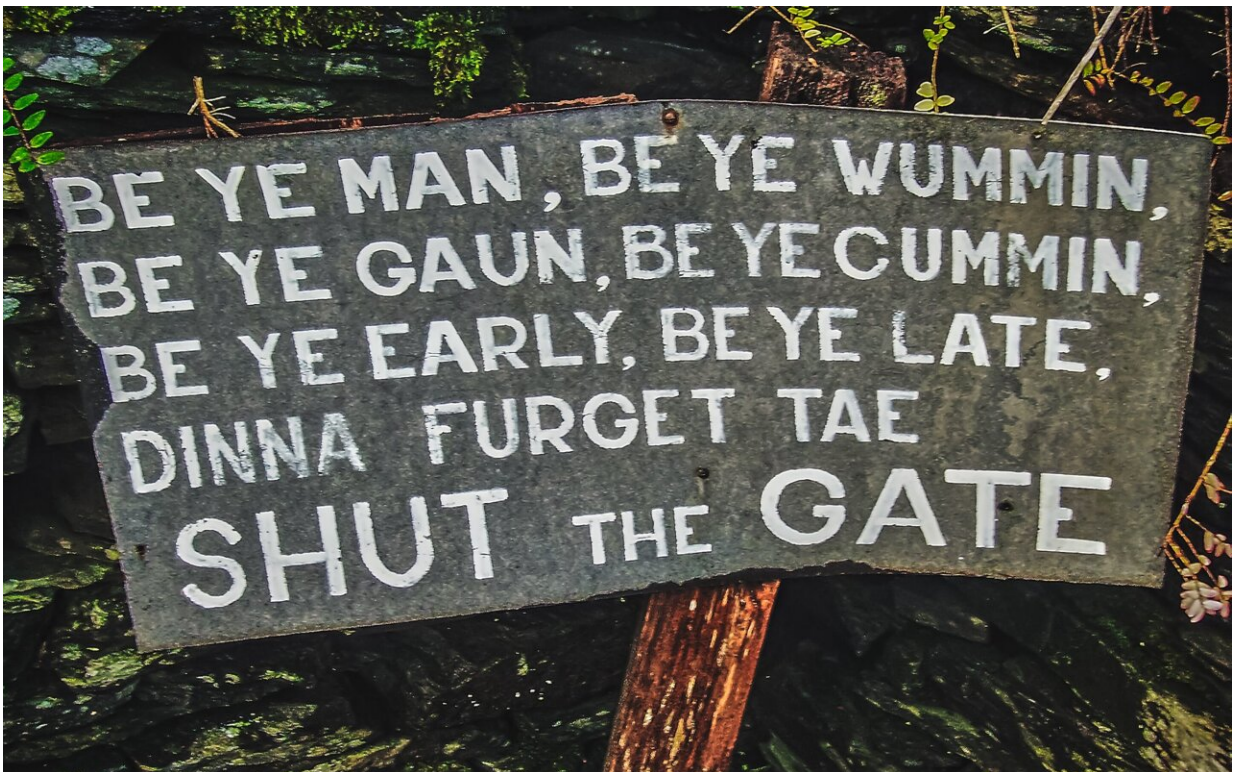


English dominates scientific research—here's how we can fix it, and why it matters

March 27 2024, by Elea Giménez Toledo



Credit: Unsplash/CC0 Public Domain

It is often remarked that Spanish should be more widely spoken or

understood in the scientific community given its number of speakers around the world, a figure the Instituto Cervantes places at [almost 600 million](#).

However, millions of speakers do not necessarily grant a language strength in academia. This has to be cultivated on a scientific, political and cultural level, with sustained efforts from many institutions and specialists.

The scientific community should communicate in as many languages as possible

By some estimates, as much as [98% of the world's scientific research is published in English](#), while only around [18% of the world's population](#) speaks it. This makes it essential to publish in other languages if we are to bring scientific research to society at large.

The value of multilingualism in science has been highlighted by numerous high profile organizations, with public declarations and statements on the matter from the [European Charter for Researchers](#), the [Helsinki Initiative on Multilingualism](#), the [Unesco Recommendation on Open Science](#), the [OPERAS Multilingualism White Paper](#), the [Latin American Forum on Research Assessment](#), the [COARA Agreement on Reforming Research Assessment](#), and the [Declaration of the 5th Meeting of Ministers and Scientific Authorities of Ibero-American Countries](#). These organizations all agree on one thing: all languages have value in scientific communication.

As [the last of these declarations](#) points out, locally, regionally and nationally relevant research is constantly being published in languages

other than English. This research has an economic, social and cultural impact on its surrounding environment, as when scientific knowledge is disseminated it filters through to non-academic professionals, thus creating a broader culture of knowledge sharing.

Greater diversity also enables fluid dialogue among academics who share the same language, or who speak and understand multiple languages. In Ibero-America, for example, Spanish and Portuguese can often be [mutually understood](#) by non-native speakers, allowing them to share the scientific stage. The same happens in Spain with the majority of its [co-official languages](#).

No hierarchies, no categories

Too often, scientific research in any language other than English is automatically seen as second tier, with little consideration for the quality of the work itself.

This harmful prejudice ignores the work of those involved, especially in the humanities and social sciences. It also profoundly undermines the global academic community's ability to share knowledge with society.

By defending and preserving multilingualism, the [scientific community](#) brings research closer to those who need it. Failing to pursue this aim means that academia cannot develop or expand its audience. We have to work carefully, systematically and consistently in every language available to us.

The logistics of strengthening linguistic diversity in science

Making a language stronger in academia is a complex process. It does

not happen spontaneously, and requires careful coordination and planning. Efforts have to come from public and private institutions, the media, and other cultural outlets, as well as from politicians, [science diplomacy](#), and researchers themselves.

Many of these elements have to work in harmony, as demonstrated by the Spanish National Research Council's work in [ES CIENCIA](#), a project which seeks to unite scientific and and political efforts.

Academic publishing and AI models: a new challenge

The global academic environment is changing as a result the digital transition and new models of open access. Research into publishers of scientific content in other languages will be essential to understanding this shift. One thing is clear though: making scientific content produced in a particular language visible and searchable online is crucial to ensuring its strength.

In the case of academic books, [the transition to open access has barely begun](#), especially in the commercial publishing sector, which releases around 80% of scientific books in Spain. As with online publishing, a clear understanding will make it possible to design policies and models that account for the different ways of disseminating scientific research, including those that communicate locally and in other languages. Greater linguistic diversity in book publishing can also allow us to properly recognize the work done by publishers in sharing research among non-English speakers.

Making publications, datasets, and other non-linguistic research results easy to find is another vital element, which requires both scientific and technical support. The same applies to expanding the corpus of scientific literature in Spanish and other languages, especially since this feeds into generative artificial intelligence models.

If linguistically diverse scientific content is not incorporated into AI systems, they will spread information that is incomplete, biased or misleading: a recent Spanish government [report on the state of Spanish and co-official languages](#) points out that 90% of the text currently fed into AI is written in English.

Deep study of terminology is essential

Research into terminology is of the utmost importance in preventing the use of improvised, imprecise language or unintelligible jargon. It can also bring huge benefits for the quality of both human and machine translations, specialized language teaching, and the indexing and organization of large volumes of documents.

Terminology work in Spanish is being carried out today thanks to the processing of large language corpuses by AI and researchers in the [TeresIA](#) project, a joint effort coordinated by the Spanish National Research Council. However, 15 years of ups and downs were needed to get such a project off the ground in Spanish.

The Basque Country, Catalonia and Galicia, on the other hand, have worked intensively and systematically on their respective languages. They have not only tackled terminology as a public language policy issue, but have also been committed to established terminology projects for a long time.

Multilingualism is a global issue

This need for broader diversity also applies to Ibero-America as a whole, where efforts are being coordinated to promote Spanish and Portuguese in academia, notably by the [Ibero-American General Secretariat](#) and the [Mexican National Council of Humanities, Sciences and Technologies](#).

While this is sorely needed, we cannot promote the region's two most widely spoken languages and also ignore its diversity of indigenous and co-official languages. These are also involved in the production of knowledge, and are a vehicle for the transfer of scientific information, as demonstrated by efforts in Spain.

Each country has its own unique role to play in promoting greater [linguistic diversity](#) in scientific communication. If this can be achieved, the strength of Iberian languages—and all languages, for that matter—in academia will not be at the mercy of well intentioned but sporadic efforts. It will, instead, be the result of the scientific community's commitment to a culture of knowledge sharing.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: English dominates scientific research—here's how we can fix it, and why it matters (2024, March 27) retrieved 27 April 2024 from <https://phys.org/news/2024-03-english-dominates-scientific.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.