

Revisiting the Spanish blogosphere

J. J. Merelo, U. Granada (Spain),
J. L. Orihuela, U. Navarra (Spain),
Ruiz (Six Apart, Spain),
F. Tricas (U. Zaragoza, Spain)

Abstract

Since our first studies the size of the Spanish blogosphere has almost trebled. New waves of bloggers, and blog hosting sites have appeared, with some of them becoming quite popular; obviously, some previously well-known sites have gone into oblivion.

Depending on how you measure, it can be said now that the Spanish blogosphere follows a power law on the sites ranked by number of incoming links, but the fit is still far from perfect.

Popular topics have also changed and, in some way, its range has enriched. Among popular topics we can point to the blogosphere itself, but also politics, free (as in free speech) software, or science are also incredibly popular. In particular, some very highly focused topical weblogs have appeared and are gaining respect from the rest of bloggers. Several meta-blogging sites have also appeared, in an attempt to become the center of the blogosphere: sites that provide lists (even feeds) of updated weblogs, directories and collective sites.

1 Introduction

We published one of our first papers on the Spanish blogosphere a year ago at BlogTalk [Tricas et al., 2003]. Since then, it has evolved to become something a bit different.

For our purposes we can say that the Spanish blogosphere can be defined as the set of blogs (or, sometimes, blog-looking web pages) that are written in Spanish (in any part of the world) or in any other of the official languages in Spain (Catalan, Basque, Galician). We also include blogs written in other languages, if they are written by people living in countries that naturally fit in the Spanish blogosphere area of influence.

Our project started at the beginning of the Summer of 2002, and it is in a 'beta' stage. All data in this study have been taken from the Blogómetro, a suite of tools whose main visible aspect is its blog (<http://blogometro.blogalia.com>), hosted in Blogalia (<http://www.blogalia.com/>). There, a list of fresh links taken from our list of blogs (ranked by the number of sites pointing to them) is published daily. The Blogómetro is an opensource collaborative project, offering our research to the community. It is open to the participation of interested people. In this sense, not only its source code is available at the project page (<http://sourceforge.net/projects/blogometro>), but also the list of sites scanned daily.

Since our first studies the size of the Spanish blogosphere has almost trebled. Last year we were studying something in between 3000 or 4000 blogs, at the moment of this writing (August-September, 2004) our list contains more than 20000. We have partial evidence that the number of readers has also increased. New waves of bloggers, and blog hosting sites have appeared, with some of them becoming quite popular. In particular, a very successful

initiative has been Blogia (<http://blogia.com/>), a hosting place started at the end of 2003, and that hosts more than 8000 weblogs now.

We will see that the Spanish blogosphere is approaching a power law on the sites ranked by number of incoming links, but that the fit is still not perfect. It is not clear if that means that the Spanish blogosphere is yet mature, but it shows that, at least, some authorities are more clearly present than they were last year, and also that newcomers are filling some existing gaps.

Popular topics have also changed and, in some way, its range has enriched. One of the most popular topic is still the blogosphere itself, but politics is also incredibly popular, followed, surprisingly, by science. In particular, some very highly focused topical weblogs have appeared and are gaining respect from the rest of bloggers. Several meta-blogging sites have also appeared, in an attempt to become the center of the blogosphere: sites that provide lists (even feeds) of updated weblogs, directories and collective sites, giving people the opportunity to gain awareness of community and also to discover sites that usually do not appear in general (English-speaking) directories, and topics not covered by the global blogosphere.

The organization of the rest of the paper is as follows. In Section 2 we are going to present a review of main trends at the Spanish speaking blogosphere between BlogTalks (May 2003 to July 2004). Section 3 is devoted to the sites that host what we call Spanish blogosphere. Section 4 will show some numerical data. Finally Section 5 will show some conclusions.

2 From BlogTalk to BlogTalk

Here we are going to sketch the main trends happened in the Spanish blogosphere during the last year. Interested readers can visit [Orihuela, 2004a, Orihuela, 2004b, Orihuela, 2004c] for a list of links related to these trends.

2.1 Trend 1: Weblogs become visible to the mass media, and eblogs sections started to appear

Most of the main newspapers, technical magazines, web portals and some radio and television shows devoted space to cover the main issues related to blogging for general audiences. It is fair to say that during this period, mass media discovered the weblogs. In most of the cases the quality of the coverage is very poor and the focus is introductory, due to the reporters lack of direct knowledge about weblogs. Interviewing bloggers is an extended format to deal with these issues. Technical magazines tend to offer a better approach with tutorials about the main blogging tools and in depth coverage including weblog reviews, lists of blogs by genre and interviews. Fixed regular sections about weblogs started at Terra portal and at the weekly technology supplement of El País, while El Mundo online edition started a sort of blog dealing with technology news. Every single news and coverage

related to weblogs in the mass media is referred by the blogosphere, mainly by the meta-bloggers but also by all of those quoted or mentioned in the news. Posts about this topics used to be very critical with the media in most of the cases, focusing in mistakes and misunderstandings, but revealing, at the same time, a strong dependence of established media recognition between the newest bloggers.

2.2 Trend 2: Weblogs become a topic research in the field of Media Studies and also an issue in most of the main academic conferences

The increasing impact of weblogs in the broad field of Media and Journalism Studies is shown by the number of references in academic publications: chapters of books and articles in journals and magazines. The most important academic conferences devoted to media and journalism included the topic of weblogs in the program and the blogosphere is studied in papers and discussed in round tables. Scholars showed interest in the nature of blogging as activity and as media, in the relations and impacts of blogging on journalism, and mainly in the study of relations, hierarchies and communities inside the blogosphere. Even when weblogs are not yet very well known or experienced by most of the scholars, there are a number of evangelists in almost every school of public communication.

2.3 Trend 3: Weblogs slowly start to become a research, learning and teaching tool in the academic world with the impulse of prime movers

Besides the Media Studies field, other areas of the academic world started to explore weblogs as a tool for research, learning and teaching: Computer Sciences, Education Sciences, Social Sciences, Librarians and Documentalists, just to mention a few. Prime movers or evangelists are pushing to include weblogs in workshops and seminars, and to explore the potential of the medium as a teaching tool to publish repertories, references and materials of interest for the students. In some cases students are blogging as course assignment.

2.4 Trend 4: A strong sense of community arises in the blogosphere with the impulse of sites of reference and tools for bloggers, and meta-blogging is reinforced.

New weblog portals, directories of blogs by genre, lists of recently updated blogs, meta-blogs and a number of new hosting services are contributing strongly in the process of expansion of the Spanish speaking blogosphere, and at the same time are operating as global references for the bloggers community. Traditions such as the weblog awards and the invisible friend game (second year both), a PDF weblog magazine, IRC channels, an Internet Radio Station and a number of local, national and regional directories are the evident signs of a dynamic community. The self reference discourse about the blogosphere and the meta-blogging initiatives used to generate debate and criticism. Studies, metrics, popularity, influence, awards and rankings happens to generate division between those interested in understand and promote the blogosphere, and those who feel in the shadow or not well represented in the figures. In any case, the Spanish speaking blogosphere, small and diverse is building its own identity and is becoming to be conscious about it.

3 Main blogging sites

We have studied main hosting sites in our list of blogs. If you compare with our results from last year you will observe that some successful sites have appeared and one can guess that language is a barrier for many Spanish-speaking people. It could be also possible that media attention also helped to attract people that in other case would have not known about blogging.

The list has been reduced in order to show only hostings with 50 or more weblogs. The first column shows the domain of the hosting. The second one shows the total number of weblogs as present in our list on August, 1st.

We can see that bloggers not only host their sites on specialized sites, but they also host their blogs at general hosting sites like iespana.es, geocities.com, cjb.net.

Furthermore, let us note that since our last visit to the Spanish blogosphere [Tricas and Merelo-Guervós, 2004], new hosting sites have appeared as universia.net (focused on the academic community) and blogs.ya.com that are helping to introduce blogging in new and different audiences. Moreover, the second one, sponsored by one of the main contenders in the Spanish WWW, has introduced also blogs in chueca.com which is a gay portal.

Site	Number of blogs
blogia.com	8166
zonalibre.org	2376
ya.com	2195
blogspot.com	1744
barrapunto.com	1047
lacomunitat.net	342
blogalia.com	272
diariogratis.com	116
bloxus.com	115
escomposlinux.org	100
barcelonablogs.com	92
bitacoras.com	87
iespana.es	77
geocities.com	77
motime.com	76
chueca.com	70
lamevaweb.info	68
deusto.es	68
cjb.net	59
pitas.com	54
...	
Total	18125

4 Towards the PowerLaw?

During the period of the study, more than 1700000 links were observed (more than 281000 in our last study), with around 450000 live links in any given day (links in the front page of weblogs studied).

All of this numbers have been checked until August 2004, first, so they are a bit different of the ones provided in the presentation at the conference. Until now were only aware of similar studies for the global blogosphere, but in this edition of BlogTalk (see proceedings) other authors are starting to do similar work about other local blogospheres. There is also another work on the Dutch blogosphere recently published [Schaap, 2004], but it is a qualitative approach, instead of quantitative.

We have presented results about this in [Merelo et al., 2003, Tricas et al., 2003, Tricas and Merelo-Guervós, 2004]. In this paper we provide an update on the evolution that can be seen in Figure 1. It is shown in a log-log scale, and we can see there that the fit is still not good, since there is a set of under-linked sites in the high part of the graph, and a big set of over-linked (always according to an hypothetic powerlaw) in the middle part.

The root mean square (RMS) of the residuals is of 4.94 that is neither better nor worse than the last one obtained for [Tricas and Merelo-Guervós, 2004]. Anyway it is evolving to a better fit with time (or, at least, is is not going worse), so, perhaps, it will be better in the future.

The most linked to sites are still banners and selfpromotion links of popular sites (Blogia, Blogger, BarraPunto, MT, CC, ...). If we do not consider these links, the first real links is <http://www.rae.es/>, with 159 links. This shows some interest from Spanish bloggers about language.

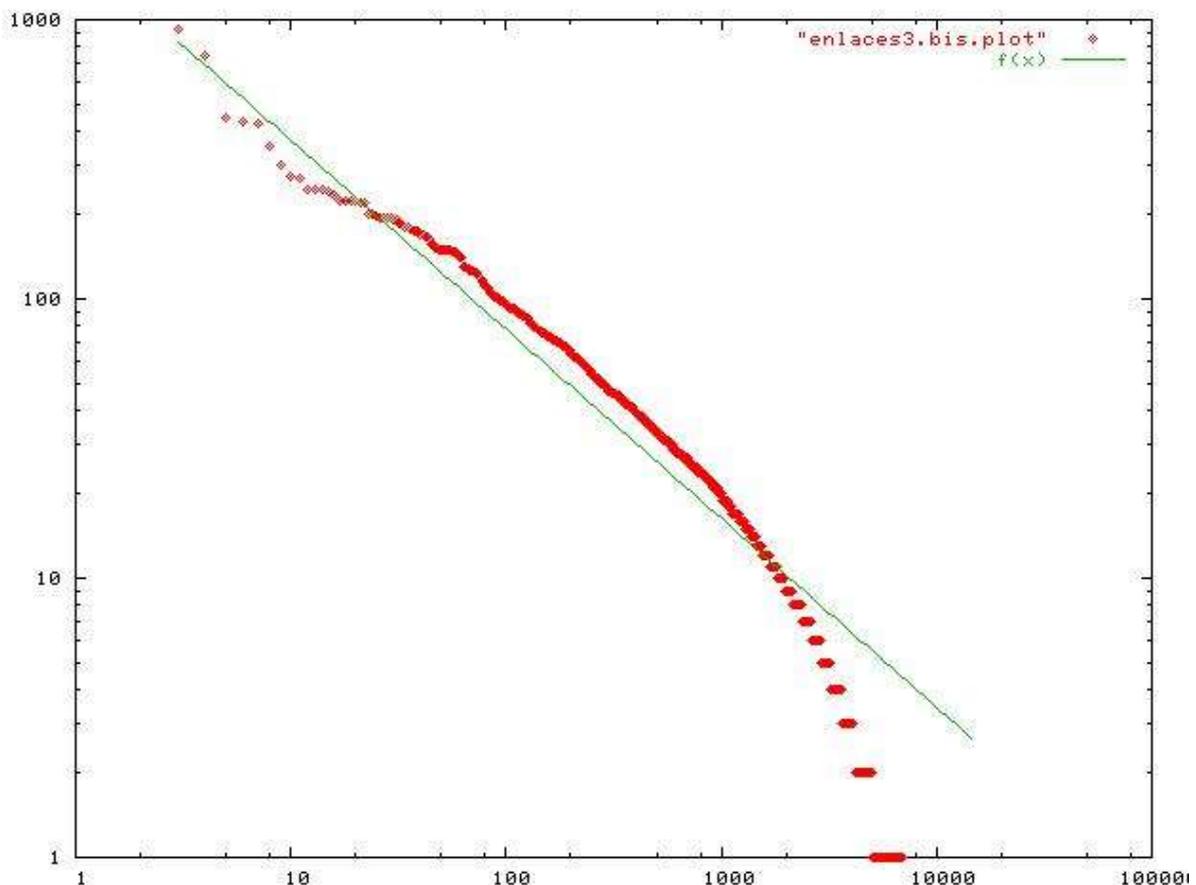


Figure 1: Fit a powerlaw of links in the Spanish blogosphere

Second is last year's first real link, <http://www.librodenotas.com/mt/prestige.html> (128 links), a (quite critical) page on the Prestige wreck, which was part of a campaign to Google

bomb the word prestige. Last year it had 75 so it is still growing (last links are from May-June of 2004).

The third one is this year's Google-bomb, <http://www.sgae.es/>, with 125 links. This is the webpage of the entity who manages author's rights in Spain, and has been criticized because of the imposition of a tax to CDROMs to compensate author's for their reduced income due to illegal copies. People proposed to link them with the word 'ladrones' (thieves). It is worth noting that this is not a special link prepared for this action, so other uses for the link can appear in the blogosphere (the Google bombing would have helped to increase the total number of links but the link appears in our database even before the date when someone proposed the bombing, and also after this date, but with other purposes).

Other successful sites are digital journals. For example, El Mundo (<http://www.elmundo.es/>, 350 links), El País (<http://www.elpais.es/>, 346), Periodista Digital (<http://www.periodistadigital.com/>, 174), to name the first three ones. But, to our surprise, there are bloggers that have even more links than these sites. For example, Mini-D (<http://www.minid.net/>, 471), Las cinco del viernes (<http://lascincodelviernes.blogspot.com/>, 344), Blogpocket (<http://www.blogpocket.com/>, 319) eCuaderno (<http://www.ecuaderno.com/>, 225 -this is José Luis Orihuela main one-),

Acknowledgements

This paper has been funded in part by project TIC2003-09481-C04-04, of the Spanish ministry of science and technology, and a project awarded the Quality and Innovation department of the University of Granada. Fernando Tricas is with the Group of Discrete Event Systems Engineering (GISED), and his work is partially supported by TIC2001-1819, financed by the Spanish Ministerio de Ciencia y Tecnología.

5 Conclusions and Future work

The Spanish blogosphere is quickly evolving. New hosting sites appear and, with them, new bloggers that are providing diversity but also a more strong sense of community. Its size is growing at a good pace and it seems that the fit to an hypothetical powerlaw is better with time. It is raising the attention in several different environments (media, academy, ...) and our feeling is that this is helping to attract more people.

References

[Merelo et al., 2003] Merelo, J. J., Tricas, F., and Ruiz, V. (2003). Measuring the spanish blogosphere. In Proceedings of Towards New Media Paradigms (COST A20 Conference).

[Orihuela, 2004a] Orihuela, J. L. (2004a). Datos para un anuario de la blogosfera hispana.. Available from <http://www.ecuaderno.com/archives/000344.php>, downloaded September, 3rd 2004.

[Orihuela, 2004b] Orihuela, J. L. (2004b). La blogosfera hispana en conferencias académicas. La blogosfera hispana en conferencias académicas. Available from <http://www.ecuaderno.com/archives/000345.php>, downloaded September, 3rd 2004.

[Orihuela, 2004c] Orihuela, J. L. (2004c). La blogosfera hispana en publicaciones. Available from <http://www.ecuaderno.com/archives/000346.php> , downloaded September, 3rd 2004.

[Schaap, 2004] Schaap, F. (2004). Links, lives, logs: Presentation in the dutch blogosphere. Into the Blogosphere: Rhetoric, Community, and Culture of Weblogs, Available at http://blog.lib.umn.edu/blogosphere/links_lives_logs.html, downloaded July 2004.

[Tricas and Merelo-Guervós, 2004] Tricas, F. and Merelo-Guervós, J. J. (2004). The Spanish-speaking blogosphere: towards the powerlaw? In Proceedings of Web Based Communities 2004, pages 430433, Lisboa, Portugal.

[Tricas et al., 2003] Tricas, F., Ruiz, V., and Merelo, J. J. (2003). Do we live in an small world? measuring the spanish-speaking blogosphere. In Burg, T. N., editor, BlogTalks, pages 158.173 ISBN 3901806008.