

AN ANALYSIS OF THE LINGUISTIC DIVERSITY OF CYBERSPACE

Cezar VASILESCU

Professor dr. habil., eng.

Regional Department of Defense Resources Management Studies, Brasov, Romania

The article proposes an analysis regarding the relationship between Internet usage and language. Moreover, it highlights the impact the latter have on the human interactions depicted by future knowledge societies within the framework of the Information Age. This endeavor explores from a linguistic perspective how cyber users' native language affects their Internet usage patterns. Hence, its final goal is to determine whether the Internet is expected to remain overbalanced in English usage. In this respect, the relationship between web users' native language and the language content of the Internet websites they access is also examined based on statistic data.

Key words: *language analysis, websites, Internet users, statistics.*

1. INTRODUCTION

The focus of this article is to present the relationship between Internet usage and language and to highlight the impact it is creating on the human interactions among future knowledge societies, within the framework of the Information Age.

A perception generally accepted by most of the people is that during the last 20 years, English language becomes the *de facto* standard for Internet communication (both for business and academic purposes) and attained the status of a so called "global language". The Internet is for sure an effective instrument for circulating English all over the world and there is much more Internet content available in English than in any other language.

2. THE RELATIONSHIP BETWEEN WEB USERS' NATIVE LANGUAGE AND THE LANGUAGE CONTENT OF THE INTERNET WEBSITES

In this paragraph we will examine how native language of cyber users

affects their Internet usage patterns from a linguistic perspective. The final goal is to determine whether the Internet is expected to remain overbalanced in English.

The English language advantage in terms of Internet content was determined by the fact that early web users were native English speakers. However, this advantage is very likely to diminish in the future, because of the faster growing usage of Internet resources among non-native English speakers.

A key determinant of whether the web content will remain significantly in English is if non-English speakers will continue to use English language websites [1]. In such a case, the existing web content providers would have a strong (economically driven) interest to provide their content in English, and almost no enthusiasm to create non-English versions of their sites.

The usage of language on the Internet can be viewed as a three step process:

- *In the short-term* - individuals decide which websites to visit (based on their interests, language skills and on the available relevant

offerings in other languages), which determines actual Internet usage;

- *In the medium term* – websites' owners decide which language to use for their site, and whether to offer it in multiple languages, based on the amount of traffic a site expects to attract in one language or another;
- *Over the long term* - individuals make decisions about which languages to learn, based partially on the interest to access specific content, as well as the desire to communicate directly with speakers of other languages.

Let us make a translation of the well known network economic effect (the value that consumers place on a particular product increases as the total number of consumers who use identical or compatible goods increases) to the topic of language used on the Internet. The principle stands correct and we can observe a clear network effects application here: the value of the network depends on the total number of subscribers who have access to the network. The more widely a language is used, learning it as a second language is more valuable.

The value of speaking English increases as the number and content (books, journals, movies) of English language websites increases. This will lead to an increase in the number of non-English speakers learning English in order to have access to English language websites, since individuals who speak English will have more websites to use.

Consequently this will lead to an increase in the number of English language websites [2]. It might be conceivably to experiment in the future the *snowball effect* [3]: once a system has gained an initial lead, there is "a natural tendency towards de-facto standardization", which resembles the organizational behavior related concept of *groupthink* [4]. For this reason, it is possible that a large portion of Internet

to remain in English (as a dominant language on the web), even if studies shown that there are more native Chinese and Spanish speakers than there are native English speakers.

The statistics done over the years are particularly interesting when it comes to estimate the number of Internet users by language. Because of the importance of the subject, and due to the lack of sources, in the paragraphs below we are presenting three statistical data sets from 2011, 2013 and 2015, which might help us finding trends about the hierarchy of the top ten languages on the Internet.

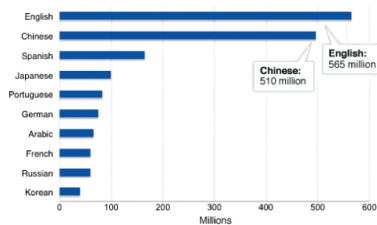


Fig. no. 1. Top ten languages on the Internet as of May 2011 [5].

This United Nations Broadband Commission's study suggests that Chinese could soon replace English as the most commonly spoken language on the Internet. The graph shows that English remained the world's most widely spoken language on the web, with 565 million people (27 %) of the global online population using it as their main language, while Chinese web users continue to grow rapidly, reaching 510 million people (24%) of the world's total. The study predicted that if current growth rates continue, Chinese users may outnumber English-speaking users by 2015, result invalidated by a more recent study [6].

The third study was also made by Internet World Stats, which presents its latest 2015 estimates for Internet Users by Language. The figure below publishes details for the top ten languages in the Internet.

Table 1. Estimates of the number of Internet users by language as of December 31, 2013 [7]

Rank	Language	Internet users	Percentage
1	English	800,625,314	28.6%
2	Chinese	649,375,491	23.2%
3	Spanish	222,406,379	7.9%
4	Arabic	135,610,819	4.8%
5	Portuguese	121,779,703	4.3%
6	Japanese	109,626,672	3.9%
7	Russian	100,700,000	3%
8	German	81,139,942	2.9%
9	French	78,891,813	2.8%
10	Malay	75,459,025	2%
11 -36	Others	440,087,029	15.7%
Total		2.81 Billion	100%

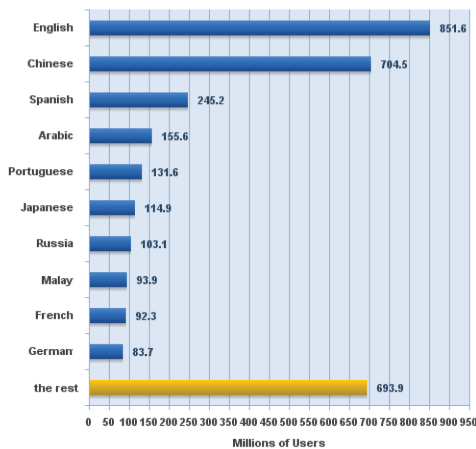


Fig. no. 2. Internet world users by language - top 10 languages [6].

The large number of English language websites might determine non English speakers to learn English so that they can access them.

Another conventional wisdom suggests that English as a second language will be a working language that will enable collaboration without threatening the survival of native tongues. The idea is that the non-native children, who learn English, will work and think in their own language, but English will allow them to communicate and interact, being a bridge language of economic opportunity.

To analyze the future of language in a connected world we have to discover precisely enough what percent of the Internet’s content is written in English.

Early studies [8] estimate that 80 percent of the World Wide Web’s content in 1997 was in English, while another study from 2003 was estimating that only 72 percent of online content was in English. Both percentages mislead researchers to suggest that English had a head start other languages would find difficult to overcome. The reason is that fact it is incredibly difficult to generate a believable estimate of language diversity online, because it is hard to choose a credible and representative test sample of websites.

A study regarding the usage of content languages for websites was made by w3techs.com [9]. It includes only the top 10 million websites (top 1 million before June 2013) based on the website popularity rankings provided by Alexa (an Amazon.com company) using a 3 months average ranking. In this case, Alexa rankings serve the purpose of providing a representative sample of very well established sites.

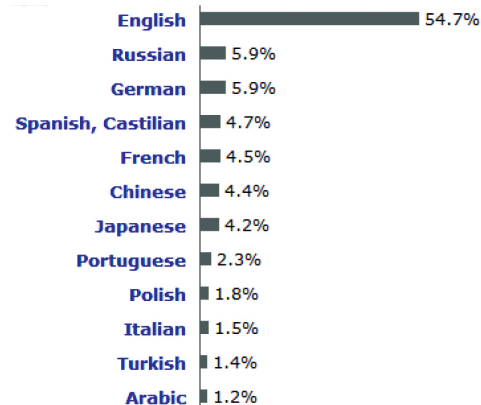


Fig. no. 3. Usage of content language for websites [9].

The study also revealed that the Russian Internet domain .ru became the most popular country-code top level domain in September 2011 and has been growing ever since. However, on top 1000 sites, Russian language is less used (it is only language no. 6 behind English, Chinese, French, German and Japanese).

The trend in the last few years showed that it seems very likely that Russian will stay at rank no. 2 for the moment, while Chinese would be the logical next candidate for the position.

Estimating the number of speakers of the world's languages is an increasingly complex task, particularly with the increased trend present in many countries to teach English in public schools. In the entire world many people are bilingual or multilingual, and it is estimated to exist a total of 750 million speakers of English (first or native language, plus second language) [10], while India and China combined have over 500 million users of English [11].

Some researchers warn that that search engines may no longer offer a representative sample of content online, because social networks (such as Facebook and Twitter) are difficult to be fully indexed. Also search engines now index less than 30 percent of the visible web, and the indexed subset skews toward English language sites, often because those sites are the most profitable places to sell advertising [7].

Statistics about Internet usage show much faster growth in countries where English is not the dominant language. In 1996, more than 80 % of Internet users were native English speakers, but in 2010 that percentage dropped to 27.3 %.

While the number of English-speaking Internet users has almost triples since 2000, twelve times as many people in China use the Internet now comparing with 1996, regularly using Renren (a social media platform equivalent to Facebook) or Sina Weibo (a site similar with Twitter). Growth is even more impressive in the Arabic speaking world, where twenty-five times as many people are online in comparison with 1996.

In 2015, the difference in growth is even more evident (the statistics covered the time interval from 2000 to 2015) [6]. The biggest percentage belongs to Arabic (6091.9 %), followed by Russian (3227.3

%), Chinese (2080.9 %) and Portuguese (1637.3 %) mainly due to South America's countries. English obtains a decent growth rate of 505.0 %.

Also the number of non-English pages is rapidly expanding (statistics covered from 2001 to 2011): the use of English online increased by around 281%, a lower rate of growth than that of Spanish (743%), Chinese (1,277%), Russian (1,826%) or Arabic (2,501%) over the same period [12].

3. HOW LITTLE OF THE WORLD WE SEE THROUGH OUR BROWSERS

"While it's easier than ever to share information and perspectives from different parts of the world, we may now often encounter a narrower picture of the world than in less connected days". [13]

The existence of cyberspace offers the premises for a better connected world and better outcomes for people across the globe, which elevates the importance of living dual lives, as citizens of nations and citizens of the world. The Internet Age we are living is a time of connection. The ubiquitous technology embedded in our lives often leads to the following assumption: while the number of people online grows, it will inevitably leads to a smaller, more connected world, in which we will think, understand, communicate and share more with people from other cultures.

The reality contradicts this assumption, because the technological ability to communicate with someone does not automatically guarantee better human interaction or an increased exchange of information. While it is easier than ever to access information from all over the world, the lenses through which we view the social environment have become narrower. People are less open to cross-cultural experiences that may stimulate new learning and understanding.

The world is complex and interconnected, and the evolution of our communications system, from a broadcast model to a networked one, has added a new dimension to the mix [13]. We search for information or find new things through people we know, and since these people tend to resemble ourselves, a lot of things within the global flow of information that happen in the world that did not catch our attention.

There is a difference between what networked technology *could do* (in principle), and what it *actually does*. “*You can’t always get what you need*” [14], but you usually get what you want - and it’s not necessarily good a good thing for you.

To ensure that electronic connection bring people closer and increase the human connection, everybody need to experiment in all digital life’s areas: online language, personal connection, and discovery [15]. Regarding the sphere of online language, it worth mentioned the ongoing development of automated translation, which could allow us to smoothly follow conversations unfolding in languages unknown for us.

4. CONCLUSIONS

We can expect that the English language status of usage in the Internet to remain unchanged during the forthcoming period of social and economic global change. English is widely scrutinized as becoming the “global language”, due to the fact that the world in which it is used is in continuous social, economic and demographic transition, which implies connection and smoother communication.

There is a growing belief that the future will be a bilingual one, in which an increasing proportion of the world’s population will be fluent speakers of more than one language (probably English, as a second, additional or foreign language) [10].

Taking in consideration those three kinds of English speakers (as a first, as a second or additional language and as a foreign language), the future of linguistic communication on the Internet will be not determined by native ones, but by the latest categories of people.

It will be interesting to observe in the next 25 years if they will embrace English or they will simply use it as a vehicular language and complementary what role the language will play in their lives. To conclude, even if English is unlikely to be displaced as the world’s most important language, the future is less certain and even more complex than we could possibly imagine.

REFERENCES

- [1] GANDAL, N. - Native Language and Internet Usage, December 2003.
- [2] CHURCH and GANDAL - Network Effects, Software Provision and Standardization, Journal of Industrial Economics, XL: 85-104, 1992.
- [3] KATZ, M., and SHAPIRO, C. - Systems Competition and Network Effects, Journal of Economic Perspectives, 8: 93-115, 1994.
- [4] ZLATE, M. – Tratat de psihologie organizațional-managerială, vol II, editura Polirom, 2007.
- [5] The State of Broadband 2012: Achieving Digital Inclusion for All. A report by the Broadband Commission, September 2012.
- [6] Internet World Users by Language - Top 10 Languages, Internet World Stats, Miniwatts Marketing Group, 30 June 2015, retrieved from <http://www.internetworldstats.com/stats7.htm>.
- [7] Number of Internet Users by Language, Internet World Stats, Miniwatts Marketing Group, 31 May 2011
- [8] NUNBERG, G. - Languages in the Wired World. In The Politics of Language and the Building of Modern Nations, Institut d’Etudes Politiques, Paris, 1998, <http://www.parc.xerox.com/istl/members/nunberg/WebPaper.html>
- [9] http://w3techs.com/technologies/overview/content_language/all.

[10] GRADDOL, D. - The future of English – A guide to forecasting the popularity of the English language in the 21st century, The British Council, 2000

[11] KACHRU, Braj B. - Asian Englishes: Beyond the Canon, Hong Kong University Press, 2005

[12] ROTARU, A. - The foreign language Internet is good for business, <http://www.scottmclay.co.uk/foreign-language-internet-good-business>, archived by the Wyback Machine website.

[13] ZUCKERMAN, E. - Rewire: Digital Cosmopolitans in the Age of Connection, W. W. Norton & Company; 1 edition (June 17, 2013).

[14] You can't always get what you want, song by the Rolling Stones, album Let It Bleed, 1969

[15] TAYLOR, A. - A Small World After All, Bookforum, June/July/Aug 2013, retrieved from http://www.bookforum.com/inprint/020_02/11685.