

The Effect of the Linguistic Landscape of Today's Science on the Performance Indicators of Researchers from a Latin American Country: A Trend for the Region?

Sonia Vasconcelos¹, Martha Sorenson¹, Pablo Batista², Maurício Sant' Ana³ and Jacqueline Leta¹

¹ *sonia@peq.coppe.ufrj.br, sorenson@bioqmed.ufrj.br, jleta@bioqmed.ufrj.br*

Federal University of Rio de Janeiro, Av. Brigadeiro Trompowsky s/ nº, Prédio do CCS, Bloco B – sala 39, CEP 21941-590, Rio de Janeiro (Brazil)

² *batista@cbpf.br*

Brazilian Center for Physics Research, Rua Dr. Xavier Sigaud, 150, CEP: 22290-180, Rio de Janeiro (Brazil)

³ *mauricio.correiasantana@gmail.com*

National Agency of Supplementary Health (Brazil)

Abstract

The scientific databases that are most used nowadays confirm the prevalence of English as the language of today's science. The effect of this linguistic landscape, however, has only recently been the focus of scientometric studies. Here we correlate language competence and research performance, focusing on the writing skills of Brazilian researchers and their publication output in English-, Portuguese-, Spanish-, French-, and German-language journals. We also compare performance indicators such as number of papers (*Nps*) and citations (*Ncs*) for English and French writing skills. A significant correlation was found between English proficiency and publication output for English-language journals, but not Portuguese-, Spanish-, French- or German-language journals. Additionally, there was a correlation of writing skills in English (but not in French) with *Nps* and *Ncs* for publications in the Web of Science; higher *Nps* and *Ncs* were found for researchers with "good writing skills" in English. Our results suggest that a linguistic factor underlies the research performance of Brazilian researchers. Accordingly, other Latin American countries should assess the effect of this linguistic issue on their research communities. We propose to create a network of scientometricians, linguists and educators in Latin America to evaluate whether a similar trend prevails in other countries of the region.

Introduction

The scientific databases organized by Thomson Reuters and Pubmed show the prevalence of English as the language of science. For both of these databases, more than 90% of the content is in English (Kotzin, 2005; Piccoli & Procionoy, 2007), an upward trend that is likely to continue (Sousa Escandón, Gonzalez & Fernández, 2000). However, the effect of this linguistic landscape on research productivity has only recently been the focus of academic research (Man et al, 2004; Vasconcelos, Sorenson & Leta, 2005; Freeman & Robbins, 2006; Benfield & Feak, 2006; Vasconcelos, Sorenson & Leta, 2007; Ammon, 2008). Current research shows that English proficiency plays a significant role in the publication output for some European and Latin American countries (Man et al, 2004; Vasconcelos et al, 2008; Bauwens, Mion & Thisse, 2008). In Brazil, recent data show that for publications in English-language international journals (2001-2004), a strong correlation with researchers' writing abilities in English can be identified (Vasconcelos et al, 2009).

For publications in the Web of Science (1945-2005), Brazilian researchers with better developed skills in written English show better performance indicators; more papers, more citations and a higher h-index (Vasconcelos et al, 2008).

Here we show data on the correlation between Brazilian researchers' English proficiency in writing and their publication output not only in English but also in Portuguese-, Spanish-,

French- and German-language journals. We also compare the relationship of Brazilian authors' writing skills in English and in French to their number of papers (*Nps*) and number of citations (*Ncs*), based on publications in the Web of Science. Our aim is to assess the extent to which a linguistic capital associated with English may impact research productivity (measured by *Nps*), and visibility (measured by *Ncs*), of non-native-English speaking (NNES) researchers from Latin America. We take Brazil as an example in the region. This linguistic capital is taken here in the Bourdieuan (1991) sense, meaning that there is a "linguistic market [that] creates the conditions for an objective competition in and through which the legitimate competence can function as linguistic capital, producing a *profit of distinction* on the occasion of each social exchange." According to Bourdieu, this legitimate competence distinguishes those who own it. Concerning the scientific community, we share the idea that researchers who lack "the legitimate competence are *de facto* excluded from the social domains in which this competence is required". In today's science this legitimate competence in English is fundamental for the social exchanges that take place in the research community. According to our results, the linguistic competence of Brazilian researchers measured by their writing skills has a significant role in their visibility in the "social domains" of science, in which getting published and being cited are of particular importance. As stated by Van Raan (2004), "communication, i.e., exchange of research results, is a crucial aspect of the scientific endeavour. Publications are not the only, but certainly very important elements, in this process of knowledge exchange". In our analysis, those with better developed writing skills are those with marked roles in this exchange, measured by their higher *Nps* and *Ncs*. Here we discuss these findings and suggest the creation of a network of scholars working in scientometrics, linguistics and education to assess the effect of this linguistic factor on the performance of researchers from other Latin American countries.

Methodological Approach

Using a database compiled in 2005 by the Brazilian National Research Council (CNPq), with information on the academic profile and linguistic abilities of 51,223 Brazilian researchers, we collected information on the writing competence of these researchers in English. As described elsewhere (Vasconcelos, Sorenson & Leta, 2007), we identified their writing skills, which were based on their self-evaluation. Among these 51,223 researchers, 44.4% classify their writing skills as good, 35.2% consider their writing skills as reasonable and 13.0% admit to poor writing skills. We correlated these data with publications in English-language international journals for a 4-year time frame (Figure 1), including all of those mentioned in these researchers' *Curriculum Vitae* from 2001 to 2004. The correlation between English proficiency in writing and publication output in international journals was tested for Portuguese (Vasconcelos et al, 2009), Spanish, French, and German. The statistical analyses were carried out using the *Statistica*® Software package (version. 7.1, StatSoft 2005) to evaluate the findings. To assess the statistical significance (significance level of 5%) we considered the following hypotheses (these null and alternative hypotheses were also tested for Portuguese-, Spanish-, French-, and German-language journals):

Null Hypothesis (H_0): *Nps* in English-language international journals and writing skills are independent, i.e., one variable does not influence the other.

Alternative Hypothesis (H_1): *Nps* in English-language international journals and writing skills are dependent.

For data comparison, we considered the performance indicators of these researchers in the CNPq's database who had publications in the Web of Science, a total of 22,900 authors with

publications from 1945 to 2004. These authors' self-evaluation was as follows: 51.4% classified their writing skills as good, 34.0% as reasonable and 9.5% admitted to poor writing skills. We thus correlated their *Nps*, *Ncs* and h-indices with their writing skills in English (Vasconcelos et al, 2008) and in French. Here we show the correlations with *Nps* and with *Ncs*. The database that compiled information on the publications by these authors, the Brazilian Science Indicators (BSI), contained 188,909 references, with information on the type of publication, the full reference, the citations per author per year up to June 2005, the authors' names and addresses, institutions, cities, states and countries. A total of 150,323 research articles, 24,164 meeting abstracts, 5,541 notes, 3,577 letters and 2,333 reviews were considered for this statistical analysis, which was the complementary cumulative distribution function (CCDF) of researchers with different writing skills in English and in French. In all graphs, $P(X > x)$ is the probability that the variable X —the total number of researchers with a given number of papers in each group—will be higher than the given value on the x -axis. For the CCDF analysis, the following equation was used: $F_c(x) = P(X > x) = 1 - F(x)$.

Results

Correlation between Writing Skills in English and Nps in English-, Portuguese-, Spanish-, French- and German-journals (2001-2004)

Figure 1A shows the correlation between *Nps* in English-language international journals and English proficiency in writing, considering “good” and “poor” writing skills (Vasconcelos et al, 2009). As can be seen, p value is < 0.05 , and we thus adopt the alternative hypothesis, which assumes dependence between the two variables (*Nps* and writing skills). The contingency coefficient obtained was 45,5%, indicating a moderate to strong correlation among these variables.

Figure 1B shows this correlation for Portuguese, *i.e.*, publications in national journals (Vasconcelos et al, 2009). Note that a weak correlation between these variables ($r = 0.05$; $p = 0.8983$) was found for “good writing skills” and a weak to moderate correlation was obtained ($r = 0.69$; $p = 0.0401$) for “poor writing skills”. The contingency coefficient (8.14%) showed weak association between the variables, *i.e.*, *Nps* in Portuguese-language journals and writing skills in English. For Spanish, French and German, the contingency coefficients were as follows: 2.38% for Spanish, 3.12 % for French and 2.80% for German, showing that Brazilian researchers' writing skills in English do not correlate with publications in Spanish-, French- and German-language journals. To illustrate the results for these three languages, Figure 1C shows a weak correlation for Spanish* ($r = 0.40$; $p = 0.3232$) for “good writing skills” and $r = 0.27$; $p = 0.5106$ for “poor” writing skills).

* The data for Spanish-language publications are easily comparable because they cover the same range of publications on the abscissa as in our analysis for English- and Portuguese-language journals. Some gaps are found for French- and German-language publications.

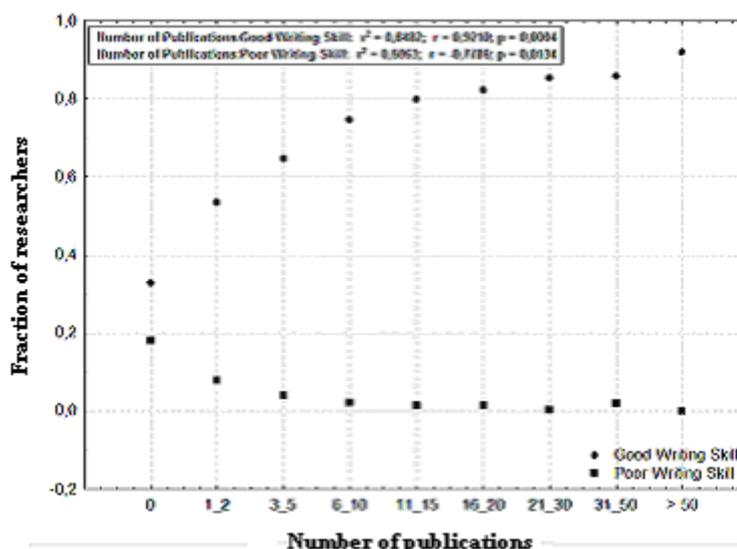


Figure 1A. Fraction of researchers with “good” and “poor” writing skills in English vs publication output in English-language international journals (Vasconcelos et al., 2009)

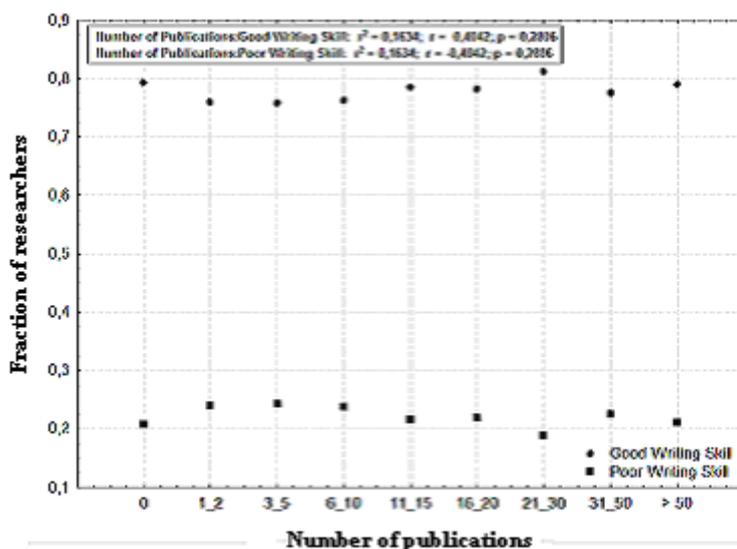


Figure 1B. Fraction of researchers with “good” and “poor” writing skills in English vs publication output in Portuguese-language international journals (Vasconcelos et al., 2009)

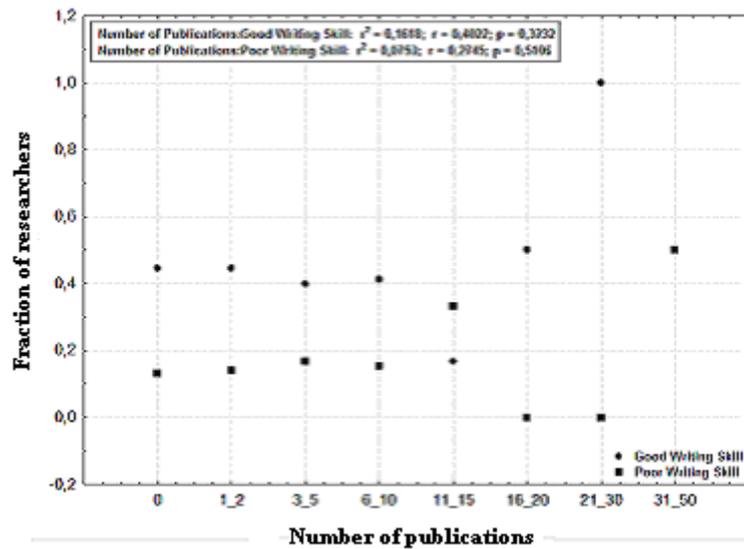


Figure 1C. Fraction of researchers with “good” and “poor” writing skills in English vs publication output in Spanish-language international journals

Correlation of Writing Skills in English and in French with Nps and Ncs for Publications in the Web of Science (1945-2004)

Figure 2 shows the comparison between productivity in the Web of Science for researchers with “good”, “reasonable” and “poor” writing skills in English and in French (inset of the Figure). As can be noted, authors with “good” writing skills in English are more evenly distributed along the whole range of publications and have a marked presence for higher *Nps*. This trend is not observed for “good” writing skills in French, where the three curves mostly overlap and the pattern of distribution obtained is rather different from that for English, indicating that writing skills in French do not play a role in the performance of these researchers. The same trend (Figure 3) is observed for English and French (inset of the Figure) when it comes to *Ncs*. The prevalence of more frequently cited papers is clearly observed for authors with “good” writing skills in English.

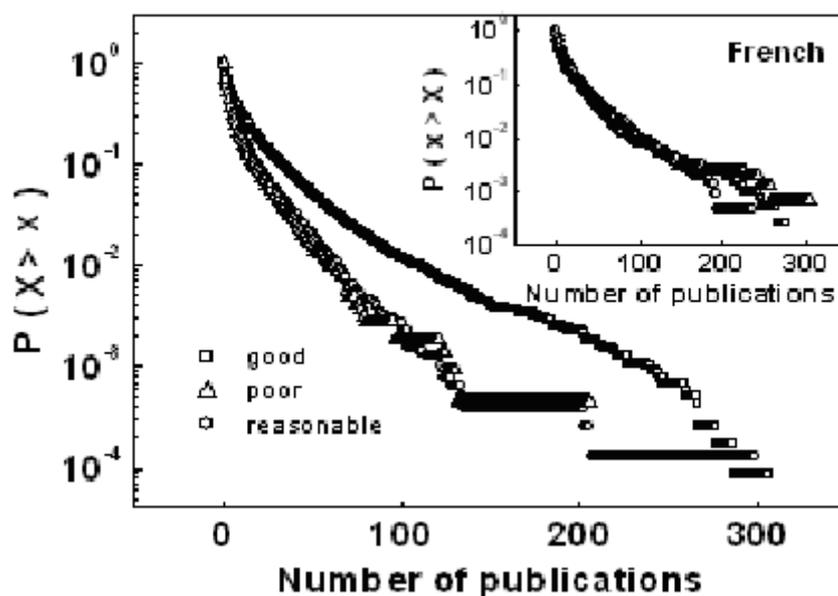


Figure 2. Complementary cumulative distribution function of researchers with different writing skills in English (Vasconcelos et al, 2009) and in French (inset of the figure): good (squares), reasonable (circles) and poor (triangles), according to Nps.

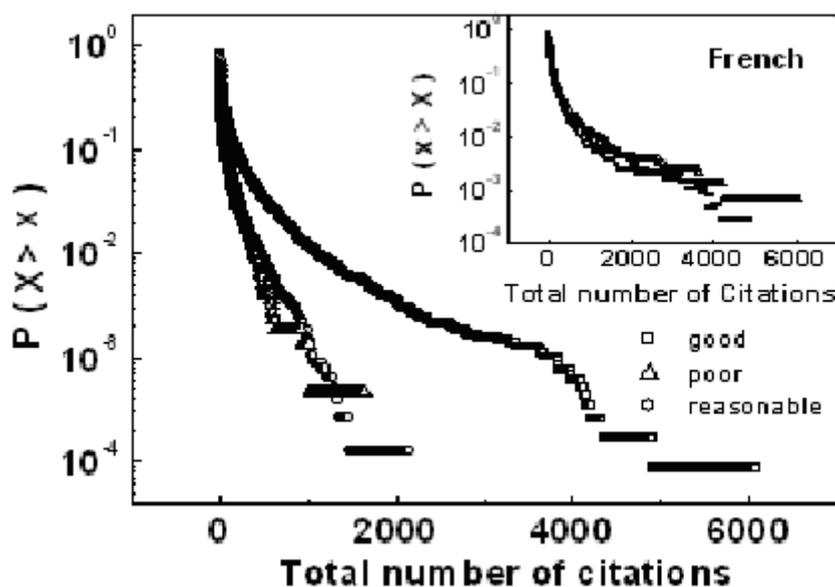


Figure 3. Complementary cumulative distribution function of researchers with different writing skills in English (Vasconcelos et al, 2009) and in French (inset of the figure): good (squares), reasonable (circles) and poor (triangles), according to Ncs.

Discussion and Conclusion

Our results corroborate the growing number of publications that point to the linguistic landscape of today's science as a relevant issue in the publication output of NNES authors (Benfield & Feak, 2006; Freeman & Robbins, 2006; Victora & Moreira, 2006; La Madeleine, 2007; Meneghini & Packer, 2007). Here we combined data we have already published with those that result from ongoing research. In fact, the comparison of data on the correlation between publication output in English-language journals and the writing skills of researchers

in that language and in other languages adds evidence of a linguistic factor permeating today's science. In our approach to this issue through Brazilian science, data on Portuguese, Spanish, French and German languages, which can be considered control groups for this analysis, the variable "linguistic competence" in the *lingua franca* of science influences research performance. This is a reasonable hypothesis when we consider the data from *Nps* and *Ncs* for publications in the Web of Science. As most of the publications appearing in this database are in English, accumulating a linguistic capital associated with this language has proved to be advantageous for these Brazilian researchers. The extent of the contribution of this capital depends on factors such as academic affiliations, research funding and number as well as type of collaborations, among others.

In fact, Pain (2007) calls attention to the role of collaboration in getting published in today's science and argues that "no matter what language you speak, building an international network is crucial to your scientific future." She mentions the words of Curry, a professor of Teachers of English to Speakers of Other Languages (TESOL) education. Pain quotes Curry, saying that "while researching the issues faced by scholars in Hungary, Slovakia, Spain, and Portugal... [he] found that "scholars sometimes don't have high-level English proficiency but publish in high [impact] journals." According to this TESOL professor, this is possible because these scholars "can draw [on] a network of people that help out." To gain a better understanding of the role of this linguistic factor in the performance of Brazilian authors, we have started to look at such a network considering not only the number and type of collaborations these authors have, but also the sequence in the list of authors for each publication in the Web of Science. So far, our data show that number of collaborations is not a factor that distinguishes authors with "good", "reasonable" and "poor" writing skills, as the patterns for these three groups overlap (data not shown). Now, we have started to collect data on the nature of these collaborations and pattern of authorship order for authors with different writing skills. Overall, our results strongly suggest that a linguistic factor underlies the research performance of these NNES researchers from Brazil. This finding makes us believe that other Latin American countries should assess the effect of English proficiency in writing on their research communities. We propose to create a network of scientometricians, linguists and educators in Latin America to evaluate whether a similar trend prevails in other Latin American countries.

References

- Ammon, U. (2008). How could international scientific communication be made fairer and more efficient? *The Scientist*, 22, 13.
- Benfield, J.R. & Feak, C. B. (2006). How authors can cope with the burden of English as an international language. *Chest*, 129, 1728-1730.
- Bauwens, L., Mion, G. & Thisse, J.F. (2007). The resistible decline of European science. Core discussion paper from the Belgian program on interuniversity poles of attraction. Retrieved January 21, 2008 from: <http://www.core.ucl.ac.be/econometrics/Bauwens/papers/CORE-DP2007-92.pdf>.
- Bourdieu, P. (1991). *Language and symbolic power*. Cambridge: Harvard University.
- Brazilian National Research Council (CNPq). Diretório de grupos de pesquisa. Available at: <http://dgp.cnpq.br/>.
- Freeman, P. & Robbins, A. (2006). The publishing gap between rich and poor: The focus of AuthorAID. *Journal of Public Health Policy*, 27, 196-203.
- Kotzin S. Journal selection for Medline (2005). World Library and Information Congress: 71th IFLA General Conference and Council "Libraries - A voyage of discovery." Available at: <http://www.ifla.org/IV/ifla71/papers/174e-Kotzin.pdf>.
- La Madeleine, B. L. (2007). Lost in translation. *Nature*, 445, 454-455.
- Lattes Platform. Available at: <http://lattes.cnpq.br/>.

- Man, J.P., Weinkauff, J.G, Tsang, M & Sin D.D. (2004). Why do some countries publish more than others? An international comparison of research funding, English proficiency and publication output in highly ranked general medical journals. *European Journal of Epidemiology*, 19, 811–817.
- Meneghini, R. & Packer, A. L. (2007). Is there science beyond English? Initiatives to increase the quality and visibility of non-english publications might help to break down language barriers in scientific communication. *EMBO Reports*, 8, 112–116.
- Pain, E. (2007). Moving out of the shadows: Publishing from the rest of the world. *Science Careers*, 2007. Retrieved April 13, 2007 from: http://sciencecareers.sciencemag.org/career_magazine/previous_issues/articles/2007_04_06/credit.a0700048.
- Piccoli, N. B. & Procianoy, R. S. (2007). Brazilian scientific journals: An overview. *European Science Editing*, 33, 73-74.
- Sousa Escandón, M. A., González, G. C. & Fernández, M. M. (2000). Which language will Medline speak in the next millennium?. *Archivos Españoles de Urologia*, 53, 93-99.
- Van Raan, A. F. J. (2004). Measuring science. In: Moed, H. F.; Glänzel, W.; Schmoch, U. (Eds.). *Handbook of quantitative science and technology research* (pp.19-50). Dordrecht: Kluwer Academic Publishers.
- Vasconcelos, S. M. R., Sorenson, M. & Leta, J. (2005). English proficiency and time to publication: The case of Brazilian science. In *Proceedings of ISSI 2005* (pp. 664-666). Stockholm: Karolinska University Press.
- Vasconcelos, S.M.R., Leta, J. & Sorenson, M. M. (2007). Scientist-friendly policies for non-native English-speaking authors: Timely and welcome. *Brazilian Journal of Medical and Biological Research*, 40, 743-747.
- Vasconcelos, S.M.R., Sorenson, M.M., Leta, J., Batista, P.D. & Sant'Ana, M. (2008). Researchers' writing competence: A bottleneck in the publication of Latin-american science? *EMBO Reports*, 9, 700-702.
- Vasconcelos, S.M.R., Sorenson, M.M. & Leta, J. (2009). A new input indicator for the assessment of science & technology research? *Scientometrics*, 80, in press.
- Victora, C. G. & Moreira, C. B. (2006). North-south relations in scientific publications: Editorial racism?. *Revista de Saúde Pública*, 40, 36-42.