

# El español: segunda lengua romance en la Red. Nuevo estudio de *Funredes*

Por José Antonio Millán

## La métrica de las lenguas en la Red

LA WEB PÚBLICA ES UN OBJETO DE ESTUDIO abierto a todos (por definición). En una Red dominada inicialmente por el inglés, los hablantes de otras lenguas con pujanza cultural pronto intentaron descubrir cuál era su posición (Millán, 1997).

Los primeros estudios se desdoblaron en dos aspectos, los cuantitativos y los cualitativos. Desde el primer punto de vista interesaba ver qué cantidad de páginas había en cada lengua. Cualitativamente importaba ver la presencia de determinados elementos de una cultura en las páginas web de distintas lenguas. El pionero de ambos estudios fue **Daniel Pimienta**, desde la *Fundación Redes y Desarrollo* (*Funredes*, 1996a y b), inicialmente para el español y el francés (en relación con el inglés), y que hoy en día abarca (con el apoyo de *Unión Latina* y de la *Délégation Générale à la Langue Française*) las otras len-

País	% de páginas en español	Productividad	
España	47,7%	2,38	
EUA	14,3%	0,40	Sigue muy baja la productividad a pesar del crecimiento
Argentina	10,6%	1,85	
México	7,4%	0,53	Sigue baja la productividad
Chile	3,9%	1,05	
Venezuela	1,3%	0,71	
Perú	1,0%	0,30	
Uruguay	0,8%	1,51	

guas romances y además el alemán. El catalán ha estado ausente de estas medidas, porque *Unión Latina* se ocupaba sólo de las lenguas oficiales de los estados miembros (este año acaba de adherirse Andorra) y porque (a pesar de las reiteradas invitaciones) ninguna institución catalana ha requerido hasta ahora la colaboración de *Funredes*.

<http://www.unilat.org/>

<http://www.culture.gouv.fr/culture/dglf/>

Para el estudio cuantitativo, que es el que nos ocupará, **Pimienta** utilizó inicialmente el buscador *Altavista* y un algoritmo de reconoci-

miento de lenguas (mediante palabras exclusivas de cada una) para determinar las cifras totales (*Funredes*, 1996a). Estudios posteriores como el de *Alis Technologies* (1997), con el apoyo de la *Internet Society*, utilizaron un muestreo muy pequeño de sitios.

«El francés crece más que el español y por primera vez supera a nuestra lengua»

Las medidas más fiables, que se extraían utilizando buscadores generales, tenían el problema fundamental de saber si el buscador indicaba la totalidad del espacio web, o si se centraba en el conjunto más antiguo de páginas, dando poca cuenta de las nuevas incorporaciones a la Web, y primando de esta manera las cifras de recursos en inglés (para una perspectiva global de la situación de los buscadores en aquel momento con respecto a la totalidad, véase **Millán**, 2000). El catalán (que se había iniciado en la Red muy tempranamente: **Partal**, 2004) acogió una iniciativa temprana para decir a los buscadores en qué lengua estaban sus sitios mediante etiquetas meta (**Yzaguirre**, 1997), como forma de aumentar su visibilidad.

La evolución de los estudios con una misma metodología (los de *Funredes*) a lo largo de los distintos

### Cuadro 1. Presencia de personajes: culturas anglosajonas/neolatinas

Para la anécdota, estos indicadores sociológicos:

—En ciencia no hay cambios notables: **Einstein** y **Marie Curie** es la pareja más citada.

—En letras hay muchos más cambios. **Descartes** pasa a **Víctor Hugo** y se instala en los cinco más citados, en el orden: **Shakespeare**, **Descartes**, **García Márquez** (quien se transforma en el autor vivo más citado), **Wilde** y **Tolkien**.

—En cine sólo **Depardieu** y **Almodóvar** compiten con la supremacía anglófona.

—En artes, el trío ganador es: **Dalí**, **Warhol** y **Khalo**.

—Para los personajes históricos, **Napoleón** gana 17 lugares y se coloca en cuarta posición detrás de **Washington**, **Cristo** y **Jefferson**, delante de **Lincoln**, **Franklin...** y **Bolívar**.

—Batman pasa bien encima de Satán y la Cenicienta hace su aparición entre los personajes imaginarios de cabeza. Mientras **Mafalda** sigue subiendo.

# Facilitando el acceso integrado a contenidos

## ZPORTAL

ZPORTAL facilita un entorno de búsqueda federada que lanza búsquedas a diversos recursos de información de forma simultánea, recuperando los resultados en una única pantalla.

Con ZPORTAL se ha desarrollado el portal de salud del Reino Unido, [www.library.nhs.uk](http://www.library.nhs.uk), la herramienta que personal sanitario, documentalistas, personal de administración y pacientes utilizan para la búsqueda integrada de información.

fdi

fretwell-downing  
informatics

telephone: +44 114 281 6040  
email: [enquiries@fdisolutions.com](mailto:enquiries@fdisolutions.com)  
[www.fdisolutions.com](http://www.fdisolutions.com)

Distribuidor en España:

doc6

consultores en recursos de información

Mallorca, 272, planta 3 - 08037 Barcelona  
Tel. 93 215 43 13 Fax: 93 488 36 21

Orense, 14, 5º A - 28020 Madrid  
Tel. 915 535 207 Fax: 915 346 112  
[mail@doc6.es](mailto:mail@doc6.es) - [www.doc6.es](http://www.doc6.es)

años ha conducido a refinar no sólo el uso simultáneo de varios buscadores y los procedimientos de cálculo, sino también las medidas absolutas, poniéndolas en relación con el número de hablantes y también con la cantidad de personas con conexión a la Red (productividad de un espacio web). Estos parámetros se pueden utilizar también para cada país, lo que proporciona una visión muy rica de comunidades geográficamente dispersas como la hispanohablante.

### «Nicaragua y Cuba: las cifras más altas de productividad en español»

#### El español y las otras lenguas

Según el último estudio (Pimienta, 2005), la presencia relativa del inglés no ha dejado de bajar desde 1998 (cuando tenía el 75% del total) hasta el 45% en la actualidad. El resto de las lenguas estarían así:

- Inglés: 45,0%
- Alemán: 7,0%
- Francés: 5,0%
- Español: 4,5%
- Italiano: 3,0%
- Portugués: 2,0%
- Rumano: 0,1%
- Resto: 35,0%

La novedad respecto a las cifras anteriores (2003) es que el español pierde el primer lugar de las lenguas romances frente al francés.

### «EUA produce más páginas en español que Argentina o México»

Éste es el resumen que aporta *Funredes* de los principales datos detectados:

El francés crece más que el español y el portugués, los cuales confirman una parada en el fuerte crecimiento de los años pasados. El despertar tardío de Francia (y Bélgica) parece explicar este auge, además de una política voluntarista de la francofonía para la producción de contenidos. Los nuevos internautas de los países con fuerte penetración de internet ya son más consumidores que productores. Eso daría aún más razón a una política de apoyo a la producción de contenidos. No aparece ninguna reducción de la brecha en África. Recordemos que la productividad es la relación entre el número de internautas y la cantidad de páginas producidas.

### «García Márquez, Almodóvar, Dalí y Mafalda, las presencias hispánicas más destacadas en cultura»

Es de notar el despertar de la producción de contenidos en español procedente de EUA y la alza notable de Chile. España sigue estable y con una alta cifra de productividad, perteneciendo las cifras más altas a Nicaragua y Cuba. Una curiosidad: el dominio .pr de Puerto Rico, aunque poco utilizado, parece concentrar la producción de páginas en español de este estado asociado de EUA.

#### Conclusión

Con casi una década de datos tal vez estemos en situación de aventurar evoluciones cuantitativas para las lenguas que se incorporan a

la Red: inicios de crecimientos rápidos, seguidos de mesetas de estabilización. El inglés se estabilizó hace años, el español habría alcanzado ya una de esas mesetas y el francés estaría llegando a la suya. Pero para éstas, y tras muchas cosas más, es mejor leer el último estudio de *Funredes*. Y en cuanto a los estudios cualitativos, véase el cuadro adjunto.

#### Referencias

*Alis Technologies*. *Palmarès des langues de la toile*, 1997.  
<http://alis.isoc.org/palmares.html>

*Funredes*. *El primer estudio de las lenguas y la internet*, 1, 1996a.  
<http://funredes.org/LC/espanol/L1.html>

*Funredes*. *El primer estudio de la cultura y la internet*, 1996b.  
<http://funredes.org/LC/espanol/L1.html>

*Funredes*. *Observatorio de las lenguas y culturas*, octubre 2005.  
<http://funredes.org/LC/espanol>

Millán, J. A. «El español en las redes globales». En: *Actas del I Congreso internacional de la lengua española*, 1997.  
<http://cvc.cervantes.es/obref/congresos/zacateca/tecnologias/ponencias/millan.htm>

Millán, J. A. «El libro de mil millones de páginas. La ecología lingüística de la Web». En: *Revista de libros*, 2000, n. 45. Versión ampliada en la web:  
<http://jamillan.com/ecoling.htm>

Partal, V. *El català a la xarxa: història i raons d'un cas d'èxit*, 2004.  
<http://www.softcatala.org/articles/article39.htm>

Pimienta, D. Comunicación personal. El último estudio aparecerá también en la Web, 2005.  
<http://funredes.org/LC/espanol/ultimas.html>

Yzaguirre, Ll. *Indexem en català*, 1997.  
<http://www.iula.upf.es/altres/pub/indexem/indexem.htm>

Una primera versión de este artículo se publicó el 21 de septiembre de 2005 en Thinkipi.  
<http://thinkipi.net/repositorio/la-metrica-de-las-lenguas-en-internet/>

Primera publicación en *Libros y bitios*: 11 de octubre del 2005  
<http://jamillan.com/espse.htm>

**Leer EPI es como asistir a un curso de formación continua pero con el horario acomodado a tus necesidades.**