

NOMINAL GROUPS AS AN INDICATOR OF NON-NATIVE ENGLISH COMMUNICATION PROBLEMS IN TOP-RANKED BRAZILIAN SCIENCE JOURNALS

GRUPOS NOMINAIS COMO UM INDICADOR DE PROBLEMAS NÃO NATIVOS DE COMUNICAÇÃO EM INGLÊS EM REVISTAS CIENTÍFICAS BRASILEIRAS DE ALTO *RANKING*



William F. Hanesⁱ
(Doutorando PGET, UFSC)
billeza@gmail.com

Abstract: This study is an attempt to document the problematic nature of an intermediary linguistic system, the lingua franca used by the scientific community, on the production and impact of science from the broad area beyond the inner circle of native English speakers. To this end, a random cross-sectional sample (n=5) of current English-language articles from top-ranked journals in the Brazil-based metapublisher Scientific Electronic Library Online (SciELO) was examined for grammatical issues, especially nominative group construction. In the studied sample, varying and in some cases elevated levels of L1 interference were found, indicating that on the best collective level, there are proficiency problems with the lingua franca, that these problems are not evenly distributed and that systematic language management yielded vastly different language quality outcomes.

Keywords: globalization; scientific discourse; lingua franca; SciELO

Resumo: Este estudo é uma tentativa de documentar a natureza problemática de um sistema linguístico intermediário, a lingua franca usada pela comunidade científica, na produção e impacto da ciência da ampla área para além do círculo de falantes nativos da língua inglesa. Para este fim, uma amostra transversal aleatória (n=5) de artigos em língua inglesa atuais retirados de revistas de alto ranking do *metapublisher* com sede no Brasil Scientific Electronic Library Online (SciELO) foi examinada com relação a questões gramaticais, especialmente a construção de grupos nominais. Na amostra estudada, foram encontrados níveis variados e, em alguns casos, altos de interferência da L1, indicando que, no melhor nível coletivo, há problemas de proficiência com a língua franca, que estes problemas não estão uniformemente distribuídos, e que o gerenciamento linguístico sistemático levou a resultados de qualidade linguística vastamente diferente.

Palavras chave: globalização; discurso científico; língua franca; SciELO

Introduction

Linguist M.A.K. Halliday writes that “learning science is the same thing as learning the language of science” (2004, p. 138), which means to say that the scientific method is inextricably embedded in discourse, more specifically, “a syndrome of features” that characterizes “scientific”

HANES. Nominal groups as an indicator of non-native English communication problems in top-ranked Brazilian science journals

Belas Infêéis, v. 2, n. 2, p. 127-139, 2014.

language (*ibid*: p. 141). Both the acquisition and transmission of knowledge, methodology, and results require proficiency in this linguistic genre. However, the increasing globalization of scientific endeavor, which appears to have entered an asymptotic curve in the last two decades due to a combination of (at least) advances in communication technologyⁱⁱ and business pressures from both the university (internally) and the publishing industry (externally), the target audience for scientific production is becoming increasingly international.

And, with almost seamless concomitance, this system has required a virtually exclusive dedication to a single language as the language of production (see DE SWAAN, 2002), such that “the language of science” has now become “scientific English”. Of the top 1000 journals in the 2011SCImago Journal Ranking, the current issue of only one was not exclusively in English.ⁱⁱⁱ Even Chinese Optics Letters (ranked 433rd) and Acta Physica Slovaca (ranked 456th) were completely in English: website and all articles.^{iv} In these journals, no mention is even made that submitted manuscripts must be in English- it goes without saying.

Such a situation is not unproblematic regarding the concept of ethnolinguistic democracy advanced by Fishman (1993), given that the majority of the world’s scientists are not native speakers of English and that competition for a place at the high impact table hinges, unfortunately, on more than solid research. Invisibility, the boogeyman (or perhaps operating strata) of translation, also haunts the scientific community. And although there has been a consistent thread on this subject in the literature at least since Gibbs’ 1995 article “Lost Science in the Third World”, effective steps toward development in what linguist Braj Kachru calls the expanding and outer circles (i.e., the periphery) of World English (1985) have not come from the center.

Rather, cooperative organizational efforts by scientists from developing countries have signaled the most impressive change. One such example is the SciELO collection, which currently represents over 1000 journals from 38 countries in Latin America, the Caribbean, Europe and Africa. Its database includes over 400,000 open-access articles and has recently become indexed in the PubMed Central archive, the largest worldwide science database. Under SciELO’s guidance, the operating procedures of member journals have been standardized and marked increases in impact factor have been observed^v. Although the metapublisher does not specifically regulate the language policy of member journals, its founders have indicated the

HANES. Nominal groups as an indicator of non-native English communication problems in top-ranked Brazilian science journals
Belas Infêís, v. 2, n. 2, p. 127-139, 2014.

primacy of the language question in articles such as “Is there science beyond English?”, stating that “the ability of scientists to communicate in the scientific lingua franca is part of a country’s scientific capabilities”(MENEGHINI & PACKER, 2007). In the same article, they also pointed out a recent trend toward increasing strictness in English standards in the literature:

During the past three decades, editors have become increasingly tough and demand better English in scientific manuscripts...Reviewers might also be more inclined to reject a paper because of poor English. Furthermore, the requirement for clear and understandable English increases with the prestige and/or impact factor of the journal, thus creating a language barrier that many scientists find difficult to overcome (*ibid*: p.114).

The grammatical metaphor

Alongside the various components of scientific discourse that Halliday lists (2004:162), i.e., interlocking definitions, technical taxonomies, special expressions, lexical density, syntactic ambiguity and semantic discontinuity, all of which contribute their own special impediments to translation, he pays special attention to a singular class called the *grammatical metaphor*. This category of semantic construction nominalizes previously-described processes, forming lexically dense blocks that are, in more accessible terms, ‘verbs-turned-nouns describing other nouns’. To illustrate, the following example is given: “glass crack growth rate”. This nominal group, translated into ‘English’ would be “the rate at which a crack in glass will grow”.

129

Nominalization as discursive power

It is important to point out that, in the above example, altering the mode of expression affects its relative discursive weight and accessibility:

[moving] from the clausal mode of grammar to the nominal mode, one tends to gain in discursive power, but by the same token one tends to lose most of the ideational-semantic information...[Nominal groups] can be obscure and highly ambiguous if one does not know what they mean in advance.” (HALLIDAY, 2004, p.197)

The discursive power gained by the use of nominal groups translates into authority since, lacking ideational-semantic information, only readers able to follow and understand the structure of the text can interpret them. Thus if only those who “know what they mean in advance” (i.e., expert

HANES. Nominal groups as an indicator of non-native English communication problems in top-ranked Brazilian science journals
Belas Infêís, v. 2, n. 2, p. 127-139, 2014.

readers) can decode them, then the writer, from a discursive point of view (the puzzle-maker), becomes a de facto expert. And authors who, in the opposite case, avoid or improperly construct nominal groups are prone to *discursive weakness*, which may be confounded with scientific weakness by science publication gatekeepers such as editors and peer-reviewers. Thus, authors without such mastery (either their own or that of their translators) and the journals who publish them would thus be at a disadvantage in competition with those who have it.

Although the grammatical metaphor could be considered a component of scientific writing worldwide, it bears upon the syntactical organization of English at a critical pinch-point: the prenominal adjective. Of course, anyone reading this article is familiar with the fact that, in English, adjectives normally precede the words they modify. In such nominal groups, however, the ‘modifier first’ model is also accompanied by the loss of grammatical words such as prepositions and articles. These features are not characteristic of a number of languages, notably Romance languages, spoken in several of the top-producing nations of scientific literature. Let’s exemplify this by translating *glass crack growth rate* into Portuguese. A literal result would look like “a taxa *de* crescimento *de* fissuras *em* vidro”,^{vi} which cannot suffer further condensation (e.g., preposition loss) without losing cohesion.

Given SciELO’s agenda to bring visibility to Brazilian science, its founders’ recognition of the question of power involved in the language issue, as well as the translation difficulties involved, especially the unnatural style inherent in English nominal group construction for native speakers of Romance languages, a study of grammatical issues, particularly nominative group construction, in a sample of recent English-language articles from some of its top-ranked journals might begin to clarify the linguistic sufficiency of the collection’s interface with the lingua franca.

Methods

Sample selection

Of the 18854 worldwide journals listed in the 2011 Scimago Journal Ranking (SJR), 239 (1.3%) are from Brazil. One original research article from the current issue of each of the five top-ranked Brazilian journals was selected by randomized drawing for analysis. Since the current issue of fifth-ranked *Cadernos de Saúde Pública* included no articles in English, the sixth-ranked

HANES. Nominal groups as an indicator of non-native English communication problems in top-ranked Brazilian science journals
Belas Infêis, v. 2, n. 2, p. 127-139, 2014.

Brazilian Journal of Medical and Biological Research, whose articles, site and submission process are exclusively in English, was used instead. The analysis was conducted by a native speaker of English who is a professional science translator. Greater detail about the sample of journals is provided in Table 1.

Table 1. Top-ranked Brazilian scientific journals according to the 2011 Scimago Journal Ranking.

Rank Brazil	Rank World	Title	SJR rating	Impact Factor 2011 (2-year)	SciELO
1.	2793	Memórias do Instituto Oswaldo Cruz	0.13	0.4420	Yes
2.	3419	Revista de Saude Publica (USP)	0.105	1.3356	Yes
3.	3795	International braz j urol : official journal of the Brazilian Society of Urology	0.093	0.2279	Yes
4.	3899	Brazilian Journal of Infectious Diseases	0.091	0.2293	Yes
5.	3938	Cadernos de Saúde Pública	0.09	0.9029	Yes
6.	3943	Brazilian Journal of Medical and Biological Research	0.09	0.1790	Yes
7.	3977	Revista Brasileira de Psiquiatria	0.089	0.6462	Yes
8.	4038	Jornal de Pediatria	0.088	1.1509	Yes
9.	4111	Revista do Instituto de Medicina Tropical de Sao Paulo	0.086	0.3276	Yes
10.	4644	Clinics	0.076	1.2404	Yes

131

Only original research articles were selected due to: 1) their having undergone a complete peer review process, i.e., having passed through both the editorial staff and peer reviewers, who represent the journal's normative authority with respect to language policy and quality; 2) the fact that they are a single genre with similar structure and objectives should result in more homogenous register and standards. Only articles whose authors were all affiliated with Brazilian institutions were included to help isolate the source of any L1 interference.

Based on a profile of characteristic Portuguese L1 interference developed from a pilot study, 14 error categories were formulated, which are shown with their respective results in Table 2. For the 'Nonexistent or weakly supported terminology' category, all unfamiliar or questionable

terms or usage were checked against results from the National Institutes of Health via Google search with the unfamiliar term in quotation marks plus “.nih.gov”. When any term yielded less than 100 hits, individual results were checked for their number of citations and, if need be, institutional affiliation of the involved authors. If these didn’t appear authoritative, the term was then considered “nonexistent or weakly supported”. Nominal groups were counted as incorrect if they were not completely condensed or the prenominal modifiers were out of order. Evaluation of the other categories followed general grammatical convention and any questions of style were submitted to the above-described search procedures; the ‘Oxford comma’, or lack thereof, was not considered an error.

Results

Table 2 presents the mean total error results by category for the five analyzed articles. Table 3 presents total errors vs. nominative group errors according to article. The error rate per thousand words calculated to make the articles comparable in spite of word count differences. Table 4 presents the nominative group error results vs. total errors by article as a bar graph. The full report for each article, including a quotation of every nominal group error, is available in the appendix.

132

Table 2. Mean error results by category for the five analyzed articles.

Awkward collocation (weak style/word choice unclear unintelligible)	25.5%
Nominal group error (subset of awkward collocation)	15.2%
Punctuation/capitalization error (unnecessary or missing comma run-on sentences all others)	12.5%
Preposition use error	10.8%
Article error (unnecessary or lacking; definite vs. indefinite)	9.7%
Connecting word/phrase (i.e. conjunction) error	4.4%

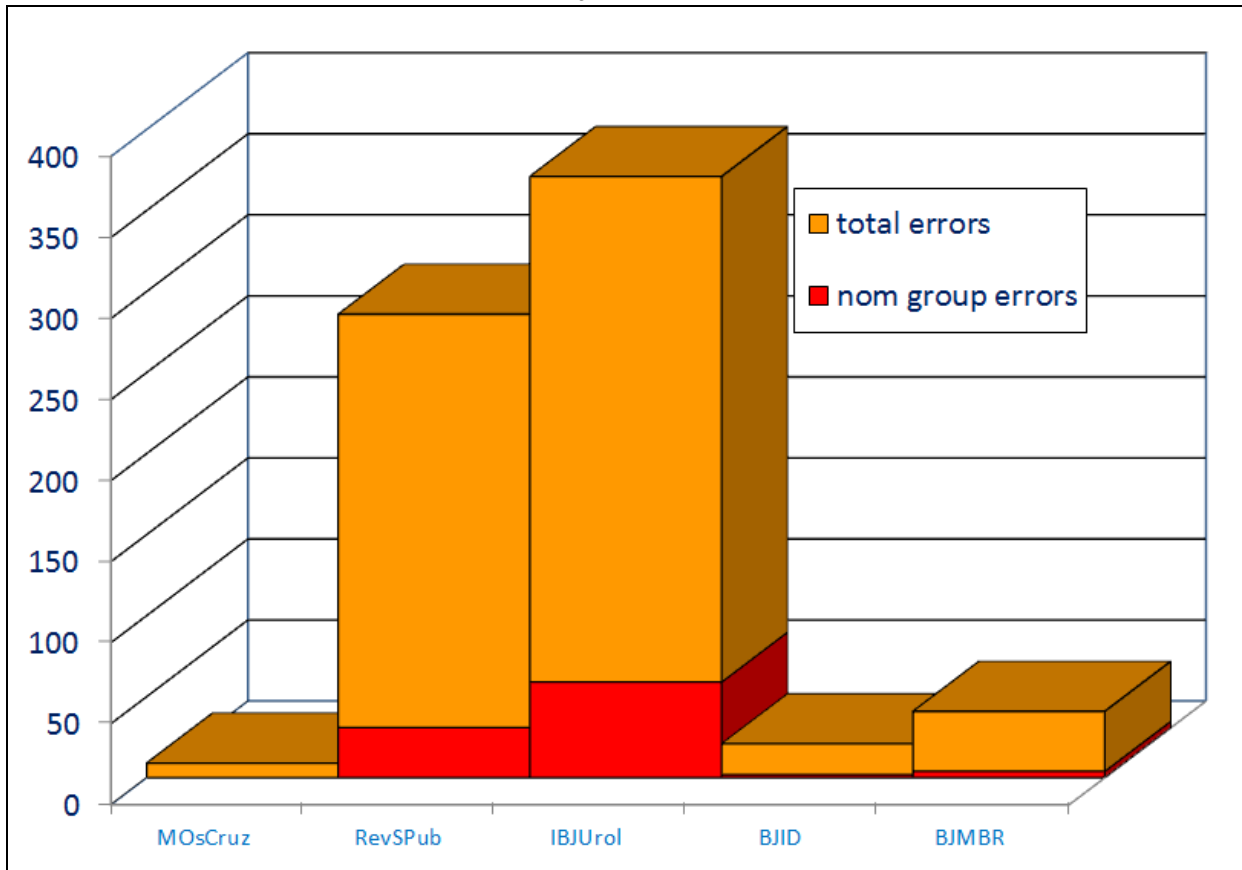
Verb error (present perfect all others)	4.4%
False cognate error	4.1%
Nonexistent or weakly supported terminology	4.1%
Infinitive/gerund error	3.2%
Plural error (noncount all others)	3.0%
Pronoun error (all others inverted syntax)	2.1%
Orthographic/typographical error	1.3%
Comparative/superlative error	0.4%

Table 3. Error rate by article.

Journal	Article	Word count	Non-NG Errors	NGE	Error rate per 1000 words	NGE/Error ratio
Memórias do Instituto Oswaldo Cruz	PRESTES et al. Vol. 107(6): 713-719, September 2012	3734	9	0	2.41	0.0
Revista de Saúde Pública	SILVA and RIBEIRO 2012; 46(4): 694-701	2780	255	31	102.88	0.122
International Brazilian Journal of Urology	PEGOLO et al. (plus editorial comment) Vol. 38 (4): 448-455, July - August, 2012	3266	312	59	113.59	0.189
Brazilian Journal of Infectious Diseases	TUON et al. 2012; 16(5): 416-419	2001	19	2	10.49	0.105
Brazilian Journal of Medical and Biological Research*	MARIANO et al. (2012) 45: 1017-1024	4315	37	4	9.50	0.108

Legend: NG= nominal group; NGE = nominal group error

Table 4. Nominative group errors relative to total errors by journal.



Legend: MOsCr= *Mémoires do Instituto Oswaldo Cruz*; RevS Pub= *Revista de Saúde Pública*; IBJUrol= *International Brazilian Journal of Urology*; BJID= *Brazilian Journal of Infectious Diseases*; BJMBR= *Brazilian Journal of Medical and Biological Research*.

As shown in Tables 2 and 3, a conspicuously high error rate was observed in two of the five articles. However, the quantity of errors was not correlated with the journal's ranking or impact factor, since the *Revista de Saúde Pública* had the highest impact factor on the list and the second-highest error rate. Nominative group errors varied from 0% to 19% of the total errors, with a combined mean of 15.2%. The most prominent error classes (i.e. those over 10%) were awkward collocation, nominal group construction, punctuation/capitalization and preposition use.

Discussion

Nominal group hypothesis

The fact that nominal group errors were so prominent corroborates the initial hypothesis that this important type of construction, which Halliday identifies as foundational to scientific English, is especially problematic for non-native speakers. The fact that nominal group errors were only outstripped by awkward collocation of all types (25.5% vs. 15.2%) only serves to reinforce that questions of collocation (the two combined categories represented 40.7% of all errors) seem to be the primary source of proficiency problems in this sample of articles. Moreover, since the present perfect, which represented half of all verb errors, is little used in Portuguese and since preposition usage error (there are +/-28 prepositions in Portuguese vs. +/-61 in English) was a top category also points to interference due to structural differences between the languages, rather than to other issues such as spelling (only 1.3% of the total- spellchecking tools?), verb conjugation or other particulars of the language such as noncount nouns, false cognates, etc. That the results for weakly supported or nonexistent (i.e. invented or literally-translated) terminology were relatively low may also be indicative of the internationalization of research, i.e., that Brazilian scientists are consumers of the international scientific literature, either alongside or instead of locally-produced Lusophonic science, which will be discussed in greater detail below.

135

Error rate variance between journals and language management

The error rate distribution among articles was very uneven, leaping from virtually nil (2.41 errors per 1000 words) in *Memórias do Instituto Oswaldo Cruz* to over 100 errors per thousand (yes, 10%) in two others. This brings up the topics of language brokers and journal editorial policy. An important clue to the reason why the article from *Memórias do Instituto Oswaldo Cruz* was virtually error-free is revealed on their webpage on editorial policy: “Papers submitted to *Memórias* will undergo the Premium Editing of English language review made (sic) by American Journal Experts, further suggestions will enhance nonnative English writing style”.^{vii} American Journal Experts is an academic editing and translating firm based in Durham, NC, USA that employs only “subject-expert editors” with advanced degrees.^{viii} This notice only

HANES. Nominal groups as an indicator of non-native English communication problems in top-ranked Brazilian science journals
Belas Infêis, v. 2, n. 2, p. 127-139, 2014.

appeared on the *Memórias* site, however, after December 2012.^{ix}Besides foreign outsourcing, A variety of approaches exist for journals to manage language quality, which depend on goals, funding and available staff. Although it is beyond the scope of this study to debate the merits of competing editorial policies, the results clearly indicate the drastic difference in proficiency that results from not systematically implementing some form of language quality control. Thus, the question remains about the availability and access to competent local personnel for handling the steadily increasing volume of Brazilian scientific literature. Another of the relatively error-free articles, TUON et al., was published in the *Brazilian Journal of Infectious Diseases*, which identifies itself as a (paid-access only) Elsevier journal since 2010, although it still appears on SciELO.^xBesides the fact that this journal has come under the control of one of the leading commercial science publishers, these authors acknowledged funding from a number of pharma giants including Merck, Pfizer, Novartis, Roche, Bristol-Myers Squibb and GlaxoSmithKlein, who would have a vested interest in guaranteeing the quality of the language.

136

Ranking and (quality) English

Another interesting aspect of the results is that SJR ranking had no apparent association with English quality in this sample of journals. Of course, the validity of both the impact factor and ranking in general have been called into question (see LAMBERT's forthcoming essay "Does ranking rhyme with banking?", SEGLEN 1997, OPTHOF 1997, etc.), but given the above-mentioned increasing strictness in language quality that top journal editors are currently requiring, it seems curious not to find a 'trickle-down' quality effect. A simple explanation for this may be a "for-us-by-us" posture among lower-ranked non-inner-circle journals. Packer clearly demonstrates a heterogeneity of citation, and hence consumption, among Brazilian scientists in that of the 20 most-cited titles in SciELO-Brasil articles from July 2006, the *Revista de Saúde Pública* and *Memórias do Instituto Oswaldo Cruz* were numbers two and three, respectively, after the Lancet and before the New England Journal of Medicine. In fact, eight of the top 20 were Brazilian journals (2007, p. 57), so local science has not yet lost all relevance in this community. This is all the more obvious since *Cadernos de Saúde Pública*, ("even") with no English articles, was number 10 in this list as well as the fifth-highest SJR ranked Brazilian journal, which may be due to the sheer number of citations mainly by other Brazilians. Thus, English is not necessarily a

HANES. Nominal groups as an indicator of non-native English communication problems in top-ranked Brazilian science journals
Belas Infêéis, v. 2, n. 2, p. 127-139, 2014.

prerequisite to at least some level of international ranking or a place as a regional reference, although, as mentioned above, not at the pinnacle of what Packer calls the “prestige pyramid”.

Study limitations

In the first place, as an exploratory study, the small sample size limits its applicability. Because sample size calculations were not carried out, neither the statistical significance of the results nor the power of detection were considered. Moreover, a sole rater was involved and, thus, there was no inter-rater agreement, which has been demonstrated as problematic elsewhere (e.g., TETREAULT & CHODOROW, 2008). However, this problem is partially mitigated by the fact that the rater is a native-English-speaking science translator.

Implications for further research

The implications for follow-up research are many. In the first place, as mentioned above, for more conclusive evidence, studies with larger samples and a more thorough statistical design (including the involvement of more than one rater) are called for in the case of both SciELO and top journals published in the outer and expanding circle for a meta-regional perspective. To obtain baseline error data, a sample of highly-cited articles could undergo the same criteria of error analysis. Moreover, an in-depth analysis of the corpus of key journals could be traced diachronically (e.g. before and since going online), as well as comparison of language policy and management with editorial board composition.

137

Conclusion

The nominalization of processes necessary for scientific discourse is indeed a problematic area for Brazilian scientists writing in English, ranking as the second most common error type, although such errors were few in journals either controlled by commercial publishing interests or who resorted to native-English-speaking outsourcing services to control language quality. The mere fact of being a SciELO member journal, even among the top five, had no apparent effect on language quality. However, English quality, in turn, or even the use of English at all, was not directly correlated with journal ranking.

REFERENCES

- DE SWAAN, A. **Words of the world: the global language system**. Cambridge: Polity Press, 2001.
- FISHMAN, J. Ethnolinguistic democracy: varieties, degrees and limits. In: **Language International**. 1993, v. 1, p. 11-17.
- GIBBS WW. Lost science in the third world. **Scientific American**. 1995; 273:76-83.
- HALLIDAY, M.A.K. HALLIDAY, M. A. K/ WEBSTER, J. **The Language of Science** (Collected Works of M.A.K. Halliday). Continuum: London/New York, 2004.
- ILIESCU GHEORGHIU, C.; LAMBERT, J. (Eds.) **Universe-cities as Problematic Global Villages: Continuities and Shifts in our Academic Worlds**. EDUFSC, forthcoming.
- KACHRU, B. B. Standards, codification and sociolinguistic realism: the English language in the outer circle. In: QUIRK, R. & WIDDOWSON, H.G. (Eds.). **English in the world: Teaching and learning the languages and literatures**. Cambridge: Cambridge University Press, 1985. p. 11-30
- MENEGHINI, R.; PACKER, A. L. Is there science beyond English? **EMBO Reports**, 2007, v. 8, p. 112-116. Available at: <<http://www.nature.com/embor/journal/v8/n2/pdf/7400906.pdf>> Accessed 02.02.2012.
- ONG, W. **Orality and literacy: the technologizing of the world**. New York: Routledge, 2002.
- OPTHOF, T. Sense and nonsense about the impact factor 1997. *Cardiovascular Research* 33 (1997) I-7
- PACKER, A. **CODATA 2007 - Strategies for Open and Permanent Access to Scientific Information in Latin America: Focus on Health and Environment Information for Sustainable Development**. Atibaia, São Paulo, 8-10 May 2007
- SCIMAGO JOURNAL RANKING 2011. Available at: <<http://www.scimagojr.com/journalrank.php>>. Accessed 02.02.2012.
- SEGLÉN, P. **Why the impact factor of journals should not be used for evaluating research**. *BMJ*: 1997, 314, p. 498–502.
- TETREAUULT, J.R.; CHODOROW, M. Native judgments of non-native usage: Experiments in preposition error detection. **Proceedings of the Workshop on Human Judgements in Computational Linguistics**. Association for Computational Linguistics, 2008.

ⁱ Currículo Lattes William F. Hanes. Disponível em: <http://lattes.cnpq.br/5710028292619217>

ⁱⁱ in agreement with Walter Ong's theory (1982,2002)

ⁱⁱⁱPhlebologie published in Germany, ranked 638th

^{iv}Moreover, of the top 100 scientific journals, only 18 are published outside the US/UK; of these 18, only 2 were not affiliated with the giant science publishing houses (i.e. Elsevier, Wiley, Springer or Cell Press), but that's another story (or two).

^v <http://www.scielo.org/php/index.php?lang=en> accessed 2013.08.04

^{vi} However, this literal translation was unclear to the native speaker of Portuguese with whom I checked. A more acceptable version with respect to comprehension would be “a velocidade na qual uma fissura em vidro expandirá”, which involves altering the main process.

^{vii} <http://memorias.ioc.fiocruz.br/editorial--policy>

^{viii}<http://www.aje.com> the language menu of whose site, interestingly, appears in Portuguese first followed by Chinese, English and then Spanish. Besides the headquarters in Durham (not too far from Duke University), they have one other office in São Paulo. accessed 2013.08.04, all of which is revealing with respect to the market.

^{ix}Was such a policy in place before, but left unstated? If so, how long have they been outsourcing? Who pays for the service- the authors or the journal?

^x <http://www.bjid.org.br/> “Since March 25, 2010, the site of the Brazilian Journal of Infectious Diseases has restricted access. Only members who have paid the current membership fee to Sociedade Brasileira de Infectologia (SBI) will be able to view full articles.” I was able to browse and download freely from the current issue on SciELO, however.