El español en cifras

AQUILINO SÁNCHEZ
Universidad de Murcia

En esta nueva sección de la revista *Carabela* se ofrecerán datos cuantitativos relativos al uso del español actual en todos sus ámbitos (geográficos, sociales, tipológicos...), en las modalidades oral y escrita. Los datos cuantitativos son útiles para que el profesor pueda tomar algunas decisiones razonablemente fundadas en relación con lo que debe o conviene enseñar en determinados casos.

Los datos de esta sección están basados en el corpus *Cumbre* (propiedad de SGEL). *Cumbre* consta de 20 millones de palabras (equivalentes a unos 200 libros de 300 páginas cada uno) y es representativo del español oral y escrito usado en todos los países de habla hispana en la última década del siglo XX. Este corpus ya ha servido de base para dos obras importantes: un diccionario del español actual (*Gran Diccionario de uso del español actual*, SGEL, 2001) y una gramática (*Gramática práctica del español actual*, SGEL, 2005).

Frecuencia de lemas y palabras en el español actual

Se ofrecen a continuación dos listados relativos al español usado actualmente:

El **listado 1**, recoge los 500 lemas más frecuentes del español. La agrupación contiene todas las formas flexivas de cada lema, como si se tratara de voces o artí-

culos de un diccionario. Se han eliminado del listado las palabras funcionales, por considerar que tales palabras son poco relevantes de cara a la utilidad de esta lista en el ámbito de la enseñanza del español como LE.

El **listado 2** recoge las 500 palabras o formas más usadas en el español actual. Se observará que ambos listados no son coincidentes, ya que el listado 2 recoge la frecuencia de cada forma tal cual aparece, sin asociarla o incorporarla al lema al cual pertenece.

Los estudios realizados para obtener este tipo de listados se iniciaron ya a finales del siglo XIX y en el caso del español se concretaron por vez primera en la obra de Víctor García Hoz (*Vocabulario usual, común y fundamental,* 1953)[1]. Pero su fiabilidad era escasa, debido al reducido número de palabras y textos sobre los que se fundamentaban los resultados. En la actualidad, el cálculo de la frecuencia léxica es fácil de obtener mediante un ordenador si se dispone de una recopilación representativa del español actual. En el presente caso nos basamos en el corpus *Cumbre*, que ofrece un alto grado de fiabilidad en este ámbito. Los dos listados que se ofrecen a continuación pueden ser utilizados, por tanto, con plena garantía en relación con el español usado en la última década del siglo XX.

¿Qué utilidad pueden tener estos listados para el profesor de E/LE? La respuesta a esta pregunta está relacionada con el grado de eficacia o eficiencia del aprendizaje propiciado en el aula: no es de igual utilidad aprender primero los verbos *ser, saber, haber, estar, tener, hacer*, que son los lemas más utilizados en español, que las palabras *demostrar, científico, causar, funcionar,* que figuran en la posición 500 de ese mismo listado. En realidad el aprendiz encontrará más menudo '*ser o hacer*' que '*científico o funcionar*'. Naturalmente, el caso es más llamativo si contrastamos esos verbos con palabras como '*altea, especialidad*', que aparecen en la lista en la posición 3.777, o con '*reconsiderar*', que aparece en la posición 10.000.

Ambas listas constituyen, pues, puntos de referencia indispensables para tomar decisiones sobre qué palabras enseñar primero o en los primeros niveles. Con ese fin se incluyen en esta sección, que seguirá aportando datos numéricos relevantes en relación con los usos del español actual.

---

[1] Véase 27. Sánchez, A. (2000), "Language Teaching before and after 'Digitilized Corpora'. Three main issues", *Cuadernos de Filología Inglesa,* Servicio de Publicaciones, Universidad de Murcia, 9.1, pp. 5-37.

1. Los 500 lemas léxicos más frecuentes del español[2]

ser	poner	trabajar
saber	país	volver
haber	vida	nada
estar	hombre	contar
tener	sobrar	problema
hacer	parecer	durante
más	sobre	hoy
todo	nuevo	pensar
poder	mucho	luego
ir	hablar	creer
decir	cada	conocer
unir	dejar	salir
otro	siempre	hijo
año	llevar	propio
comer	entonces	social
ver	parte	además
muy	casar	sentir
también	tan	lugar
bueno	persona	claro
vez	seguir	hora
mismo	quedar	tal
dar	entre	derecho
parar	encontrar	cosa
querer	después	español
día	mujer	mano
alguno	mundo	niño
primero	venir	mirar
deber	menos	importante
grande	aquí	ejemplo
bien	momento	gente
entrar	aunque	padre
llegar	general	cierto
sólo	formar	contra
donde	último	hacia
pasar	señor	nacional
tiempo	llamar	caso
político	crear	obra
ahora	algo	público

[2] Según el corpus *Cumbre*, de 20 millones de palabras. No se incluye la mayor parte de palabras 'funcionales' (artículos, preposiciones, demostrativos, pronombres personales, posesivos, numerales y similares), aunque sean de frecuencia muy alta.

cual
tipo
grupo
vivir
poco
medio
tratar
santo
noche
casi
tomar
ninguno
cualquier
producir
presentar
realizar
según
cuanto
mientras
millón
empezar
manera
nunca
existir
preguntar
coser
agua
económico
empresa
sociedad
buscar
historia
cambiar
fin
ciudad
permitir
verdad
siglo
dentro
partir
ahí
ley
presidente
forma

esperar
determinar
alto
joven
gustar
palabra
aún
pequeño
situación
gobierno
casa
gobernar
fuente
varios
información
desarrollar
único
utilizar
relación
nombre
servicio
cuatro
mañana
trabajo
posible
abrir
allí
lado
cuyo
libro
considerar
oír
recibir
escribir
pedir
ojo
resultar
número
cuerpo
antes
gracia
familia
nadie
mantener

tierra
madre
solo
todavía
final
pesar
perder
realidad
idea
punto
morir
convertir
entender
actividad
pueblo
incluso
mediar
acabar
guerra
siguiente
arte
modo
comenzar
semana
dios
malo
distinto
suponer
(a) través (de)
aparecer
cabeza
don
mercado
necesario
segundo
estado
servir
economía
conseguir
recordar
paso
viejo
terminar
establecer

explicar
acordar
tardar
tras
puerta
amigo
igual
ocurrir
universidad
estudiar
real
temer
voz
largo
efecto
nivel
lograr
amor
muerte
sistema
comunidad
ofrecer
diferente
necesitar
producto
anterior
cambio
caer
zona
interés
caro
europeo
dinero
comunicación
sector
papel
bajar
razón
clase
intentar
internacional
blanco
educación
edad

vestir
director
fondo
principal
quizá
música
próximo
personal
estudio
luz
negro
sacar
ayer
rey
necesidad
banco
usar
movimiento
orden
acción
valor
obtener
posibilidad
imagen
humano
alcanzar
señalar
especial
tercero
leer
mostrar
elemento
programar
centro
bastante
acuerdo
tampoco
artículo
proceso
seguridad
incluir
sentar
iglesia
dirigir

ganar
futuro
organización
nadar
tarde
bajo
seguro
referir
pagar
difícil
demás
decidir
cultural
médico
desear
consejo
mayoría
respecto (a)
color
principio
chico
cortar
época
escuchar
ambos
aspecto
medir
tema
levantar
cumplir
reconocer
representar
apoyar
técnico
programa
línea
actual
junto
meter
traer
dedicar
recurso
constituir
calle

dirección
decisión
responder
interior
comprender
militar
función
miembro
población
pie
total
comprar
antiguo
investigación
capital
continuar
popular
película
proyecto
aumentar
partido
pronto
dato
francés
echar
fuerte
ocupar
texto
reír
carácter
salud
olvidar
plan
serie
correr
forzar
carta
periodo
interesar
objeto
civil
natural
producción
minuto

autor
consecuencia
opinión
capacidad
hermano
televisión
común
cine
acercar
cantidad
ciencia
cerrar
sentido
alguien
favor
mes
caminar
término
mediante
fuera
contener
allá
superior
frente
experiencia
aceptar
faltar
diverso
emplear
indicar
condición
profesor
asegurar
evitar
proponer
nacer
ciento
centrar
región
ayudar
institución
dormir
iniciar
destacar

disponer
atención
unidad
memoria
materia
profesional
observar
andar
enfermedad
trabajador
significar
animal
comisión
subir
contrario
material
ciudadano
juntar
afirmar
mar
fuerza
informar
conocimiento
central
calidad
práctico
suceder
notar
procesar
realmente
callar
demostrar
científico
causar
funcionar

2. Las 500 formas o palabras más frecuentes del español[3]

de	sin	mismo
la	son	otro
que	sobre	día
y	está	cada
el	también	hacer
en	esta	siempre
a	hay	entonces
los	sí	nada
se	entre	donde
no	ser	esa
un	era	hace
las	mi	bueno
del	dos	decir
por	había	tan
con	nos	otra
una	años	esto
es	tiene	después
lo	hasta	ella
para	desde	menos
su	te	tanto
al	eso	otros
como	fue	mundo
más	todos	aquí
o	puede	va
pero	pues	poco
me	han	mucho
le	así	usted
ha	bien	e
sus	vez	estaba
si	ni	estado
yo	sólo	sea
ya	ahora	ver
este	él	gran
porque	uno	cómo
muy	parte	aunque
todo	ese	están
cuando	tiempo	tres
qué	vida	les

[3] Según el corpus *Cumbre*, de 20 millones de palabras.

antes
gobierno
casa
sido
algo
dijo
país
sino
hombre
forma
estos
año
caso
hecho
durante
hoy
luego
he
momento
unos
general
creo
ellos
cosas
todas
además
mayor
mejor
España
trabajo
toda
días
tienen
política
tenía
contra
dice
hacia
vamos
tal
otras
veces
gente

ejemplo
lugar
tener
casi
estas
personas
parece
mujer
primera
según
mientras
mí
nosotros
poder
nunca
tu
algunos
pueden
claro
cuenta
sé
nuevo
debe
tú
medio
nacional
ante
muchos
señor
Madrid
millones
dentro
nuestro
ahí
manera
sistema
misma
fin
verdad
quien
muchas
tengo
social

tipo
cosa
punto
cualquier
primer
sociedad
agua
presidente
embargo
historia
pasado
grandes
nuestra
noche
hemos
cual
importante
dicho
cuatro
estar
fuera
allí
información
mis
siglo
ello
tenemos
acuerdo
padre
tarde
mañana
haber
unas
nueva
estamos
ciudad
nadie
sentido
todavía
estoy
situación
grupo
alguna

ley
lado
eran
van
será
frente
mano
aún
incluso
fueron
persona
problema
bajo
países
da
realidad
madre
partido
posible
problemas
número
nombre
dar
fuente
esos
mujeres
voy
ojos
cambio
desarrollo
través
cierto
gracias
cinco
guerra
habían
don
relación
obra
proceso
ir
programa
hora

mal
hizo
estados
cuerpo
tema
primero
algunas
pueblo
horas
quiere
tras
tierra
familia
modo
centro
derecho
economía
san
dios
hablar
quiero
cabeza
pasa
determinar
algún
hombres
sería
empresa
final
cuanto
arte
quién
esas
público
iba
llegar
igual
meses
dinero
amor
mil
sabe
educación

calle
falta
soy
muerte
buena
podía
último
largo
ayer
música
segundo
saber
cultura
aquel
decía
éste
mercado
partir
demás
nivel
manos
tampoco
edad
ellas
español
voz
semana
libro
propio
digo
seguridad
hijo
ningún
junto
niños
paso
pesetas
podría
universidad
dado
luz
vista
palabras

jóvenes
visto
respecto
obras
hijos
comunicación
real
puesto
haya
puerta
trata
española
unidos
mayoría
hubiera
tenido
aquella
principio
podemos
buen
bastante
idea
único
interés
niño
propia
ustedes
papel
razón
político
director
cara
nuestros
siguiente
casos
dónde
Europa
servicios
condiciones
eh
ninguna
orden
total

medios
seis
tuvo
estaban
media
campo
salud
dirección
difícil
personal
viene
siendo
pasar
proyecto
pronto
joven
haciendo
camino
existe
época
buenas
consejo
diferentes
buenos
aire
quienes
segunda
fuerza
hacen
hacía
quería
rey
derechos
empresas
interior
comunidad
dio
población
acción
medida
servicio
alguien
últimos

iglesia
económica
mediante
base
internacional
televisión
pesar
México
varios
sociales
datos
puedo
allá
palabra
última
valor
diez
relaciones
serie
especial
futuro
cine
producción
padres
fondo
anterior
favor
carácter
queda
ministro
actividad
grupos
ésta
nuevos
libertad
artículo
has
llama
alto
aquellos
salir
estudio
capital