

La selección del vocabulario didáctico: un nuevo enfoque

José Manuel Santiago
IES "Samaniego", Laguardia (Álava)

1. LA IMPORTANCIA DE UN VOCABULARIO BÁSICO

Como hemos ya expuesto en otra parte (Santiago 1998), el léxico constituye uno de los pilares fundamentales sobre los que debe asentarse todo el proceso de enseñanza y aprendizaje de una lengua, con una importancia superior incluso a la gramática. La inclusión en los programas dota a estos de un entramado lingüístico que facilita la comprensión de la cultura que transmiten los textos. Los alumnos abandonan la tediosa dependencia del diccionario y se hacen con una herramienta que les va a permitir una comprensión más rápida de los mensajes y que rentabilizarán posteriormente con la incorporación de nuevas palabras relacionadas con éstas a través de diversos procedimientos (sinónimos, antónimos, campos léxicos, familias léxicas). Sin embargo, el problema llega cuando hay que decidir qué palabras son las que deben estar incluidas en los primeros cursos de enseñanza de una lengua.

Esta selección, realizada de un modo intuitivo la mayor parte de las veces, se ha ido dejando en manos de la estadística, pues hechas las comprobaciones oportunas se vio que había un número relativamente pequeño de palabras que aparecían con mucha frecuencia, hasta el punto de que las 4.000 palabras más frecuentes de una lengua abarcaban el 97,5% de un texto. Así se establecieron léxicos para algunas lenguas modernas y para el latín (Lodge 1907, Mathy 1952), que fueron de gran utilidad para los profesores, pero que no promovieron una discusión sobre el papel que podían jugar.

1.1. LOS INCONVENIENTES DE LA FRECUENCIA

Por otra parte, aunque la frecuencia es un criterio matemático y los resultados son objetivos, estos dependen de la preparación de la muestra sobre la que se haya aplicado. Las primeras críticas en el campo de las lenguas modernas surgieron porque generalmente las muestras estaban tomadas de textos escritos, por lo que las diferencias entre lengua hablada y lengua escrita se manifestaban también en la selección. Gougenheim (1958) intenta paliar este problema utilizando además grabaciones, pero aun así no se libró de otras críticas que atañen más a la sociolingüística. Para el latín las críticas se centran en cuál debe ser el *corpus* de autores que debe incluirse en la muestra, pues no es lo mismo una muestra con textos de historia que con cartas.

Las objeciones¹ generales hechas a la frecuencia se pueden resumir en las siguientes:

- 1º) No hay relación entre poca frecuencia e inutilidad de una palabra para el aprendiz (*civitas* y *legio* pueden no aparecer entre las más frecuentes, pero son indispensables al traducir a César; lo mismo pasa con el verbo *for* y *puppis* en Virgilio).
- 2º) Palabras muy frecuentes pueden resultar inútiles para el aprendiz (*bellum* o *civitas*, muy frecuentes en César, apenas aparecen en Fedro y Ovidio).
- 3º) Una palabra poco útil para un aprendiz concreto, puede serlo mucho para otro (los términos que hay que enseñar a un turista son muy distintos a los que se enseñarán a un inmigrante).
- 4º) Las listas muchas veces no se ponen de acuerdo en el orden de las palabras, por lo que la elección de una lista u otra afectará a los resultados.
- 5º) El orden en que aparecen las palabras en la lista no es aplicable directamente a la enseñanza.

1 Estas críticas se pueden encontrar en Gougenheim *et alii* (1964), Savard (1970), Ghadessy (1979), Nation (1983) y Carter (1987), por citar sólo a algunos.

6º) A veces, las listas no ofrecen un nivel de confianza suficientemente alto.

Con todo, la mayoría de los vocabularios nucleares que se conocen para inglés, francés y español hasta mediados de los años cincuenta han sido elaborados teniendo en cuenta este criterio².

Después, este criterio empezó a ser complementado por otros como la *dispersión* y la *disponibilidad*. Mediante el primero se trata de medir la extensión de una palabra en los diversos ámbitos en que se había dividido previamente la muestra, mientras que con el segundo se da entrada a la subjetividad de los hablantes, ya que son ellos los que dicen qué palabras les parecen más importantes al tratar determinados temas. Así en el diccionario de Gougenheim (1958) la palabra *fourchette* está entre los últimos lugares por su frecuencia, pero en el recuento de las entrevistas a hablantes, es indispensable para tratar un tema como la comida. Aplicado al latín, es lo que ocurriría con las palabras *for* y *puppis* ya comentadas.

1.2. ORGANIZACIÓN INTERNA DE LA LENGUA

El desarrollo de la lingüística en este siglo ha revelado la estructura fonológica y la estructura morfosintáctica de las lenguas, siendo la semántica la que todavía no ha conocido un grado de descripción paralelo. Los estudios de Trier y sus seguidores descubrieron todo un sistema en el estudio de los campos semánticos, así como Coseriu logró perfeccionar un programa de descripción semántica del léxico y estableció una primera estructuración del mismo desde tres puntos de vista: expresión léxica, expresión y contenido, y contenido. Pero es sólo desde el contenido que se puede descubrir una estructura coherente del mismo siguiendo el método aplicado en fonología: conmutación y distribución. Mientras que la semántica es la ciencia referente a cualquier problema de la significación, la lexemática sería la rama de ésta que estudiaría las relaciones estructurales de los significados en un mismo sistema lin-

2 Vid. Gougenheim *et alii* (1964), donde se comentan algunos de éstos, especialmente para inglés y francés.

güístico. El concepto de lexema como unidad de contenido léxico le sirve para llegar al archilexema, que sería la unidad cuyo contenido es idéntico al de dos o más unidades de un campo léxico y el de clasema como rasgo distintivo que funciona en la misma categoría funcional (p. ej., [ser humano] en *miles, magister, senex*) para distinguir la clase léxica. Pero, ¿cómo es posible explicar que a pesar de todas las variantes diatópicas y diastráticas, el vocabulario de una lengua se mantenga con cierta unidad? Coseriu (1977: 92) lo expresa de la siguiente manera:

Por otra parte, si bien es verdad que la variedad del léxico es mayor que la variedad fonológica o gramatical, la comprensión recíproca, aún en comunidades muy amplias, también es un hecho innegable, lo que implica que al menos una buena parte del léxico de las lenguas correspondientes debe presentar una organización más o menos homogénea.

De la misma opinión es Stubbs (1986), quien responde a la pregunta afirmando que existe un área central del vocabulario de una lengua, al que llama *vocabulario nuclear*, que permite mantener la homogeneidad del contenido del léxico. Por ejemplo, el archilexema *bos* mantiene la unidad de contenido como ‘animal bovino’ para otras unidades de la lengua como *taurus* y *vacca*, del mismo modo que *asiento* lo es para *banqueta, sofá, sillón, otomana*. En estos ejemplos el archilexema se puede conmutar por cada una de estas unidades, mientras que éstas entre sí ofrecen una distribución de significado diferente. Así, estos archilexemas recogen un valor unitario que mantiene homogéneo el campo semántico.

La Gramática Funcional, especialmente el Modelo Léxico Funcional (Martín Mingorance 1990) parte precisamente de la existencia de una jerarquía en el eje paradigmático para ayudarse luego a establecer los patrones de complementación de cada lexema, lo que pertenece ya al eje sintagmático. En el diccionario de una lengua no existen unidades abstractas, sino léxicas de la propia lengua, organizadas jerárquicamente, de modo que cada dominio (campo semántico) o subdominio tiene su archilexema y unos hipónimos que se pueden definir a partir de los hiperónimos. *Decir* está en la definición de *prometer*, ésta en la de *jurar*, y ésta última en la de *abjurar*, formando una sucesión jerár-

quica. Ogden (1930) ya vio esta posibilidad al concebir su BASIC (British American Standard English Comunication) como un *corpus* cerrado de palabras por combinación de las cuales se podían definir otras mediante paráfrasis. Por ejemplo, no aparece en su lista *to walk*, pero sí *to have* y el sustantivo *walk*, por lo que expresar aquella acción se conseguiría fácilmente mediante la perífrasis *to have a walk*.

Otro indicio de la homogeneidad del léxico es la posibilidad de establecer oposiciones entre unidades en razón de que una de las dos contenga el rasgo marcado. De este modo, *audio* tiene el rasgo no marcado de la percepción auditiva frente a *accipio*, que sí lo tiene, por lo que también puede indicar este sentido. Las palabras *altura* y *old* tienen el rasgo no marcado de la oposición a juzgar por expresiones como *Todo depende de la altura* o *How old are you?*, donde abandonan su significado opositivo frente a *bajura* y *young*, respectivamente.

Así pues, el hecho innegable de que exista comprensión en comunidades muy amplias, la existencia de unos archilexemas para cada campo semántico, así como de unos lexemas en la lengua marcados con el rasgo negativo en oposiciones de contenido, que permiten que funcionen con el significado del término positivo, y, finalmente, la posibilidad de definir lexemas a partir de sus hipónimos hace que la idea contenida en Ogden (1930) y planteada más modernamente por Stubbs (1986) de que exista un vocabulario nuclear en todas las lenguas nos parezca bastante razonable.

1.3. POSIBILIDAD DE CUANTIFICACIÓN DE LA NUCLEARIDAD

Una vez demostrada la existencia de un vocabulario nuclear, Stubbs se plantea cómo descubrirlo, pues ello acarrearía indudables ventajas para otros campos de la investigación lingüística, pero especialmente para la didáctica de las lenguas. En efecto, saber de dónde parte la lengua para formar ese todo orgánico que es un idioma, puede ayudar a organizar de un modo más eficaz el aprendizaje del mismo.

Se trataría en suma de inventar unas pruebas que cuantifiquen en alguna medida la dimensión lingüística de las palabras, lo cual sería com-

plementario de la aplicación de la frecuencia, que se fija más en la dimensión estadística de las mismas.

Así enuncia una serie de pruebas que pueden servir para descubrir el vocabulario nuclear. Las pruebas están divididas en cuatro bloques:

- A) Neutralidad del discurso. El vocabulario nuclear es pragmáticamente neutro, lo que indica que las palabras que lo forman no tienen por qué tener connotaciones emocionales o evaluadoras, sino un sentido más conceptual y descriptivo. Con el criterio que denomina *culture-free* recoge palabras de la lengua tan unidas a la cultura que reflejan que no son objeto de préstamo, por ejemplo, los términos de funciones fisiológicas, partes del cuerpo o flora local. El hecho de que haya palabras que no indique el campo del discurso, p.ej., *izquierda/derecha* frente a *habor/estribor* también las acredita como más nucleares; las palabras que no afectan al tenor del discurso (no restringidas a un argot) o al modo del discurso (lenguaje hablado o escrito) también son nucleares; y, finalmente, las palabras con poco grado de expresividad, como ocurre al redactar el resumen de un texto literario, donde el tono y el estilo del autor se pierden en favor de enunciados mucho más proposicionales.
- B) Relaciones semánticas y sintácticas entre las palabras nucleares. Las palabras nucleares pueden organizarse en series hiponímicas, de modo que la más general será también la más nuclear, como en la serie *flor, rosa, rosa híbrida*. Otra prueba para comprobar si una palabra es nuclear consiste en observar si en la definición de otros términos se incluye, p. ej. *caminar, marchar y pasear* pueden expresarse incluyendo en la definición el verbo *andar* seguido de una determinación, por lo que éste es más nuclear, ya que en los otros no se puede operar igual; además, este término podría figurar en frases donde estuvieran los otros, pero no aquellos sin cambiar el sentido. Una tercera prueba está constituida por el número de colocaciones de la palabra nuclear, así *blanco* se puede combinar con muchas más palabras que *albino*. Las palabras nucleares tienen mayor extensión considerando ésta el número de

entradas del diccionario en que figura la palabra. Finalmente, el número de compuestos en los que puede aparecer también es un indicio de la nuclearidad de la palabra.

- C) Estructura del vocabulario nuclear. Desde este punto de vista los propios hablantes mediante test son capaces de asignar antónimos a las palabras nucleares de un modo muy claro como con los términos *bueno*, *gordo*, *limpio*, pero es mucho menos predecible cuando los términos son *excelente*, *obeso*, *higiénico*.
- D) Existen otros tipos de pruebas que pueden demostrar la pertenencia de una palabra al vocabulario nuclear, como son la preferencia por la palabra de origen germánico (en el caso del inglés) a la correspondiente románica, siempre que tengan el mismo sentido, de la más simple a la compuesta y de las palabras monosémicas, que no sean hipónimo y que no estén restringidas a una clase de objetos (p.ej., *blond* sólo para cabello).

El autor ofrece a lo largo del artículo la bibliografía en que se han discutido estas pruebas, que por otra parte, no es tan abundante como pudiera parecer, pero aquí nos interesa resaltar el hecho de que no es hasta principio de los ochenta que se empieza a indagar en las posibilidades didácticas del vocabulario (Meara 1980) y que estas pruebas están concebidas para lenguas modernas, lo que afecta profundamente a nuestra investigación.

2. EL MÉTODO DE SELECCIÓN DE UN VOCABULARIO NUCLEAR PARA EL LATÍN

Partimos, pues de dos limitaciones muy importantes si queremos hallar un vocabulario nuclear para el latín al modo en que se ha descrito en el punto anterior.

En primer lugar, el latín es una lengua de *corpus*, lo que implica que estamos considerando un conjunto cerrado de unidades, por lo que puede resultar más accesible el conjunto de datos, pudiendo incluso llegar a una muestra donde estuviera toda la población (al menos, la conservada), lo que con una lengua moderna resulta imposible. Pero, por otro

lado, selecciona enormemente el tipo de pruebas comentadas que se pueden aplicar, descartándose aquellas en las que hace falta la intervención del hablante (pruebas del tipo C) y las de resumen del tipo A ya comentadas.

En segundo lugar, no debe perderse de vista que el latín es una lengua clásica con una tradición gramatical muy fuerte, desde hace tiempo más inclinada a la comprensión que a la expresión, lo que ha limitado mucho los instrumentos que se pueden utilizar a la hora de elaborar un vocabulario de este tipo. Nos estamos refiriendo a diccionarios de sinónimos y antónimos, a diccionarios monolíngües redactados a partir de un lista limitada de palabras como pueden ser el de Gougenheim (1958) y el *Longman Dictionary of Contemporary English* (1978) y a la limitada existencia de materiales para el aprendizaje y evaluación del léxico.

2.1. LAS VARIABLES DE UN VOCABULARIO NUCLEAR

Así pues, dadas todas las posibilidades que se nos ofrecían, según Stubbs, y teniendo en cuenta las limitaciones comentadas, decidimos seleccionar un conjunto de cinco pruebas, denominadas *acepción*, *construcción*, *extensión*, *sintema* y *sinonimia*. Para cada uno de estos criterios anotábamos el número de acepciones de una palabra, sus construcciones sintácticas posibles, idiomatismos, sintemas y sinónimos. Para las acepciones se anotó el número de ellas que hay en el *Oxford Latin Dictionary (OLD)*, de donde también se han recogido las construcciones sintácticas que figuran a lo largo del artículo y los idiomatismos que aparecen en cursiva dentro del cuerpo de cada sentido. En el caso de las acepciones, se nos puede reprochar el no haber tenido en cuenta el *Thesaurus Linguae Latinae (ThLL)*, pero estimamos que éste adolece de cierta falta de sistematicidad al consignarlas, pues unas veces se organizan a partir de criterios de significado (*accedo*, *ager*), otras de construcción (*amicus*) e, incluso, de la mezcla de ambos criterios (*aio*). Compárense ahora los datos de número de acepciones recogidos de los dos diccionarios citados y del *Diccionario Latino (DL)* y el de Gaffiot (*DLF*):

| | ThLL | DL | DLF | OLD |
|---------|------|----|-----|-----|
| Acerbus | 4 | 6 | 4 | 8 |
| Accedo | 15 | 20 | 9 | 17 |
| Accuso | 4 | 4 | 3 | 3 |
| Acies | 18 | 24 | 11 | 7 |

Basta una breve mirada para darse cuenta de la disparidad de criterios con que cada uno divide el número de sentidos de una palabra, lo que no quiere decir que sean arbitrarios. Más bien hay que pensar en concepciones lexicográficas distintas, por lo que la exhaustividad del *Thesaurus Linguae Latinae* se ve mermada por la falta de unos criterios fijos para todos los lexicógrafos que han colaborado. Sin embargo, el *Oxford Latin Dictionary*, aunque mucho menos exhaustivo, da los datos de una forma más homogénea y sencilla.

Respecto a los sintemas tenemos que aclarar que hemos tomado el término de Martinet (1987: 52) quien lo define como ‘un signo lingüístico que la conmutación revela como resultante de la combinación de varios signos mínimos, pero que se comporta con relación a los otros monemas de la cadena como un monema único’. Bajo este criterio se recogen tanto compuestos, lo que ya se había hecho en otras listas (Savard 1970) y está propuesto por Stubbs entre los criterios del tipo B, como derivados. Hemos tomado el número de sintemas de cada una de las palabras consideradas ayudándonos del libro de Bréal y Bally *Les mots latins groupés d’après le sens et l’étymologie* junto con las notas sobre etimología del *Oxford Latin Dictionary*. No es útil utilizar aquí el diccionario etimológico sea el de Ernout y Meillet, sea el de Walde y Hofmann, porque no todas las palabras que figuran como derivaciones en cada entrada lo son del étimo que la motiva. Por ejemplo, de *miles* deriva *milito*, *militia*, *militarius*, *militaris* y *commilito* (sust.), pero no *commilito* (verbo), *commilitium* ni *militariter*, pues no derivan de *miles*, sino de alguno de sus derivados.

Al tratar los sinónimos tomamos el punto de vista de Lyons (1981: 441), para quien dos unidades son sinónimas si al conmutarse en una

frase el sentido no cambia. También así lo interpreta Savard (1970: 25) al afirmar:

Pour nous, le critère fondamental qui permet d'établir une serie de synonymes, c'est la possibilité de substituer un terme à un autre.

Esto no es rigurosamente cierto, ya que cada palabra distinta conlleva un significado distinto, sea conceptual, emocional o pragmático. No es lo mismo utilizar la palabra *entrada* que el anglicismo *hall*, como para un inglés tampoco lo es utilizar *to end* que *to terminate*, a pesar de que el objeto o la acción a la que se refieren es la misma, pues se añaden matices socioculturales o pragmáticos que diferencian los términos³. En cambio, a efectos prácticos de comprensión ambos términos, a los que habría que considerar co-hipónimos, son sustituibles. Incluso, ésta no se ve dificultada si se utilizan hiperónimos, lo que no ocurre al revés. La frase *¿Dónde resides?* puede ser menos comprensible para un principiante de español que *¿Dónde vives?*, en virtud de que el verbo *residir* tiene unos semas muy concretos que sólo recogen el hecho de habitar en un lugar, mientras que *vivir* abarca un espectro mucho mayor de semas entre los que está el de habitar un lugar, por esto puede substituir a *residir* en casi todos sus contextos, mientras que a la inversa no es posible.

Para el latín no existen diccionarios de sinónimos del estilo de los que conocemos para las lenguas modernas, si no más bien léxicos (Menge 1822, Dumesnil 1777) que aclaran en sentido de palabras cuyo uso se ha ido oscureciendo. Generalmente se trata de diccionarios del siglo pasado, útiles en una época en que todavía se utilizaba el latín como lengua de comunicación. El *Thesaurus Linguae Latinae* contiene también los sinónimos, pero la ausencia de criterios tipográficos claros para todos los artículos hace muy laboriosa la recolección. Finalmente, la obra de R. de Miguel (1867) contiene un tratado (*sic*) que es lo más parecido a un diccionario moderno de sinónimos, por lo que recogimos de aquí el número que tenía de ellos cada una de las palabras analizadas por nosotros.

3 Para centrar el problema de la sinonimia desde una perspectiva estructuralista, vid. García-Hernández 1997.

Una vez determinadas las pruebas que íbamos a aplicar, seleccionamos un conjunto de 1.000 palabras de las más frecuentes de la lista de Gardner (1971) y aplicamos cada una de las pruebas recogiendo el número de acepciones, construcciones, idiomatismos, sintemas y sinónimos de cada una de ellas. Obtenidos los datos brutos para cada lema correspondientes a los cinco criterios considerados, se hacía necesario agrupar todos estos en un valor que permitiera la comparación entre lemas a efectos de valorar su importancia, es decir, debíamos clasificarlos según un valor total. Para hacerlo hay que sumar los cinco valores asignados a cada palabra, pero no se puede hacer esto directamente, pues la media y la desviación típica de cada una de las columnas es distinta. Para poder comparar los valores de cada columna hay que tomar en consideración tanto la media de cada una como su *desviación típica* y *tipificar* los datos, con lo que los valores resultantes de todas las columnas lo serán respecto a una media de 0 y una desviación típica de 1, es decir, que todos estarán en un intervalo que irá de -1 a 1, excepto los valores muy extremos, ya que la desviación típica es una media de la desviación de cada valor respecto a la media de todo el conjunto.

Con la suma del valor correspondiente a cada prueba obtenemos un índice, que nos sirve para clasificar las palabras en orden decreciente, de más nucleares a menos nucleares o lo que es lo mismo, de más importantes lingüísticamente a menos importantes.

2.2. LAS CONCLUSIONES⁴

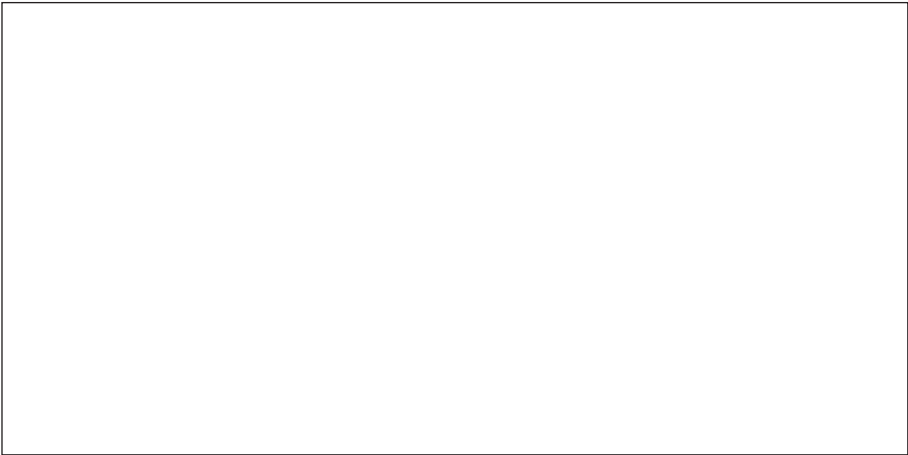
Desde un punto de vista estadístico el vocabulario basado en los criterios que conforman este tipo de selección siguen la misma tendencia que el basado en la frecuencia, confirmándose en cierto modo la adecuación de nuestro planteamiento y la posibilidad de utilizar este método para otras clasificaciones distintas a la frecuencia de las palabras de una lengua. La diferencia está, como hemos dicho en la aplicación de unas pruebas más lingüísticas en nuestro vocabulario. En la siguiente

4 Los resultados de esta investigación están expuestos más detalladamente en Santiago 1998. Aquí nos limitamos a exponer un breve resumen.

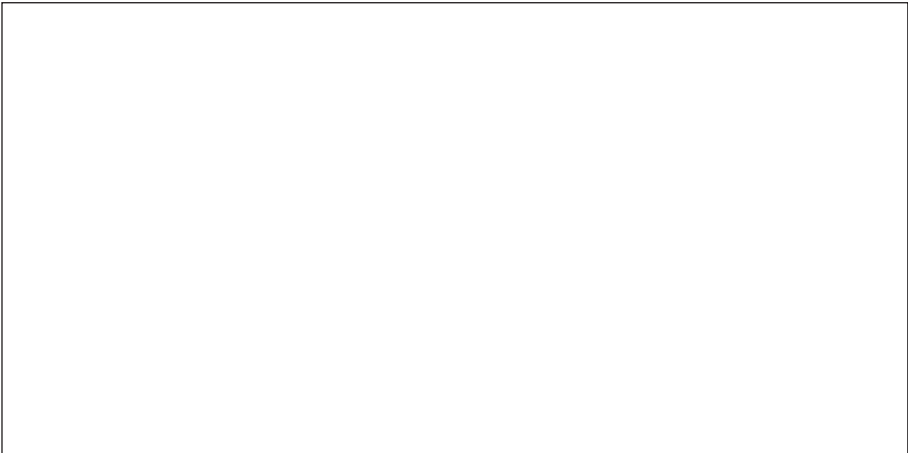
gráfica se comparan las tendencias de las 100 primeras palabras de una lista de frecuencia, como la de Govaerts y Denooz (1974), y de las 120 que figuran en la lista confeccionada por nosotros según los criterios aplicados.

Comparación curvas de frecuencia de Govaerts-Denooz (1974) y Santiago (1998)

LASLA: Histograma de los 100 primeros casos en una lista de frecuencias.



VOCABULARIO NUCLEAR: Histograma de las 120 primeras palabras .



Se puede observar cómo la línea en ambas gráficas está sesgada hacia la derecha, es decir, el número de palabras nucleares o frecuentes disminuye paulatinamente. Sin embargo, mientras que en el vocabulario de frecuencia el salto de valores altos a bajos de frecuencia es muy brusca, en el vocabulario basado en la nuclearidad se produce gradualmente. Esto quiere decir que, aunque las diferencias entre las palabras en cuanto a frecuencia pueden ser muy acusadas, la diferencia en cuanto a la importancia que da la lengua a partir de los criterios lingüísticos no es tan grande.

Respecto a las características morfológicas de las palabras de este vocabulario latino podemos decir que está formado en un porcentaje bastante alto por palabras simples, si son adjetivos, serían de la 1ª clase y de 3 terminaciones, si son sustantivos, de la 3ª declinación y de la 2ª, y si son verbos, serían casi todos de la 3ª conjugación. Por otra parte, a la hora de expandirse, el sistema nominal prefiere la derivación, mientras que el sistema verbal lo hace por medio de la composición. Esto nos puede dar pistas sobre la morfología de las palabras que integren un lista didáctica de palabras en latín.

Si dividimos en grupos las palabras que hemos considerado según hayan superado determinado número de estas pruebas, observamos que los sinónimos suelen encontrarse en distinto grupo, mientras que los antónimos se encuentran en el mismo. El verbo *cipio* del grupo A, el más nuclear tiene posibles sinónimos en el grupo B (*accipio*, *rapio*, *emo*), que representan distintas formas para ‘tomar’. Igualmente ocurre con *quaero* e *invenio* del mismo grupo cuyos posibles sinónimos se encuentran en el grupo B, *requiro* y *reperio*, respectivamente. *Magnus* en el grupo A tiene como sinónimos a *amplus* y *grandis* en el grupo B. *Committo* en el grupo A, con el significado ‘unir’, ‘poner en contacto’ puede tener como sinónimos varios verbos en el grupo B: *iungo*, *contineo* y *vincio*.

Por el contrario, los antónimos hay que buscarlos en el mismo grupo de nuclearidad cuando los tienen. Así están en el mismo grupo los siguientes pares: *magnus/parvus*, *ostendo/occulo*, *committo/divido*.

Aunque se pueden encontrar excepciones, no deja de resultar intere-

sante considerar la hipótesis de que un vocabulario nuclear bien descrito no debería admitir sinónimos (llamémosles co-hipónimos), ya que van contra la economía del lenguaje, pero sí antónimos; que una primera expansión del vocabulario nuclear supondría ampliar con palabras más específicas los conceptos contenidos en las palabras nucleares del primer nivel (hipónimos); y que los antónimos se encontrarían en el mismo nivel de nuclearidad.

3. LAS IMPLICACIONES DIDÁCTICAS

El vocabulario nuclear nos puede servir tanto para seleccionar textos con un nivel determinado como para adaptarlos. A la hora de juzgar si un texto es adecuado o no para el nivel de los alumnos hay que recurrir al *índice de densidad* (Clemente 1995: 197), que es la relación entre el número el número de lemas (*types*) y el de ocurrencias (*tokens*) para después comprobar en qué medida se encuentran las palabras nucleares en el texto. Puesto que la nuclearidad no es una medida estadística, he recogido en un cuadro la repartición de las palabras del vocabulario nuclear en varias muestras de autores (en concreto, *Liber Gromaticus* de Frontino, *Fragmenta* de Nepote y *De coniuratione Catilinae* de Salustio) y de un libro de texto según estas palabras cumplieran con las pruebas pruebas indicadas en el punto II.1.

Presencia relativa de grupos del VN en la muestra

| | Fr.-Nep. | Sal. Cat. | Anaya |
|---|----------|-----------|-------|
| A | 39,8% | 47% | 47,7% |
| B | 26,1% | 34,4% | 35,3% |
| C | 24,2% | 28% | 24,2% |
| D | 13,1% | 29,5% | 32,8% |
| E | 11,1% | 44,4% | 33,3% |

Se observa aquí que hay más palabras del grupo A, el que contiene las palabras más nucleares que del B, y así sucesivamente con los otros grupos, especialmente en Frontino y Nepote, que son textos completos.

Sin embargo, en las otras dos calas formadas por diez muestras de unas 150 palabras cada una se mantiene esta tendencia hasta el grupo C, pero se invierte en los grupos D y E. Creemos que se debe al hecho de que se trata de calas, por lo que necesitaríamos analizar muestras bastante más amplias para comprobar la presencia por grupos de palabras del vocabulario nuclear. Así pues, esto demostraría que los textos también se organizan en torno a unas palabras que consideramos lingüísticamente más importantes.

La adaptación y selección de textos adecuados para traducir es una de las utilidades primeras del vocabulario nuclear, pero no se queda ahí, pues creemos que debe formar parte del programa léxico del libro, con lo que decimos que hay que aprovecharlo también para aclarar paradigmas y para confeccionar ejercicios específicos para el aprendizaje del léxico. La propia estructura de la lengua proporciona la base para jugar con sinónimos y antónimos, como hemos visto anteriormente, pero también se puede recurrir a redes asociativas (Erdmenger 1985) y a otros ejercicios que requieren mayor elaboración (Carbonell 1997, Santiago 1997).

No obstante, quisiera recalcar que no debe olvidarse la importancia de una lista consensuada de palabras a la hora de evaluar la consecución de cualquier objetivo que implique al léxico, sea de frecuencia, sea de nuclearidad, aunque estimo que ésta es más eficaz al considerar criterios lingüísticos.

4. CONCLUSIÓN

Los inconvenientes que puede presentar la frecuencia, así como el poco aprovechamiento que se ha hecho de ella en la didáctica del latín, nos ha llevado a buscar en otra parte una lista de palabras que sea realmente útil desde el primer momento. Todas las observaciones hechas por la lingüística sobre la organización estructural de los contenidos de las lenguas han llevado a plantear el léxico como una estructura en la que el centro estaría constituido por una serie de palabras que presentan características comunes. A partir de ahí, la labor ha sido descubrir cuáles son estas palabras y cómo hacerlo. Puesto que están marcadas lin-

güísticamente, Stubbs y otros crearon una serie de pruebas, de las que nosotros hemos adaptado cinco dado el carácter de lengua de *corpus* del latín y los instrumentos de que disponemos. La finalidad obvia de encontrar estas palabras es didáctica, pues puede hacer que la enseñanza sea más eficaz.

En nuestra investigación aplicamos estas pruebas a 1.000 palabras, resultando de la suma un índice total que llamamos de nuclearidad, que nos sirve para clasificarlas por orden decreciente según su importancia lingüística. Además, hemos visto que la curva de *nuclearidad* tiene gran parecido con la de frecuencia y que todo el conjunto observa ciertas características morfológicas, así como alguna regularidad en la posición de antónimos y sinónimos.

Pero donde mayor resulta su utilidad es en la propia didáctica del latín, ya que al tratarse de criterios lingüísticos tiene mayor eficacia a la hora de trabajar el léxico. Una de las aplicaciones que nos parece fundamental es la selección y adaptación de textos con un criterio más científico de lo que se ha hecho hasta ahora, así como la inclusión del léxico en los programas de los libros de texto, lo que redundará en beneficio del aprendizaje del latín.

BIBLIOGRAFÍA

- BRÉAL, M. y BALLY, CH. (1937): *Les mots latins groupés d'après le sens et l'étymologie*, París.
- CARBONELL, J. (1997): "El reto de una metamorfosis: del latín *per se* al latín instrumental", en *Aspectos didácticos del latín. 5*, Zaragoza, ICE, Universidad de Zaragoza, pp. 169-215.
- CARTER, R.A. (1987): *Vocabulary applied linguistics perspectives*, Londres, George Allen and Unwin.
- CLEMENTE ESTEBAN, R.A. (1995): *Desarrollo del lenguaje. Manual para profesionales de la intervención en ambientes educativos*, Barcelona, Ediciones Octaedro.
- COSERIU, E. (1977): *Principios de semántica estructural*, Madrid [1966]. *Diccionario Latino*, Madrid, CSIC [en curso desde 1984].
- DUMESNIL, G. (1815): *Synonymes latins, et leurs différentes significations*

- avec des exemples tirés des meilleurs auteurs, à l'imitation des synonymes de la'abbé Girard*, París [1777].
- ERDMENGER, M. (1985): "Word acquisition and vocabulary structure in third-year EFL-learners", *IRAL*, XXIII/2, pp. 159-164.
- GAFFIOT, F. (1985): *Dictionnaire Latin Français*, París, Hachette [1934].
- GARCÍA-HERNÁNDEZ, B. (1997a): "Sinonimia y diferencia de significado", *RSEL*, 27/1, pp. 1-31.
- (1997b): "Sinonimia y diferencia de significado", *RSEL*, 27/2, pp. 381-407.
- GARDNER, D.D. (1971): *A frequency dictionary of Latin words*, Michigan, UMI, Ann Arbor, 4 vols.
- GHADESSY, M. (1979): "Frequency counts, word list and material preparations: a new approach", *Forum*, 17/1, pp. 24-27.
- GOUGENHEIM, G. (1958): *Dictionnaire fondamental de la langue française*, París, Didier.
- GOUGENHEIM, G., MICHÉA, R., RIVENC, P. y SAUVAGEOT, A. (1964): *L'elaboration du français élémentaire (1^{er} degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*, París, Didier [1956].
- GOVAERTS, S. y DENOOZ, J. (1974): *Lexique de base latin*, Lieja, (L.A.S.L.A.), Universidad de Lieja.
- LODGE, G. (1907): *The vocabulary of high school latin*, New York.
- Longman dictionary of contemporary english* (1978), P. Procter, (ed.) Avon, Longman.
- LYONS, J. (1981): *Introducción en la lingüística teórica*, Barcelona, Teide [1968].
- MARTÍN MINGORANCE, L. (1990): "Functional Grammar and lexematics in lexicography", en Tomaszczyk y Lewandowska-Tomaszczyk (eds.), *Meaning and lexicography*, Amsterdam, John Benjamins, pp. 227-253.
- MARTINET, A. (1987): *Sintaxis general*, Madrid, Gredos [1985].
- MATHY, M. (1952): *Vocabulaire de base du latin*, París, Editions OCDL.
- MEARA, P. (1980): "Vocabulary acquisition: a neglected aspect of language learning", *Language Teaching and Linguistics: Abstracts* 15, 4, pp. 221-246.
- MENGE, H. (1959): *Lateinische Synonymik*, Heidelberg, Carl Winter-Universitätsverslag [1822].
- MIGUEL, R. de (1958): *Nuevo diccionario Latino-Español etimológico*, se-

guido de un tratado de sinónimos y un vocabulario español-latino, Madrid, Victoriano Suárez [1867].

NATION, I.S.P. (1983): *Teaching and learning vocabulary*, Wellington, University of Wellington, English Language Institute.

OGDEN, C.K. (1930): *Basic english: a general introduction*, Londres, Kegan Paul, Trench and Trubner.

Oxford Latin Dictionary (1982), P.G.W. Glare (ed.), Oxford, Clarendon Press.

SANTIAGO, J. M. (1997): “Principios para el establecimiento de un vocabulario básico en latín y aplicaciones prácticas”, en *Aspectos didácticos del latín. 5*, Zaragoza, ICE, Universidad de Zaragoza, pp. 141-168.

— (1998): *Confeción de un vocabulario básico-nuclear para el latín: criterios y métodos científicos*, tesis doctoral, Universidad del País Vasco.

SAVARD, J.-G. (1970): *La valence lexicale*, París, Didier.

STUBBS, M. (1986): “Language development, lexical competence and nuclear vocabulary”, en *Educational Linguistics*, Oxford, Blackwell, pp. 98-115.

Thesaurus Linguae Latinae editus auctoritate et consilio academiarum quinque germanicarum Berolinensis Gottingensis Lipsiensis Monacensis Vindobonensis, Leipzig, Teubner [en curso desde 1900].