

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

Rogério Meneghini & Abel L. Packer



Of the past 25 winners of the Nobel Prize in Literature, only 9 wrote their masterpieces in English; the remaining 16 laureates wrote in other native languages. Many of their works were eventually translated into English, which was probably necessary for international recognition and the attention of the Swedish Nobel Prize committee. The translators faced the arduous task of transferring the splendour of the original text into a different semantic, syntactic and sometimes cultural context to make it appeal to a wider audience.

Most Nobel laureates in Physics, Chemistry, and Physiology or Medicine do not face the challenge of translating their works into another language before gaining recognition. Many speak English as their first language and even non-native English speakers usually publish their discoveries in English. Furthermore, given that English is the *lingua franca* of science, the international community—including the Nobel Prize committees at the Royal Swedish Academy of Sciences and the Karolinska

Institute—do not have to wait for a translation. The situation is reminiscent of the ancient and medieval worlds, when scholars could communicate only in Latin until great historical literates, such as Dante Alighieri, William Shakespeare and Martin Luther, promoted the use of the Italian, English and German languages, respectively, by writing in their native languages.

Any scientist must therefore master English—at least to some extent—to obtain international recognition and to access relevant publications

In fact, English has become the modern *lingua franca* in a world that is economically, scientifically and culturally largely dominated by Anglo-American countries. Any scientist must therefore master English—at least to some extent—to obtain international recognition and to access relevant publications. But although this makes

communication between scientists much easier, it also creates problems for non-English-speaking countries. Even if their scientists are able to read English publications, to reap the societal benefits they must still translate this knowledge into a national context.

Conversely, many scientists in Africa, Asia, Latin America and Europe still publish their work in national journals, often in their mother tongue, which creates the risk that worthwhile insights and results might be ignored, simply because they are not readily accessible to the international scientific community. To overcome this dilemma, several initiatives now aim to strengthen the impact and quality of national journals with the goal of gaining greater international visibility for articles published in a language other than English.

Scientific English has come a long way since the first publications appeared in the eighteenth century, notably by national scientific academies. Those who have read Charles Darwin's *On the Origin*

of *Species* might remember sentences such as "...man with all his noble qualities ... still bears in his bodily frame the indelible stamp of his lowly origin" in the middle of the discussion on theoretical tenets. More recently, James Watson and Francis Crick's publication on the structure of DNA closes with the coy phrase, "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material" (Watson & Crick, 1953). These exemplify arduous efforts to translate scientific thought into artful communication.

...scientific English has become a communication tool in a less erudite world, consisting of those who want to learn about and pass on knowledge

However, such beautiful sentences have become a rarity; today's scientific English bears little resemblance to the literary form, and has developed its own rules, style, virtues and habits (Burnett, 1999). The pronouns 'I' or 'we' are rarely used; instead the passive voice is commonplace. Its structure has become less flexible to conform to the requirements of communicating information in limited space, the need for rigour and the huge amount of articles published. This trend has created a 'scientific English' that is more accessible to the reading and writing capabilities of non-English-speaking researchers. In general, scientific English has become a communication tool in a less erudite world, consisting of those who want to learn about and pass on knowledge.

However, even if English is the dominant language in science, it is certainly not the native language for the majority of scientists. This creates a dilemma for readers and authors alike. Authors usually want to attract interest to their work to enhance their reputation both nationally and internationally. Conversely, readers and users, such as healthcare professionals and journalists, find it easier to access and understand an article if it is written in their native language. This is particularly relevant for biomedical research with clinical implications because professionals need to access the latest results—usually published in English—to benefit their patients, with whom they communicate in their native language. But it also creates the

risk that important knowledge is 'lost' if it is not published in English, which might delay or prevent further research or policies. For example, as early as the 1930s, German scientists showed a significant causal link between smoking and lung cancer, and published their research—in German (Proctor, 1999). But their findings were largely ignored until the 1960s when British and US scientists rediscovered the link, which eventually triggered public health policies and programmes to reduce smoking.

Even in a country with a strong scientific tradition such as Germany, the language problem has been acknowledged and a possible solution suggested: "The real solution seems to be to wait for Germany to become a genuinely bilingual society, using English as the global language of science and German as the local language spoken and read by health professionals and patients" (Ofori-Adjei *et al*, 2006). Every non-English-speaking country must address this problem in some way. For example, several Spanish journals in clinical medicine have decided to publish their articles in both Spanish and English, to address healthcare professionals and international scholars, respectively (Bordons, 2004).

The open-access movement and the increasing importance of the worldwide web are creating new ways to deal with the language problem. Public Library of Science (PLOS) journals now encourage non-English-speaking authors to provide a version of their article in its original language as supporting material (Editorial, 2006). Conversely, an increasing number of non-English journals provide English translations of their articles. This is an additional burden on authors because they have to provide their own translation, but it is rewarding in terms of gaining larger visibility through the English version of their articles. This approach might also solve the dilemma that many authors face when they try to attract both national and international interest. At the same time, some countries are making extra efforts to gain scientific visibility in their own language. Not surprisingly, France has several initiatives in the area of agricultural research, and makes French journals in the areas of human and social sciences—strong research fields in the country—freely available to French-speaking developing countries.

It is interesting to note that in higher education, the language used by online

resources has increasingly shifted towards the national language (Holmes, 2004); however, the scientific literature is somewhat different to teaching resources. Although there are thematic areas and articles of particular interest to a national audience, authors favour publishing their articles in international journals with an attractive impact factor. Inevitably, these journals use the English language and thus further tilt the balance towards the *lingua franca*. If journals could set a trend towards a more balanced use of languages in scientific publications, this might help to reverse the decline of other languages in international scientific communication.

However, other countries have taken the opposite approach and are making efforts to publish their science in English. China, for example, has established an agreement with the scientific publishing house Springer (Heidelberg, Germany) to select the best articles from more than 1,700 Chinese university journals and translate them into English. The Japan Science and Technology Information Aggregator has indexed more than 500 Japanese journals covering all scientific areas, a large number of which are available in English. Similarly, Czech, Hungarian and South Korean journals indexed by Thomson Scientific (formerly ISI; Philadelphia, PA, USA) are almost all published in English.

If journals could set a trend towards a more balanced use of languages in scientific publications, this might help to reverse the decline of other languages in international scientific communication

These initiatives are still new and somewhat ambiguous, and it is not clear what consequences they might have for language usage in science. However, it is clear that two important forces drive these actions. Scientists look for international visibility by publishing in English either in national journals or in high-profile international journals; conversely, they hope to attract a larger regional audience by publishing in their mother tongue or they choose a national journal because they are not sufficiently fluent in English.

The latter is a growing problem. During the past three decades, editors have become increasingly tough and demand better English in scientific manuscripts. Some might remember the days before the late 1970s when editors would practically rewrite papers that they thought were worth publishing, but a growing number of submissions now makes that level of service impossible. 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.

Those who practice and teach science in English-speaking countries might not find it particularly difficult to gather information and new concepts and to disseminate knowledge to their students. Communication through congresses and journals is rapid and creates a quick convergence of scientific terminologies, so that its transfer to the classroom is relatively uniform. This, however, is not the case in non-English-speaking countries. Scientists and teachers alike must read and lecture about concepts and knowledge that originated in English publications. In biochemistry and molecular biology in particular, this creates the problem of how to keep up with a continuous flow of English information, ideas and terminologies while communicating and teaching in another language. The problem is even more critical in developing countries because few young people learn English.

In addition to the practical difficulties, science is part of culture. It is not done in an ivory tower separated from the rest of society but is recognized as an essential source of knowledge for economic and technological development. Its results, terms and ideas pervade the society surrounding it and create new products, services and expressions that ultimately enter common usage. If there is no effort to create scientific semantics in the native language within a national scientific community, the country and its culture will not be able to absorb the scientific ideas and knowledge that ultimately serve its society (Kushner, 2003). This semantic conflict between science and culture was first put forward—of course—in the English literature. Charles Snow identified the lack of a bridge between science and the humanities—‘the two cultures’—

and claimed a need for a better cultural assimilation of science (Snow, 1959). In fact, science has created an abundance of terms that were absorbed into the language by the media and by common usage. In the 1960s, terms such as DNA, genetic code, black hole and entropy were mostly unknown to laymen. Today, not only are these terms familiar to the average educated citizen, but also the ideas they implicate are easily recognizable.

...the ability of scientists to communicate in the scientific *lingua franca* is part of a country's scientific capabilities

Should we therefore make additional efforts to incorporate scientific English into our culture or should we improve our native scientific language? There seems to be no easy solution, although both alternatives present considerable challenges for any non-English-speaking countries. If we abandon efforts to create a scientific mother language, individual scientists must have some way of tapping into the main sources of scientific knowledge, namely English publications. Therefore, the ability of scientists to communicate in the scientific *lingua franca* is part of a country's scientific capabilities.

The burden of having to deal with two languages can actually be beneficial as it creates new opportunities and experiences; however, this is a distant goal for most developing countries. Most students with a PhD degree in these countries can read an English text in their area of knowledge, but they rarely have mastered English sufficiently to write a clear and concise text. By contrast, the progress of a scientific mother language also faces obstacles, as we will describe using Brazil as an example.

Brazilian scientists currently publish about 50,000 articles each year, of which approximately 60% are in Portuguese. About 18,000 of these articles are published in journals that are indexed by the Thomson Scientific Web of Science database; only 2.7% of these articles are in Portuguese. On the surface, Brazil seems to have solved the two-language problem, given that a considerable amount of its scientific output is published in the mother tongue. However, this is not quite true: the quality of the English

and the Portuguese articles differs significantly; most of the latter are published in journals that address only a small community covering peripheral interests, and that have little or no peer-review at all.

This does little to contribute to a non-English scientific language. The more refined and elaborate thoughts require a sophisticated vocabulary and semantics to be adequately expressed; conversely, publishing only poor science falls short of improving the Portuguese scientific language. This is one of the reasons why the Scientific Electronic Library Online (SciELO) was launched.

A few Brazilian journals, which are run by dedicated individuals with limited financial support from government agencies, publish good quality papers; the condition of journals in other Latin American and Caribbean countries is similar. But—with a few exceptions—this mode of operation is insufficient to improve these journals further; consequently, a new approach was conceived. In 1997, the Latin-American and Caribbean Center on Health Sciences Information (BIREME), the Pan American Health Organization (PAHO; Washington, DC, USA) and the World Health Organization (WHO; Geneva, Switzerland), in partnership with the State of São Paulo Research Foundation (FAPESP), launched SciELO—a non-profit electronic metapublisher for scientific journals—which was further supported from 2002 by the Brazilian National Science Research Council.

The SciELO programme was established with three main objectives. The first was to publish a collection of the best Brazilian journals online following the open-access model, under which full-text articles can be freely accessed to achieve wider national and international visibility (Packer *et al*, 1998; Meneghini, 2003). In October 2006, the programme included 173 journals and 65,000 articles in the SciELO/Brazil collection (www.scielo.br). The initiative was soon adopted by Chile and then spread to other Latin American and Iberian countries, and now encompasses more than 350 journals.

The second objective was to improve the quality of the journals in those countries that adopted SciELO with respect to relevance and accuracy of the articles, care in presentation and the assessment of articles by referees. All these requirements are judged for each journal by an external *ad hoc* panel.

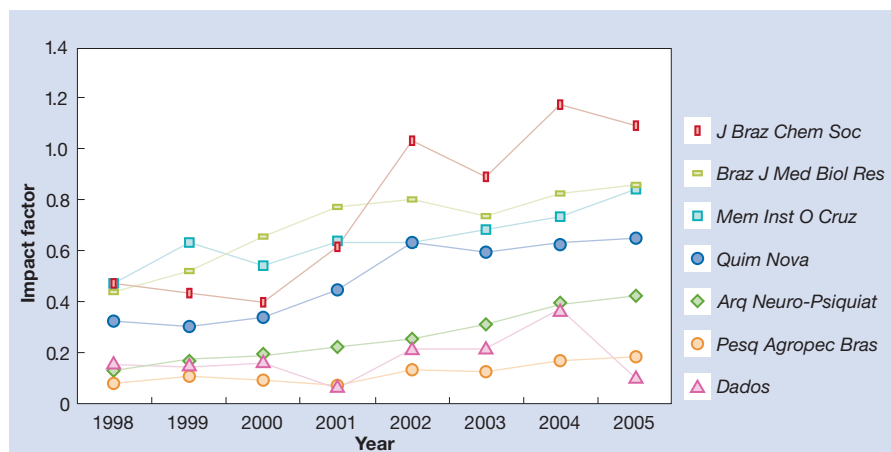


Fig 1 | Impact factor of the Scientific Electronic Library Online (SciELO) journals from 1998 to 2005. Source: Thomson Scientific Journal Citation Reports

The third aim was for SciELO to create a bibliometric database to produce performance indicators similar to those provided by Thomson Scientific's Web of Science and Journal Citation Reports. This database allows the systematic generation of data that might be important to subsidize political decisions in the scientific arena (Meneghini *et al*, 2006). At present, approximately 30% of articles in the SciELO database are published only in English. SciELO's policy has been that the editors of each journal should decide on what language the article should be published in, whereas SciELO supports the assessment of scientific quality.

Although SciELO and other initiatives might help non-English-speaking countries to establish and improve their scientific language ... they still do not fully solve the problem of how and where to find the 'lost' science...

Google Scholar has become a significant indicator of the increasing visibility of SciELO journals: after a link was set up from this portal to SciELO articles in 2005, SciELO is now ranked among Google Scholar's ten most accessed sites. Access to individual articles increased from two million in January 2006 to six million in May 2006. This visibility is likely to influence the impact factor of the 19 SciELO journals now

indexed by Thomson Scientific. Fig 1 shows the performance of seven of these journals, which have been indexed in both databases since 1998. These journals have steadily increased their impact factors by an average of 100% during this period. For comparison, the top 20 journals in the areas of biology, medicine, chemistry and physics indexed by Thomson Scientific in 2005 increased their impact factors in the same period by an average of 42%.

With the exception of China and South Korea, this is the biggest increase in impact factor for journals from any developing country with more than ten journals indexed by Thomson Scientific's Journal Citation Reports. Interestingly, China and South Korea are also the two countries with the highest percentage of journals indexed that are managed by commercial publishers. One might argue that this was an important reason for their success; SciELO, although it is non-commercial, could therefore have a similar role as an umbrella organization for Brazilian journals.

Although SciELO and other initiatives might help non-English-speaking countries to establish and improve their scientific language—and eventually the quality of science—they still do not fully solve the problem of how and where to find the 'lost' science, hidden beneath an unfamiliar language (Gibbs, 1995). Since the publication of Gibbs' article, the internet has markedly changed how we communicate with each other, but it has not changed the ability of scientists to

uncover findings of interest if they are behind language barriers. Economically successful countries—such as China, India, South Korea and various South American nations—are increasing, or plan to increase their investments into scientific research and development, which will inevitably lead to more publications, many of which will be in the local language. These could well be ignored elsewhere, just as much of the work now published in Japanese or French fails to be appreciated by other scientists, notably in the USA (Barany, 2005).

Although Fig 1 shows an increasing international visibility of Brazilian journals, the majority of SciELO articles are accessed from countries in which the primary language is Portuguese or Spanish. This is, of course, unsurprising given the similarity between these two languages, and it might further strengthen scientific collaboration among the 29 countries where either language is used. This is also reflected by the impact factor: nine out of the ten top-ranked journals in the SciELO Brazilian collection from 2003–2005 published articles predominantly in Portuguese.

...many journals would benefit if articles from non-English-speaking authors were published in their original language and in English

However, one of the goals of the SciELO programme is to achieve broader international visibility for Brazilian publications. The best science from the countries covered by SciELO—with exceptions in a few areas—is already published in international journals (Meneghini *et al*, 2006). This will certainly remain so in the near future as SciELO journals are not yet widely recognized as attractive channels of communication, owing to concerns about their scientific quality and that most articles are not published in English. The best strategy seems to be to deal with both difficulties in a stepwise manner. Scientific quality will eventually increase through SciELO's professional management of the peer review system. As more bibliometric data are made public to show the increasing quality of SciELO journals, they will then attract more attention from both the scientific community and policy makers.

The language barrier is not an exclusive problem to SciELO journals; many journals would benefit if articles from non-English-speaking authors were published in their original language and in English. However, a proper translation of a scientific text still requires a human translator, which renders the process too expensive to implement widely. Automating this process by translation software will certainly take some time (Barany, 2005; Heartfield, 2004), although some authors are confident that sophisticated software might become available in the near future (Kurzweil, 2002). At present, such technologies are useful for getting a general idea of a text, but they are certainly not able to create a precise translation, which is particularly important for scientific papers.

Online publishing of bilingual articles is neither difficult nor costly for journals in the era of the worldwide web, but it still requires human assistance and improvements to the structure of scientific communication. At this point, it would be wise to encourage authors to do so by their own efforts. Of course, only visionaries or ambitious or well-funded researchers would be inclined to do so initially. Conversely, both international and national journals should consider offering two versions of every article: one in the author's native language and one in English. In addition, the international indexes need to address the issue of how to treat multilingual versions of articles, including the presentation of metadata and full texts, inserting context-sensitive linking

and a comprehensive calculation of citations. This will require a combined and international effort by editors, index providers and authors, as well as support from national research agencies interested in developing native scientific languages and increasing both national and international visibility of national research. This truly bilingual publication system would be an important first step to overcoming language barriers in scientific communication and in moving towards a real global publication system.

REFERENCES

Barany MJ (2005) Science's language problem. *Business Week*, 16 Mar. www.businessweek.com
 Bordons M (2004) Hacia el reconocimiento internacional de las publicaciones científicas españolas. *Rev Esp Cardiol* **57**: 799–802
 Burnett DG (1999) A view from the bridge: The Two Cultures debate, its legacy, and the history of science. *Daedalus* **128**: 193–218
 Editorial (2006) Ich weiss nicht was soll es bedeuten: language matters in medicine. *PLoS Med* **3**: e122
 Gibbs WW (1995) Lost science in the third world. *Sci Am* **273**: 92–99
 Heartfield K (2004) Lost in translation. *The Ottawa Citizen*, 12 Feb
 Holmes HK (2004) An analysis of the language repertoires of students in higher education and their language choices on the internet (Ukraine, Poland, Macedonia, Italy, France, Tanzania, Oman and Indonesia). *Intern J Multicult Soc* **6**: 52–75
 Kurzweil R (2002) *Are We Spiritual Machines? Ray Kurzweil vs. the Critics of Strong A. I.* Seattle, WA, USA: Discovery Institute
 Kushner E (2003) English as global language: problems, dangers, opportunities. *Diogenes* **50**: 17–23

Meneghini R (2003) SciELO project and the visibility of 'peripheral' scientific literature. *Química Nova* **26**: 155–156
 Meneghini R, Mugnaini R, Packer AL (2006) International versus national oriented Brazilian scientific journals. A scientometric analysis based on SciELO and JCR-ISI databases. *Scientometrics* **69**: 529–538
 Ofori-Adjei D, Antes G, Tharyan P, Slade E, Tamber PS (2006) Have online international medical journals made local journals obsolete? *PLoS Med* **3**: e359
 Packer AL, Biojone MR, Antonio I, Takenaka RM, Garcia AP, da Silva AC, Murasaki RT, Mylek C, Reis OC, Delbucio HCRF (1998) SciELO: uma metodologia para publicação eletrônica. *Ciência Informação* **27**: 109–121
 Proctor RN (1999) *The Nazi War on Cancer*. Princeton, NJ, USA: Princeton University Press
 Snow CP (1959) *The Two Cultures and the Scientific Revolution*. Cambridge, UK: Cambridge University Press
 Watson JD, Crick FHC (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**: 737–738



Rogerio Meneghini and Abel L. Packer are at BIREME, PAHO/WHO and at the Universidade Federal de São Paulo, Brazil. E-mail: rogmene@bireme.br

doi:10.1038/sj.embor.7400906