

MOTIVACIÓN LINGÜÍSTICA Y CRITERIOS EN EL
ESTABLECIMIENTO DE LA PARCELACIÓN LÉXICA:
EXAMEN DESDE LA PERSPECTIVA
DE LA LINGÜÍSTICA INGLESA

MIGUEL FUSTER
MARÍA DEL MAR MARTÍ
Universitat de València

1. LA RIQUEZA LÉXICA Y LA LEXICOGRAFÍA

Averiguar la suma total de vocablos de una lengua o decidir los criterios selectivos preocupa, no sin razón, más a la lexicología y la lexicografía que a otras parcelas lingüísticas. La riqueza lingüística es un elemento valorado como positivo ideológicamente, como aspecto sociolingüístico, pero estructuralmente es menos relevante. Unas lenguas tienen más vocabulario que otras, sin embargo un menor número de vocablos en una lengua no es indicador de que por ello dicha lengua sea más o menos compleja.

Los hablantes de lenguas occidentales, gracias a una larga tradición literaria, a su gran expansión geográfica y demográfica, así como al liderazgo que ejercen en lo científico y lo tecnológico han aumentado el volumen de su vocabulario en proporciones gigantescas; contrastemos si no la situación a finales del medievo y la situación actual. Por proporcionar algunos detalles, el período que transcurre entre 1530 y 1660 es el de mayor crecimiento del vocabulario en la historia del inglés, y ello fundamentalmente como consecuencia de dos factores: (1) la necesidad de expresar nuevas

ideas en inglés en dominios anteriormente reservados al latín y (2) el deseo de aumentar el potencial retórico de la lengua inglesa (Görlach 1991:136-8). Más que percibirlo como rasgo evolutivo de vital importancia en el desarrollo de la lengua inglesa, lo debemos entender como cambio en las condiciones socio-históricas de las naciones y de las necesidades comunicativas de los hablantes.

Pecaríamos de ingenuos si realizásemos sin más un recuento de vocabulario castellano acudiendo al *Diccionario de la lengua española* y lo comparáramos, digamos, con el volumen léxico de un diccionario inglés de prestigio. Es más, se nos presenta el escollo de diferentes tradiciones lexicográficas y, de hecho, la inglesa es bastante distinta a la española. Parte de la responsabilidad de esta divergencia se debe a las opiniones y a la obra del doctor Samuel Johnson, el cual en su famoso *Preface* a la primera edición del *Dictionary of the English Language* del año 1755, asestó un golpe de gracia a la propuesta tiempo atrás acariciada de constituir una autoridad institucional normativa sobre la lengua inglesa auspiciada por la *Royal Society*. A su parecer las academias se contraponen al espíritu libre de los ingleses. Entre otras razones, para el doctor Johnson el vocabulario es un componente sujeto al incesante vaivén histórico:

When we see men grow old and die at a certain time one after another, from century to century, we laugh at the elixir that promises to prolong life to a thousand years; and with equal justice may the lexicographer be derided, who being able to produce no example of a nation that has preserved their words and phrases from mutability, shall imagine that his dictionary can embalm his language, and secure it from corruption and decay, that it is in his power to change sublunary nature, or clear the world at once from folly, vanity, and affectation (*apud* Baugh y Cable 1993:264-265).

En efecto, la inestabilidad se nos antoja más perceptible en el vocabulario que en otros aspectos, si bien ningún componente lingüístico es ajeno al cambio lingüístico. Nuevas palabras son añadidas continuamente mientras que unos vocablos consolidados aumentan su polisemia y otros caen en desuso. Frente al limitado inventario de fonemas de una lengua, el léxico parece incontrolado.

Varios problemas resisten a cualquier intención de cuantificación precisa. Admitámoslo, si debemos juzgar por lo observado en diccionarios al uso, éstos no son sino el resultado de criterios

selectivos de autores o editoriales, entre los cuales figura el usuario. Sólo basta examinar la amplia gama de diccionarios ingleses para darnos cuenta de la variedad de tamaños y diseños. Por consiguiente, es conveniente no confundir un hipotético cómputo total de palabras en el idioma en un momento concreto de su historia, con las decisiones de los lexicógrafos. He aquí algunos problemas lexicológicos que acechan habitualmente al lexicógrafo:

- a) Indecisión a la hora de incluir los nombres propios –no refrendada por la aportación de la teoría lingüística–. Únicamente los nombres propios de personajes relevantes son ubicados en los márgenes del léxico como conocimiento de índole enciclopédica.
- b) Multitud de vocablos con suficiente transparencia tampoco tendrán cabida, o no tienen entrada propia –¿qué hacer con adverbios terminados en *-ly*, o sustantivos en *-ness*?–.
- c) La terminología científica, excesivamente técnica, suele ser apartada ya que, ¿consideraríamos *equus asinus*, para referirnos a un “asno” como un vocablo inglés, a pesar de que sólo sea corriente entre científicos? Razones no nos faltarían, ya que lexemas como *locus classicus* o *sine prole*¹ son recogidos en diccionarios ingleses, a pesar del aparente “cambio de código” o falta de naturalización exhibida. ¿Es precisa la carta de naturaleza en cualquier vocablo para su consideración como vocablo del idioma?
- d) Un problema de distinta índole sería cómo actuar frente a la homonimia y la polisemia y llegar a una clara distinción entre ambas sin recurrir a la etimología. Sabemos que para multitud de vocablos que aparecen como polisémicos, las relaciones de sentido son distantes (Katamba 2005:124-5).
- e) En el inglés, ¿qué hacer con la increíble cantidad de vocablos que se produce mediante la *conversión*, donde la derivación sin afijo produce una identidad formal prácticamente absoluta en una lengua como la inglesa de morfología bastante aislante? ¿Debería, por ejemplo, contarse *walk* dos veces, una como verbo y otra como sustantivo, para cada una con las acepciones que le corresponden?
- f) ¿Qué hacer con el medio millón de formas abreviadas, muchas de ellas, según Crystal (1995:119), con estatus léxico?

¹ Vocablos recogidos en el *Webster's Encyclopedic Unabridged Dictionary of the English Language*, New York-Avenel: Gramercy Books, 1994.

- g) Otra razón que permite sospechar de un recuento basado en los diccionarios se refiere a su vertiente ideológica: ¿qué fiabilidad concederíamos a un diccionario *normativo*? Los diccionarios normativos no describen normas consensuadas por la comunidad de hablantes sino que dirigen o controlan la conducta léxica de los hablantes. Este problema es más evidente en el caso del castellano que en el inglés, cuyos diccionarios tienden a ser *descriptivos*.

Estas y otras cuestiones nos hacen dudar de que podamos alcanzar un fiable cómputo total aunque en realidad se han hecho. Dupuy (1974) y Goulden, Nation and Read (1990) han examinado el vocabulario inglés a partir del contenido del *Webster's Third International Dictionary*, publicado en 1963. Se trata del mayor diccionario no histórico del inglés publicado hasta dicha fecha. Las palabras fueron clasificadas por familias consistentes en una palabra básica, sus formas de inflexión y las derivaciones más transparentes. En total este diccionario contiene 54000 familias de palabras o, siguiendo un recuento tradicional, unas 450000 palabras. Si añadimos vocabulario científico suficiente, nos acercáramos al millón de palabras (McCrum, Cran y McNeil 1992:1). Dejando a un lado ciertas idiosincrasias, una enorme proporción de dicho léxico es formalmente equivalente al que poseen otras lenguas occidentales, y sus causas encuentran explicación en el contacto secular entre estas culturas y el desarrollo de una comunidad científica siguiendo patrones lingüísticos próximos (Fuster 1995).

2. EL LÉXICO Y OTROS COMPONENTES LINGÜÍSTICOS

Está extendida entre los lingüistas la creencia de que el léxico es la parte menos estructurada o sistematizada de la lengua. Mientras la fonología proporciona unas reglas fonológicas sistemáticas, y la sintaxis unos patrones fijos en cláusulas y oraciones, el vocabulario consiste en innumerables unidades que, en su conjunto, se resisten a cualquier riguroso análisis con relación a estructuras o sistemas. Como resultado de ello, el vocabulario es generalmente apartado de las descripciones lingüísticas o bien es considerado como el componente más externo. Sin embargo, existe una línea de pensamiento que considera que el vocabulario ni es externo ni se

trata de un bloque impenetrable. Los estudiosos de la morfología aceptan la existencia de la estratificación morfo-léxica, mientras que otros dirigen su atención al uso.

En lugar de preguntarnos qué número de vocablos posee un idioma, pregunta que buscaría satisfacer posiciones ideológicas, parece más relevante preguntarse qué vocablos utiliza o comprende un hablante nativo medio y cuál es su naturaleza. Varios experimentos realizados sobre el inglés cuantifican los vocabularios pasivo y activo de los hablantes nativos. Las cifras oscilan entre los 32000 y los 63000 como activos totales, o un máximo de 76000 como vocabulario pasivo total que correspondería a un profesor universitario (Crystal 1995:123).

John Sinclair, principal editor del *Collins Cobuild Dictionary* de 1987, destaca la importancia de un *vocabulario básico*. Observa este lexicógrafo que “una cantidad relativamente pequeña de palabras sirve para expresar un gran porcentaje de lo que se escribe o se habla diariamente”. Así, aunque admitamos que la lengua inglesa y, por ende, cualquier lengua occidental culta, pueda contener innumerables vocablos, no todos ellos tienen igual valor respecto al uso. Esta jerarquización léxica atendiendo al uso ya había sido indicada por James Murray, editor del conocido diccionario histórico *The Oxford English Dictionary*, de donde nos permitimos extraer esta cita:

That vast aggregate of words and phrases which constitutes the Vocabulary of English-speaking men presents, to the mind that endeavours to grasp it as a definite whole, the aspect of one of those nebulous masses familiar to the astronomer, in which a clear and unmistakable nucleus shades off on all sides, through zones of decreasing brightness, to a dim marginal film that seems to end nowhere, but lose itself imperceptibly in the surrounding darkness (1884:xvii, extraído de Hughes 2000:2).

Tanto el trabajo de los lexicógrafos ingleses como la lingüística aplicada, han ido consolidando a lo largo de los años una visión del léxico donde el uso ocupa un lugar central el cual además se ha potenciado gracias a los avances tecnológicos en el desarrollo de los corpus lingüísticos. La investigación llevada a cabo por Sinclair mediante la compilación de un corpus propio sobre el vocabulario inglés es crucial en su decisión sobre la inclusión de vocablos y diseño de entradas concretas. Es más, su trabajo en torno a principios léxicos ha servido de inspiración en el diseño de

materiales didácticos destinados al aprendizaje del inglés como L2. Todo ello atendiendo al uso cotidiano que los nativos hacen de su léxico.

3. IMPORTANCIA DEL LÉXICO COMÚN

Si atendemos a este principio, diremos que no todas las palabras de una lengua presentan la misma utilidad. Entre unas palabras y otras existe una diferencia de uso gradual que puede medirse. El parámetro más requerido a la hora de medir la centralidad de un vocablo es su indicación de *frecuencia* relativa dentro de un corpus suficientemente representativo. Con respecto a este parámetro, lenguas pobremente flexivas como la lengua inglesa cuentan con palabras funcionales en lugares muy destacados. En concreto el artículo invariable *the* ocupa una privilegiada primera posición como vocablo más frecuente. Y, en su conjunto, los cien vocablos más frecuentes del inglés son igualmente esencialmente vocablos funcionales.

En su estudio sobre el *Brown Corpus*, corpus pionero en lengua inglesa que contiene un millón de palabras, Francis & Kucera (1982) indican que un total de 2000 palabras inglesas son suficientes para comprender el 80 por ciento de cualquier texto. Al igual que Francis & Kucera, los autores de *COBUILD*, un corpus más reciente y mucho más amplio, base del *Collins COBUILD English language Dictionary* (Sinclair 1987) indican que unas 15000 palabras suponen una cobertura del 95 por ciento del total de palabras de todo el corpus. En su discusión sobre esta temática, Nation & Waring (1997:9) señalan que 2000 familias léxicas nos proporcionan la comprensión del 80 por ciento de textos escritos en inglés. Es decir, dos de cada diez vocablos resultarían desconocidos. Liu Na y Nation (1985) han investigado si estos vocablos desconocidos podrían deducirse contextualmente, y concluyen que para una comprensión óptima de un texto, sería preciso estar familiarizado con un 95% del léxico general. Esta cifra coincide con la proporcionada por Laufer (1988), el cual sugiere que un conocimiento de 95% del vocabulario común inglés capacitaría a un estudiante de inglés como L2 para comprender un texto. Por su parte, el *Longman Dictionary of Contemporary English* en sus ediciones recientes incluye un

“vocabulario de definición” básico de la lengua inglesa extraído de tablas de frecuencia de corpus muy conocidos, como el *British National Corpus*. Se trata de un vocabulario básico extraído de tablas de frecuencia que, como su nombre indica, sirve para definir cualquier vocablo que aparece en dicho diccionario. Otros diccionarios han seguido trayectorias próximas que hacen patente que, independientemente de la cantidad total de vocablos que posea una lengua y cuál pueda ser su cómputo, el uso que los nativos hacen del vocabulario es un dato altamente significativo. En principio nos señala la existencia de una jerarquización léxica donde distinguiremos entre un *vocabulario general* frente a otro *no general* o más específico.

4. LISTADOS DE LÉXICO COMÚN

La práctica de contar palabras observando su frecuencia se remonta a los tiempos helénicos (Fries & Traver 1960). A lo largo del siglo XX la confección de listados reducidos ha recibido el impulso del interés en la enseñanza/aprendizaje del inglés como L2. En los listados que a continuación ofrecemos la frecuencia es argumento fundamental:

1. *The General Service List* (West 1953) contiene 2000 lexemas y fue desarrollado en los años cuarenta. Las frecuencias de cada lexema están basadas en un corpus de 5000000 de palabras. El rigor en la descripción hace que se considere una de las listas de mayor aceptación de las realizadas hasta hoy.
2. El libro de trabajo del profesor de 300000 palabras (Thorndike & Lorge 1944) contiene 30000 lexemas y 13000 familias de palabras; está basado en un corpus escrito de 18000000 de palabras.
3. El libro de frecuencias del *American Heritage* (Carroll, Davies & Richman 1971) está basado en un corpus escrito de libros de texto utilizados en EE.UU. de 5000000 de palabras. Su ventaja radica en la determinación de la frecuencia de cada palabra de acuerdo con la asignatura y con el nivel de enseñanza.

Más recientemente, los corpus lingüísticos se han desarrollado y multiplicado sobremanera, especialmente para la lengua inglesa.

Sinclair, como responsable del proyecto *COBUILD*, propone considerar como básicas unas 2000 o 3000 palabras inglesas. En su opinión, no conviene reducir excesivamente el listado a 50 o 100 palabras. Estas 2000 o 3000 palabras inglesas son de gran utilidad y de uso complejo, con una acusada tendencia a formar estructuras más complejas con palabras asociadas a este núcleo. La intención es priorizar la utilización de estas palabras al esfuerzo de conocer vocablos cuya ocurrencia es baja.

Por su parte el análisis de frecuencia léxica realizado por McCarthy (1999) en el entorno del discurso hablado, utilizando el CANCODE (una sección del *Cambridge International Corpus* de 100 millones de palabras de la lengua escrita y oral), le permite llegar a conclusiones similares a las de Sinclair. Desafortunadamente, no contamos con una abundante investigación sobre el vocabulario básico en español con el que podamos establecer contrastes o comparaciones suficientemente válidos. Sin embargo, trabajos concretos como los de Ávila Muñoz sobre el habla de Málaga (1999) permiten vislumbrar que el castellano hablante nativo utiliza un reducido número de vocablos cuya alta frecuencia los distingue del resto. López Morales (1986) señala que por regla general el hablante educado de español puede utilizar de 4000 a 5000 palabras básicas, mientras que un hablante menos educado utiliza alrededor de 2000 palabras. El interés de este autor es en la *disponibilidad léxica*, que no entramos a valorar aquí. Tampoco entramos a valorar de qué manera se da entrada a los vocablos en cada uno de los cómputos y sería preciso considerar aspectos morfológicos tales como la flexión a la hora de contrastar con lenguas como el inglés. Así, mientras que en inglés el artículo determinado es invariable y figurará en primer lugar, por su morfología el español posee cuatro formas diferentes que, igualmente, ocuparán un lugar preferente.

5. CRITERIOS PARA LA DESCRIPCIÓN DEL VOCABULARIO BÁSICO

En un artículo reciente, Fuster y Martí (2000) realizan una revisión crítica de los distintos argumentos que han ido utilizándose en diversas descripciones lingüísticas del inglés a la hora de determinar la existencia de un vocabulario básico, tanto en la dimensión sincrónica como en la diacrónica. Estos autores

encuentran que el término “core” o “basic” para referirse al vocabulario central no se utiliza con criterios suficientemente justificados empíricamente. De modo particular, los listados de léxico básico utilizados en diacronía parecen buscar satisfacer posiciones ideológicas, con la intención de procurar para el inglés una historia lingüística lo suficientemente prolongada mediante la decisión de que, al igual que otros aspectos de la lengua, su vocabulario básico es eminentemente anglosajón o germánico y no adquirido por vía de contacto (Milroy 2002:7-27; Haspelmath 2003). Así, por ejemplo, afirmar la procedencia anglosajona del léxico básico inglés actual, criterio mantenido por muchísimos autores (Hock & Joseph 1996:257 y ss.; Thomason 2001:90), es un argumento dudoso a tenor de investigaciones recientes (Hughes 2000:391-394), las cuales apuntan en un sentido bastante diferente. Hagamos repaso de las características del léxico básico en clave sincrónico/diacrónica:

- I. *Aculturalidad del vocabulario básico*. El vocabulario acultural sería resistente al préstamo. Crowley (1992:168-9) señala que todas las lenguas del mundo poseen vocablos no culturales, propios de la existencia humana, los cuales teóricamente no son permeables a la transferencia mediante contacto. El problema es que, con la probable excepción de los vocablos funcionales, es difícil determinar la neutralidad de multitud de vocablos. Una propuesta válida sería acudir a la teoría de la *semántica natural* desarrollada por Wierzbicka y Goddard (Wierzbicka 1995, 1996; Goddard & Wierzbicka 1994), los cuales basan su argumentación en torno al léxico en pruebas de la existencia de un grupo reducido y básico de universales del significado, los *primitivos semánticos*. Los primitivos semánticos pueden encontrarse en las distintas lenguas como palabras u otras expresiones lingüísticas. A su favor tienen la base empírica mediante la observación de numerosas lenguas. Un cruce de estos primitivos semánticos con listados de vocablos más frecuentes revela de hecho un grado elevado de coincidencia, ya que se encuentran entre los 300 vocablos más básicos.
- II. *La distinción entre clases cerradas y clases abiertas*. Estudiosos de la gramática descriptiva identifican dos grupos distintos de palabras: *abiertas* vs. *cerradas*, en línea con Sweet (1891) y que tiene una amplia aceptación en la actualidad. Pertenecen a

las clases abiertas las denominadas palabras léxicas o de contenido, mientras que en las clases cerradas se sitúan las palabras funcionales o gramaticales. La distinción semántica entre palabras vacías y palabras llenas (Lyons 1995:65-6) se vincula a esta distinción. Un problema es que la diferencia entre términos funcionales y no funcionales no es absolutamente nítida como han señalado varios estudios. Fijémonos en que esta distinción entre dos clases descansa sobre una base diacrónica: decimos que los miembros de una clase abierta lo son porque los nativos continuamente van incorporando nuevas unidades. Podríamos considerarlos como *clase inestable* o *mutable*, en contraste con los miembros de clase cerrada, los cuales no aceptan fácilmente la incorporación de nuevas unidades y su número es reducido, siendo su rasgo fundamental la *estabilidad*. Sin embargo, la historia nos muestra que multitud de vocablos de clase abierta son resistentes al cambio (por ejemplo, *padre, madre, ojo, mano*, etc.). Además, no es difícil demostrar tampoco que con el paso del tiempo, el léxico básico de la lengua también sufre alteraciones. Así, el léxico básico del inglés antiguo, considerados éstos como sus 2000 vocablos más frecuentes, ha cambiado radicalmente hasta el presente, un milenio después.

- III. *El monosilabismo*. Un rasgo de las palabras básicas inglesas es su monosilabismo. Este rasgo está vinculado a cambios morfofonológicos, y al hecho de haber pasado de lengua flexiva a lengua aislante. Palabras como *give, get, take, tell*, eran polisilábicas en inglés antiguo, por consiguiente no sería válido en clave diacrónica. Es más, dentro de las 1000 más básicas hay multitud de vocablos polisilábicos tanto anglosajones como romances o híbridos (*e.g. about, people, many, between, because*).
- IV. *La polisemia del vocabulario básico*. Carter (1987:185) observa que la enseñanza de las palabras básicas es una tarea ardua por la multiplicidad de significados que exhiben. Sin embargo no queda claro que las palabras funcionales deban figurar como léxicas, ya que siendo gramaticales, podría pensarse que no tienen significado propio sino que forman parte de la estructura de la lengua.

- V. *El vocabulario básico pertenece al registro oral.* Para Stubbs (1996:63 y ss.) las palabras del registro hablado son informales, comunes, frecuentes, monosilábicas, de procedencia germánica, activas y adquiridas. Conocido es que el inglés contiene numerosos sinónimos, en gran medida debido al contacto (Jackson y Zé Amvela 2000:96-97). Según Stubbs (1996:63 y ss.), en estas situaciones, el vocablo más básico pertenece al registro oral, y sería nativo, mientras que su sinónimo se colocaría en el registro escrito del inglés y sería producto del préstamo. He aquí algunos ejemplos significativos:

<i>English</i>	<i>Romance</i>
answer	reply
buy	purchase
first	prime
inner	internal
land	territory
world	universe

Mientras que las palabras de la primera columna son anglosajonas, las de la segunda son préstamos romances. Sin embargo, el problema en esta observación de la *disociación léxica* del vocabulario inglés (McArthur 1992:131) no puede aplicarse a multitud de vocablos romances cuyo uso está empíricamente demostrado que es cotidiano y que se han ubicado perfectamente en el registro oral, como se aprecia en los ejemplos que hemos dado anteriormente y muchos otros que figuran entre los 1000 vocablos más comunes del inglés (Fuster 2004).

- VI. *La frecuencia.* Este argumento ha resultado ser fundamental en la determinación del léxico básico o común. Tal y como señala Scholfield, colaborador del *Longman Dictionary of Contemporary English* (traducción mía):

La frecuencia ha sido considerada durante mucho tiempo como criterio clave en la determinación de los vocablos que son más o menos importantes o qué vocablos del inglés son básicos. La frecuencia también guía a los lexicógrafos en la presentación y ordenación de la información y se utiliza en estilística en el reconocimiento de diferentes variedades del inglés.

Lexicógrafos, lingüistas aplicados y teóricos han hecho uso de la frecuencia mediante numerosos listados publicados. La razón estriba en que la frecuencia de un vocablo en un gran corpus de lengua está estrechamente relacionada con su centralidad en el uso. En su discusión sobre la productividad morfológica, Haspelmath (2002:43-44) reivindica la frecuencia como argumento pues en sus propias palabras “speakers remember a word not only if it is unpredictable, but also if it is very frequent”, siendo ésta una característica cognitiva común a animales y seres humanos. Es más, la frecuencia de los vocablos puede medirse empíricamente y puede decirse que existe un salto brusco entre los elementos más frecuentes y los que no lo son. McCarthy (1999: 234) indica:

Existe un punto en el que la frecuencia disminuye de forma brusca, desde las palabras básicas que muestran una gran frecuencia a las palabras que se dan de forma infrecuente, las frecuencias no decaen de forma regular sino que existe una caída repentina desde las palabras más frecuentes a las menos frecuentes (traducción es nuestra).

6. EL LISTADO DE SWADESH Y LA DIACRONÍA

Investigaciones como las de McMahon y McMahon (2000) o McMahon (2004) entre otras muestran un resurgimiento de aplicaciones basadas en la lexicoestadística y la glotocronología en clave diacrónica. La justificación de esta revisión se relaciona con la necesidad de establecer comparaciones con recientes investigaciones arqueológicas y avances en genética, por lo que nada podría parecer más apropiado que los listados de 100 y 200 palabras confeccionados por Morris Swadesh, y conocidos como listados de Swadesh. Un problema de estos listados es que no tienen base empírica, sino que son meramente intuitivos. A pesar de ello son muchos los lingüistas que han recurrido a los mismos para argumentar sobre el origen o trayectoria de lenguas concretas. Así, para el inglés, Thomason (2001:90), y más recientemente Winford (2003:59), por mencionar dos casos representativos, argumentan a favor de su carácter germánico fundamentándose en estos listados que, dicho sea de paso, supuestamente no contienen vocablos culturales (Thomason

2001:72). Para Haspelmath (2003:2), tanto Hock y Joseph (1996:257) como Thomason (2001:71-2) proporcionan en sus discusiones una noción vaga y poco útil del vocabulario básico o central. Pero además resulta limitada, ya que normalmente se hace alusión a los 100 vocablos más comunes para la determinación genealógica. En opinión de Campbell (2003:271) deberíamos extremar la precaución en la utilización de estos listados. Simplemente observemos el léxico del turco y persa actuales, con gran abundancia de arabismos en su vocabulario básico, o la misma lengua inglesa, con numerosos préstamos romances. Para Campbell, “the problem of loans and potential loans is very serious”. Si ampliamos la lista hasta los 1000 vocablos más comunes, los resultados cambian drásticamente (*vid.* Hughes 2000; Fuster y Martí 2000), ya que el elemento germánico-nativo se reduce hasta el 50 por ciento.

7. CONCLUSIÓN

Ante todo esperamos haber hecho patente la existencia de una parcelación léxica en las lenguas naturales. Si bien las lenguas exhiben una relativa mayor o menor riqueza en el vocabulario, cuyos límites se nos antojan difíciles de precisar, a tenor de las investigaciones llevadas a cabo, parece evidente que existe un vocabulario central donde el factor determinante sería el uso. Este vocabulario básico puede medirse y se ha medido empíricamente mediante el criterio de la frecuencia en la lengua inglesa. No debemos descartar los avances en la semántica léxica como, por ejemplo, en la búsqueda de primitivos semánticos. Sin embargo, el recurso a listados intuitivos, como los ofrecidos por Swadesh, es problemático. Coincidimos pues con Haspelmath (2003:6) en introducir y entender el vocabulario básico mediante el uso del factor frecuencia también en su aplicación diacrónica:

One way in which the notions “basic vocabulary” and “core vocabulary” can be interpreted is as *the words which are used most frequently*. It would not be surprising if they were resistant to borrowing, because it is well known that high-frequency items are resistant to other types of language change such as analogy.

Curiosamente, un cruce entre resultados de frecuencia léxica en inglés tomando como referencia el *British Nacional Corpus*, con 100 millones de palabras, indica que muchos de los vocablos de Swadesh no son los más frecuentes en el inglés actual. Muchos de ellos ni siquiera se encuentran entre los mil más frecuentes. Para dar una breve muestra de esta distancia, la siguiente tabla con algunos ejemplos extraídos de los cien más básicos de Swadesh y su posición en el *British Nacional Corpus* reflejan a las claras estas enormes distancias:

Swadesh (100 básicos)	BNC
<i>short</i>	588
<i>husband</i>	833
<i>heavy</i>	970
<i>wide</i>	649

Si atendemos a la frecuencia, muchísimos vocablos romances figuran como más frecuentes o centrales en el uso en el inglés actual que tantos otros de origen anglo-sajón o germánico.

REFERENCIAS BIBLIOGRÁFICAS

- ÁVILA MUÑOZ, A. M. (1999): *Léxico de frecuencia del español hablado en la ciudad de Málaga*, Málaga: Universidad de Málaga.
- BAUGH, A. C. & CABLE, TH. (1993): *A History of the English Language*, Englewood Cliffs, NJ.: Prentice-Hall, 4th ed.
- CAMPBELL, L. (2003): "How to show languages are related: Methods for distant genetic relationship", en B. D. Joseph y R. D. Janda (eds.), *The Handbook of Historical Linguistics*, Oxford: Blackwell, 262-283.
- CARROL, J. B.; DAVIES, P. y RICHMAN, B. (1971): *The American Heritage Word Frequency Book*, New York: American Heritage Publishing Co.
- CARTER, R. (1987): *Vocabulary: Applied Linguistic Perspectives*, London/New York: Routledge.
- CROWLEY, T. (1992): *An Introduction to Historical Linguistics*, Oxford: Oxford University Press, 2nd ed.
- CRYSTAL, D. (1995): *The Cambridge Encyclopedia of the English Language*, Cambridge: C.U.P.

- DUPUY, H. J. (1974): *The rational development and standardization of a basic word vocabulary test*, Washington, D.C: US Government Printing Office.
- FRANCIS, W. N. & KUCERA, H. (1982): *Frequency analysis of English usage*, Boston: Houghton Mifflin Company.
- FRASER, B. (1990): "An Approach to Discourse Markers", *Journal of Pragmatics*, 14, 383-395.
- FRIES, R. C. & TRAVER, A. A. (1960): *English Word List*, Ann Arbor: George Warh.
- FUSTER, M. (1995): "La romanización del inglés y la anglización del español peninsular", en M.^a T. Echenique, M. Aleza y M.^a J. Martínez (eds.), *Actas del I Congreso de Historia de la lengua española en América y España*, València: Universitat de València/Tirant lo Blanch, 245-261.
- FUSTER, M. (2004): "Ubicación del elemento romance medieval en el inglés actual", en C. Calvo, E. Casanova y F. Latorre (eds.), *Lingüística Diacrónica Contrastiva*, Valencia: Universitat de València, Quaderns de Filologia: Estudis Lingüístics IX, 69-86.
- FUSTER, M. y MARTÍ, M.^a DEL M. (2000): "Contact and basic English vocabulary", *Studies in English Language and Linguistics*, Universitat de València, vol. 2, 97-117.
- GODDARD, C. y WIERZBICKA, A. (eds.) (1994): *Semantic and Lexical Universals - Theory and Empirical Findings*, Amsterdam/Philadelphia: John Benjamins.
- GODDARD, C. y WIERZBICKA, A. (eds.) (1995): *Semantic and Lexical Universals: Theory and Empirical Findings*, Amsterdam/Philadelphia: John John Benjamins.
- GÖRLACH, M. (1991): *Introduction to Early Modern English*, Cambridge: C.U.P.
- GOULDEN, R.; NATION, P. y READ, J. (1990): "How large can a receptive vocabulary be?", *Applied Linguistics*, 11, 4, 341-63.
- HASPELMATH, M. (2002): *Understanding morphology*, London: Arnold.
- HASPELMATH, M. (2003): "Loanword Typology: Steps toward a systematic cross-linguistic study of lexical borrowability". Publicación electrónica en: <http://email.eva.mpg.de/~haspelmt/LWT-text.pdf>
- HOCK, H. H. y JOSEPH, B. (1996): *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*, Berlin/New York: Mouton de Gruyter.
- HOLMES, J. (1988): "Doubt and certainty in ESL textbooks", *Applied Linguistics*, 9, 1, 29-44.
- HUGHES, G. (2000): *A history of English words*, Oxford: Blackwell.

- JACKSON, H. y ZÉ AMVELA, E. (2000): *Words, meaning and vocabulary*, London/New York: Cassell.
- KATAMBA, F. (2005): *English words: structure, history, usage*, Abingdon: Routledge, 2ª ed.
- LASS, R. (1987): *The Shape of English: Structure and history*, London/Melbourne: J. M. Dent and Sons Ltd.
- LAUFER, B. (1988): "What percentage of text-lexis is essential for comprehension?", en C. Lauren y M. Nordmann (eds.), *Special Languages: From Humans thinking to Thinking Machines*, Clevedon: Multilingual Matters.
- LIU, N. y NATION, S. P. (1985): "Factors affecting guessing vocabulary in context", *RELC Journal* 16, 1, 33-44.
- LÓPEZ MORALES, H. (1986): *La enseñanza de la lengua maternal*, Madrid: Playor.
- LYONS, J. (1995): *Linguistic Semantics: An introduction*, Cambridge: C.U.P.
- MCCARTHY, M. (1999): "What constitutes a basic vocabulary", *Studies in Language and Linguistics*, 1, 233-55.
- MCCRUM, R.; CRAN, W. y MACNEIL, R. (1992): *The Story of English*, New York: Penguin.
- MCMAHON, A. (2004): "Language, time and human histories", en F. Brisard, S. D'Hondt y T. Mortelmans (eds.), *Language and revolution/Language and time*, Antwerpen: Universiteit Antwerpen, Antwerp papers in linguistics, 106, 155-171.
- MCMAHON, A. y MCMAHON, R. (2000): "Problems of dating and time depth in linguistics and biology", en C. Renfrew, A. McMahon & L. Trask (eds.), *Time depth in historical linguistics*, Cambridge, England: The McDonald Institute for Archaeological Research Publications, cap. 4.
- MILROY, J. (2002): "The Legitimate Language: Giving English a History", en R. Watts & P. Trudgill (eds.), *Alternative Histories of the English Language*, London/New York: Routledge, 7-27.
- NATION, I. S. P. (2001): *Learning Vocabulary in Another Language*, Cambridge: Cambridge University Press.
- NATION, I. S. P. y WARING, R. (1997): "Vocabulary size, text coverage and word lists", en N. Schmitt y M. McCarthy (eds.), *Vocabulary: Description, Acquisition and Pedagogy*, Cambridge: Cambridge University Press, 6-20.

- SCHIFFRIN, D. (1987): *Discourse Markers*, Cambridge: Cambridge University Press.
- SCHOLFIELD, Ph.: "What's New about Word Frequency?". Publicación electrónica en: <http://www.longman.com/dictionaries/teachers/articles/p-scholfield-01.html>.
- SINCLAIR, J. (ed.) (1987): *The Collins Cobuild English Language Dictionary*, London/Glasgow: Collins.
- STUBBS, M. (1996): *Text and Corpus Analysis*, Oxford: Blackwell.
- SWEET, H. (1891): *A New English Grammar: Logical and Historical*, Oxford: Clarendon.
- THOMASON, S. G. (2001): *Language Contact*, Edinburgh: Edinburgh University Press.
- THORNDIKE, E. L. y LORGE, I. (1944): *The teacher workbook of 30.000 words*, New York Teacher's College: Columbia University.
- WEST, M. (1953): *A General Service List of English Words*, London: Longman.
- WIERZBICKA, A. (1995): "Universal Semantic Primitives as a basis for lexical semantics", *Folia Linguistica*, vol. XXIX (1-2), 149-69.
- WIERZBICKA, A. (1996): *Semantics: Primes and Universals*, Oxford/New York: Oxford University Press.
- WINFORD, D. (2003): *An introduction to Contact linguistics*, Oxford: Blackwell.