO ELEMENTO ESENCIAL DEL APRENDIZAJE

Hasta ahora hemos considerado la frecuencia como una herramienta estadística que nos ayuda a seleccionar determinadas palabras de una lengua, lo que redunda en una mayor eficacia a la hora de aprender esa lengua. Pero de nada sirve si no hay una concepción del manual de enseñanza como un todo donde el léxico es una parte muy importante de él, no un apéndice en forma de lista alfabética, cuyos criterios de selección no siempre se explicitan. El mismo Coseriu (1987:236) dice a propósito de esto:

«El aprendizaje del vocabulario, en especial del de una lengua extranjera, se hacía hasta hace poco tiempo de una manera poco sistemática, según el antojo de la ejemplificación gramatical y el azar de los textos de lectura. La determinación de la frecuencia de las palabras ha introducido en este campo un primer principio de racionalización, y la clasificación onomasiológica por dominios de «cosas» designadas, un segundo. Pero estos dos principios, aunque sin duda importantes, siguen siendo insuficientes para un aprendizaje racional y sistemático del vocabulario. La clasificación de palabras según su frecuencia no dice nada con respecto al significado y la designación; y el proceder onomasiológicamente es enteramente idóneo sólo para el léxico terminológico, no estructurado desde el punto de vista propiamente idiomático (y, claro está, también para las secciones terminológicas de la lengua corriente, no científica).»

Es decir, de nada nos sirve la frecuencia si no concebimos el aprendizaje de una lengua como aprendizaje de las relaciones que establecen entre sí los lexemas. Así pues, no se puede estudiar la palabra *homo* independientemente de *mulier* y *vir*, ni ésta sin tratar de *uxor*. Al hablar del significado del verbo *audio*, que indica la capacidad auditiva, hemos de relacionarlo con *attendo* ('prestar atención'), que tiene el clasema no resultativo [+dinámico] y ponerlo en relación con *accipio*, que también es dinámico, pero tiene aspecto resultativo, por lo que puede significar también 'he oído' (García-Hernández 1977). La frecuencia, por tanto, ofrece un marco en el que empezar a trabajar el léxico desde los primeros momentos, de modo que las palabras utilizadas resulten rentables en otros ejercicios, pero es el léxico y sus relaciones el verdadero protagonista del aprendizaje.

Sin duda, la comprensión de los textos propuestos a los alumnos se ve sustancialmente influida por el grado de dominio del vocabulario y de las relaciones lexemáticas de la lengua. En el estudio de Laufer (1992) con alumnos de inglés cuya lengua materna era el hebreo, en que examina el grado de correlación entre el tamaño de vocabulario y la comprensión lectora de los alumnos, el vocabulario es el primer factor que influye, por delante del conocimiento del tema y de la sintaxis, que figura en tercer lugar. De nuevo entramos aquí en los odiosos porcentajes, pues resulta que la comprensión del texto no es proporcional al número de palabras conocidas de éste. Si Guiraud (1960:93) dice que el 60% de las palabras de un texto está constituido por las 100 primeras palabras de una lista de frecuencia y que las 1.000 primeras conforman el 85% del mismo, no quiere decir que se vaya a comprender en esta misma medida, pues el valor informativo de las palabras poco frecuentes es mayor que el de las muy frecuentes. Por esto, a continuación de estos porcentajes y salvando las fórmulas matemáticas, nos dice que las 100 primeras palabras más frecuentes nos informan del 30% del texto, las 1.000 primeras del 50% y las 4.000 primeras del 70%. Siendo pues el conocimiento del léxico uno de los principales factores de la comprensión de un texto y tan costosa la adquisición de un vocabulario, merece la pena conjugar lexemática y frecuencia a fin de seguir unos principios que rentabilicen y hagan más eficaz la adquisición de un vocabulario.

La dependencia de los alumnos del diccionario, tan criticada entre nosotros especialmente en los años cincuenta y sesenta desde la pro-

pia práctica docente (Marín Peña 1952, Merino Rubio 1956, Burgos 1957, Lafuente 1965), es estudiada por Benoussan (1992). Esta autora demuestra que el diccionario sólo es útil para estudiantes avanzados, que son capaces de una visión panorámica (macronivel) del texto y, precisamente por ello, de escoger mejor el sentido de las palabras en una frase concreta (micronivel), mientras que los estudiantes con un nivel bajo de conocimientos, que no tienen esta visión general del texto, se sitúan en el micronivel de la palabra, no sabiendo por esto qué sentido escoger. Así se explica que los mayores errores sean por elección de acepciones equivocadas. La conclusión evidente es que no es recomendable el diccionario, mientras el alumno no posea un vocabulario mínimo que permita la comprensión global del texto, pues de no ser así titubeará en la elección de las acepciones y no será capaz de hacer una traducción inteligible.

La tradición gramatical de los estudios latinos ha influido decisivamente en la importancia dada a la gramática en la enseñanza, a veces en detrimento del léxico, hecho que se paliaba con una enseñanza reglada que concedía un número de horas suficientemente amplio para luego rentabilizar ambas cosas. La carencia de un vocabulario mínimo vacía de contenido todos los textos y provoca una dependencia del diccionario absoluta y nefasta como hemos comentado antes. Basta hojear un libro de texto de latín del nivel de secundaria para darse cuenta del gran desequilibrio existente.

No quiere decir esto que el autor de manuales o el profesor no conceda el valor que merece al léxico. A nuestro juicio, se trata más bien de un problema de *doctrina recepta*, que hace que se centre el esfuerzo en la gramática, a pesar de los estudios comentados arriba y de otros más directamente relacionados con el latín. García-Hernández (1989:843) apunta esta razón para explicar la carencia y propone abordar la enseñanza del léxico desde el punto de vista de la lexemática:

«Es un hecho que la enseñanza tradicional del latín adolece de un exceso de gramática, de morfología y de sintaxis. Pensamos que conviene prestar mayor atención al estudio del léxico, particularmente en la fase inicial del aprendizaje del latín y, por supuesto, hay que enseñar el léxico de una forma más organizada. El estudio del léxico latino de una forma racional y con perspectiva histórica puede ejercer gran atractivo para el estudiante de latín, por cuanto que en él puede encontrar una utilidad de forma inmediata.»

Abordarlo de una forma más organizada supone no perder de vista el sistema y aquí el concepto de oposición lingüística aplicado al plano del contenido puede resultar muy instructivo, del mismo modo que los procedimientos de formación de palabras mediante *modificación, desarrollo y composición* (Coseriu 1977:169) resultan muy educativos para el alumno, que tiene que hacer un verdadero esfuerzo de deducción, al tiempo que eficaces, ya que adquiere un volumen considerable de palabras con un reducido número de términos. Efectivamente, el conocimiento de un término base y de los prefijos o sufijos que se le pueden añadir para modificarle sin cambiar de clase (modificación), cambiando (desarrollo) o uniéndose dos términos base (composición) puede ampliar enormemente el horizonte léxico del alumno.⁹

Pinkster (1995:77), desde un punto de vista más empírico, corrobora la importancia que puede tener el léxico en la comprensión del texto latino, aunque no da datos del estudio llevado a cabo.¹⁰ Según este autor, sólo entre un 5% y un 10% de las ocasiones la marca de caso es necesaria para determinar la función sintáctica o semántica de las palabras en la frase y concluye al final del párrafo lo siguiente:

«Sin embargo, sí puede deducirse de esta conclusión que la marca de caso no es más que uno de los bloques constructivos y que es mucho menos importante para el éxito de la comunicación que, por ejemplo, el significado de los lexemas con los cuales se construyen las oraciones.»

En este sentido, coinciden sus conclusiones con las del estudio de Laufer ya comentado, en que se deja claro que la comprensión está más relacionada con el léxico que con la gramática.

⁹ Las posibilidades didácticas de estos procedimientos están expuestas con gran claridad en el trabajo publicado en esta misma revista de López Moreda y Rodríguez Alonso (1989).

¹⁰ El mismo autor dice que los porcentajes son parecidos a los obtenidos por Hubka (1982) y Schössler (1984) para el francés antiguo.

4. CONCLUSIÓN

A lo largo del artículo nos hemos centrado en uno de los instrumentos más utilizados para el aprendizaje del latín, el vocabulario didáctico, y analizamos algunos de los criterios que han servido para seleccionarlo. Aunque la frecuencia es un criterio que aparece ya en Comenio, no cobra la suficiente importancia hasta el desarrollo de la estadística y la fijación de la traducción como ejercicio escolar principal.

Desde principios de siglo se publican diccionarios de frecuencia latinos, como los de Lodge, Wormald *et alii*, Mathy, Gardner y Delatte *et alii*, motivados en parte por la investigación en estadística literaria, pero sobre todo por la aplicabilidad didáctica de los mismos. Así lo demuestra el que los diccionarios anteriores hayan sido la base para vocabularios latinos claramente escolares, como los de Echave-Sustaeta y Podvin.

A mediados de los años sesenta, empiezan a surgir diccionarios de frecuencia para diversas lenguas romances en el seno de la Universidad de Stanford, pero incluyen además de este criterio el de dispersión, según el cual se valora la importancia de la palabra por el número de ámbitos en que aparece. Dentro de esta colección, Gardner publica su *Frequency Dictionary of Latin Words*, en el que se clasifican más de cuatro mil palabras por orden decreciente según los criterios de frecuencia, dispersión y, el más importante de todos, uso.

En el apartado 2.3 hemos puesto de manifiesto cómo los datos de este tipo de obras no se han proyectado plenamente en el tratamiento del léxico que se incluye en los libros de texto. Efectivamente, el porcentaje de palabras que están entre las cien más usuales, según Gardner, y las gramaticales, apenas llega al 40% de representación en las calas que hemos hecho, cuando en los textos examinados por Guiraud lo normal es un 60%. La frecuencia no parece, pues, que haya solucionado los problemas relacionados con el aprendizaje del vocabulario, dada la escasa repercusión que ha tenido en los manuales escolares, además de que no ofrece otra metodología de aprendizaje que la mnemotécnica.

Junto a esto, autores como Coseriu, Laufer o Benoussan han venido subrayando la importancia del vocabulario en el aprendizaje de una lengua, sugiriendo en el caso del lingüista rumano y de García-

Hernández una aproximación a la didáctica del vocabulario desde el campo de la lexemática, en concreto, a partir de los procedimientos de formación de palabras (modificación, desarrollo y composición). De este modo, el léxico, que el alumno ve a veces como una masa caótica e indomable, se convierte poco a poco en algo cercano y manejable si se estudia de una forma organizada. Conociendo el sentido direccional de preposiciones como *ab*, *ex*, *per*, *in*, *ad* y el significado del verbo *ago*, no es difícil para el alumno descubrir lo que significan *abigo*, *perago*, *adigo*, *indigo*; otro tanto ocurre, por ejemplo, si se conoce el sentido de lugar del sufijo *-arium* y la palabra *argentum* al presentarse el término *argentarium*.

Así pues, creemos que la enseñanza del latín se vería beneficiada si junto a programas gramaticales tan minuciosamente diseñados se incluyera en los libros de texto un programa léxico, basado en criterios estadísticos y lingüísticos, que sirviera de base para leer textos propios del nivel, pero también para provocar la reflexión lingüística en nuestro alumnos.

Bibliografía

- Benoussan, M. (1992), «Learners' spontaneous translation in an L2 reading comprehension task: vocabulary knowledge and use of schemata» en Arnaud Y Bejoing (eds.), *Vocabulary and applied linguistics*, Macmillan Academic and Profesional Ltd., Houndmills, Basingstoke, Hampshire and London, pp. 103-113.
- Burgos, M^a A. (1957), «La enseñanza del latín y la tiranía del diccionario», *Revista de Educación*, 21, n^o 60, pp. 1-2.
- Buridant, Cl. (1986), «Lexicographie et glossographie médiévales: esquisse de bilan et perspectives de recherche», *Lexique* 4, pp. 9-46.
- Castro, A. (1991), *Glosarios latino-españoles del la Edad Media*, CSIC, Madrid [1^a ed. 1936].
- Comenio, J.A. (1637), *Porta linguarum trilinguis reserata & aperta sive Seminarium Linguarum & Scientiarum omnium. Hoc est: compendiaria latinam Anglicam, Gallicam (& quamvis aliam) linguam una cum artium & scientiarum fundamentis sesquianni spatio ad summum docendi & perdiscendi methodus, sub titulis 100. periodis 1058. comprehensa / opera, studio & elucubratione Ioh. Anchorani*, excudebat Anna Griffin, sumptibus Michaelis Sparke, Londini.
- (1657), *Didáctica magna*, Editorial Porrúa, México, 1988.
- Coseriu, E. (1977), «Introducción al estudio estructural del léxico» en *Principios de semántica estructural*, Gredos, Madrid, pp. 47-142 [1^a ed. 1966].

- (1987), «El estudio funcional del vocabulario», en *Gramática, semántica y universales*, Gredos, Madrid, pp. 206-238 [1ª ed. 1975].
- Delatte, L., Evrard, E., Govaerts, S. y Denooz, J. (1981), *Dictionnaire fréquentiel et index inverse de la langue latine*, Liège.
- Echave-Sustaeta, J. (1953), *Lengua latina. Vocabulario básico*, Cefiso, Barcelona.
- García-Hernández, B. (1977), «El campo semántico de 'oír' en la lengua latina. Estudio estructural», *REL*, 7, pp. 115-136.
- (1989), «Cultura clásica, una disciplina para la primera etapa de la educación secundaria (14-16 años)», *Actas del VII Congreso de EE.CC.* (Madrid 1987), vol. 3, pp. 839-44.
- Gardner, D.D. (1971), *A frequency dictionary of latin words*, UMI, Ann Arbor, Michigan, 4 vol.
- Gougenheim, G. (1958), *Dictionnaire fondamental de la langue française*, Didier, Paris.
- (1970), *Études de grammaire et du vocabulaire français*, Picard, Paris
- , G., Michéa, R., Rivenc, P. y Sauvegeot, A. (1964), *L'élaboration du français élémentaire (1^{er} degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*, Didier, Paris, [1ª ed. 1956].
- Govaerts, S. y Denooz, J. (1974), *Lexique de base latin*, Laboratoire d'Analyse statistique des Langues anciennes (LASLA), Universidad de Lieja, Liège.
- Guiraud, P. (1960), *Problèmes et méthodes de la statistique linguistique*, PUF, Paris.
- Gutiérrez de Cerezo, A. (1497), *Ars Grammatica* [ed. de M.A. Gutiérrez, *Andreas Guterrius Cerasianus: Ars Grammatica*, Burgos, 1998, 2 tom.].
- Gutiérrez Galindo, M.A. (1996), «Algunas aportaciones metodológicas de la gramática protohumanista a la enseñanza del latín: el uso de los diccionarios», en J. Bartolomé, C. García Román, F. González Vega, M.A. Gutiérrez Galindo y P. Redondo (eds.), *Actas del I encuentro científico y pedagógico. Historia y métodos en la enseñanza de las lenguas clásicas*, Universidad del País Vasco, Vitoria, pp. 95-106.
- Jones, P.V. y Sidwell, K.C. (1986), *Reading latin*, Cambridge University Press, Cambridge.
- Lafuente, Mª R. (1965), «Consideraciones en torno a un vocabulario básico (I)», comunicación al *II Congreso Español de EE.CC.*, pp. 446-50.
- Laufer, B. (1992), «How much lexis is necessary for reading comprehension?» en P.L.J. Arnaud y H. Bejoint (eds.), *Vocabulary and applied linguistics*, Macmillan Academic and Profesional Ltd. Houndmills, Basingstoke, Hampshire and London, pp. 126-32.
- Lodge, G. (1907), *The vocabulary of high school latin*, New York.
- López Moreda, S. y Rodríguez Alonso, R.C. (1989), «Formación de palabras y aprendizaje del vocabulario latino. Utilidad del método estructural», *Estudios Clásicos*, 96, pp. 99-113.
- Lucas, J.M. (1986), «El vocabulario básico griego», *Estudios Clásicos*, 90, pp. 207-237.
- Mathy, M. (1952), *Vocabulaire de base du latin*, EDITIONS O.C.D.L., Paris.

- Merino Rubio, W. (1956), «La selección del vocabulario y la mecánica lingüística», *Revista de Educación*, pp. 67-70.
- Meara, P. (1980), «Vocabulary acquisition: a neglected aspect of language learning», *Language Teaching and Linguistics: Abstracts*, 15, 4, pp. 221-246.
- Michéa, R. (1953), «Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage», *Les Langues Modernes*, tomo 47, 1953, pp. 338-44.
- Owen, A. (1929), *Traité de Walter de Bibbesworth sur la langue française*, P.U.F, Paris.
- Pinkster, H. (1995), *Sintaxis y semántica del latín*, Ediciones Clásicas, Madrid [1ª ed. hol. 1984].
- Podvin, M.-L. (1981), *Les mots latins: les 2.500 mots et constructions de base du latin*, Editions Scodell, Paris.
- Purnelle, G. (1995), «Opera latina: la banque de donnés de textes latins de LASLA sur CD-ROM», *RISSH* 31, 1995, pp. 173-182.
- Rico, F. (1978), «El cuaderno de un estudiante de latín», *Historia* 16, 25, pp. 16 y en Matín, J. L. et alii, *Así nació el castellano. Cuadernos de Historia* 16, 18, Madrid, 1985, pp. 12-14.
- Riganti, E. (1989), *Lessico latino fondamentale*, Pàtron Editore, Bologna.
- Scheler, A. (1965), *Lexicographie latine au XIII^e siècle: Trois traités de Jean de Garlande, Alexander Neckam et Adam du Petit Pont*, Jahrbuch für rom. und eng. Literatur, Leipzig.
- Valla, L. (1540), *Elegantiae*, Basilea [ed. de E.Garin Laurentius Valla: *Opera omnia*, Turín 1962, 2 vols.].
- Villadei, A. , *Das 'Doctrinale' des Alexander de Villa-Dei (Kritishe-exegetische Ausgabe)* [MGH XII], Berlín 1893 [reimpr. Nueva York 1974].
- Wormald, Line, Morris y Protheroe (1949), *Latin basic vocabulary*, Centaurs Books, Blackpool.
- Zulueta, J. de. (1988), *Latín/2. Bachillerato*, Madrid.

JOSÉ MANUEL SANTIAGO ÁNGEL
 I.E.S. «SAMANIEGO» (Laguardia, Álava)
 e-mail: iessanjms@terra.es