The Relationship between the Language of Scientific Publication and its Impact in the Field of Public and Collective Health

dos Santos, Solange Maria ; Fraumann, Grischa; Belli, Simone; Mugnaini, Rogério

The Relationship between the Language of Scientific Publication and its Impact in the Field of Public and Collective Health

Solange Maria Dos Santos[1], Grischa Fraumann[2,4], Simone Belli[3], Rogério Mugnaini[4,*]

[1]*Scientific Electronic Library Online - SciELO / FapUNIFESP, São Paulo, BRAZIL.*
[2]*TIB Leibniz Information Centre for Science and Technology, Hannover, GERMANY.*
[3]*Complutense University of Madrid, Madrid, SPAIN.*
[4]*University of São Paulo, São Paulo, BRAZIL.*

ABSTRACT

The language of scientific publications is a crucial factor when seeking to reach an international audience, because it affects linguistic accessibility and the geographical reach of research results. English is the language of science and the fact that it can be understood by most readers represents an undeniable advantage. Moreover, the fact that a large proportion of Ibero-American research has been published in national languages, is often cited as one of the reasons for its limited exposure. The purpose of this study was to analyze the relationship between scientific output published in a native language and its degree of exposure and impact in the field of Public and Collective Health. This bibliometric study was carried out based on the scientific output data obtained from the most prolific countries that are members of the SciELO (Scientific Electronic Library Online) Network in Public and Collective Health, in the 2011-2018 period. The data was collected from the SciELO Citation Index database (SciELO CI), which was integrated into the larger WoS platform in 2014 and was chosen on account of its importance as one of the few regional indexes that is still scarcely used in studies of this nature. The data shows that Brazilian articles in Portuguese had the greatest citation impact on publications in its own language (48.7%), while its articles in English present practically the same impact (48.5%) on Portuguese publications, followed by 34.5% on Spanish publications. The impact on the national language is also significant in the case of both Mexican and Spanish publications, to whom the percentage of citing articles in Spanish, for documents cited in the same language, is higher than for documents cited in English (respectively 1.6 and 1.8). The same applies to Portuguese and US-American articles where, respectively 56.6% and 43.9% of the citing articles are in their native language. Cuban and Peruvian articles have more than 90% of their citing articles in the national language. In contrast, the USA and Brazil are countries that have a greater citation impact on other languages, especially when published in Spanish. The extent of exposure of a given language of the scientific publication varies per the country´s scientific output. In the case of Brazilian and US-American publications, including publications in the national languages of these countries, the effects on audiences in other languages can be measured by the citation impact. Furthermore, the degree of exposure of certain publications suggests that SciELO CI represents a useful database for evaluating local scientific output, and this can be observed, particularly, for publications in the national language.

Keywords: Citations, Native languages, National languages, Publications, Public and Collective Health, International visibility.

INTRODUCTION

The language of scientific publication is a key factor in terms of the extent of the international audience that is measured for research purposes, as it directly affects the linguistic accessibility and geographical scope of the research. English is the lingua franca of science and the fact that it can be understood by most is a great advantage.[1] It should also be noted that a significant part of Latin American research is mainly published in national languages and this is often mentioned as one of the reasons for the limited exposure – or as it is also commonly referred to, low visibility – of its scientific output.[2,3] Furthermore, some areas are characterized by their centrality on locally relevant research. This is reflected in the following facts: a) the large amount of scientific output devoted to issues of national interest, b) a low proportion of articles written in English, c) articles of predominantly national authorship,

d) low levels of international collaboration, and e) publishing mainly in national journals with limited circulation.[4-7] The above characteristics, which are already well known in the area of social sciences and humanities, lie at the heart of the debate about how to address the specific local features of some scientific communities, while at the same time making them meet the research requirements of an international audience. The social sciences and humanities tend not only to study, but also to serve and work together with culture and society, which means that it is better if research results are communicated in their original languages.[8] Moreover, this challenge is even greater at a time when science is opening up, and when the social repercussions of its findings are being established, particularly at a local level.[9] Likewise the social sciences and humanities, the field of Public and Collective Health has also some of these national and centralized features. Thus, they also face the moral dilemma of how to publish important research results in their mother tongue so that they can serve the local community (professionals, researchers and society in general), or either publish them in English to make them more widely accessible and reach the international audience. Public and Collective Health is an essential dimension in national public policies. It is an area that accounts for about 2% of journals and articles indexed internationally each year by the main bibliographic indexes, commonly used as a reference for measuring the performance of scientific research produced globally.

Brazil is the largest producer of articles in Public and Collective Health in Latin America, and it is among the seven largest producers in the world (fifth in the WoS and seventh in Scopus).[6] Despite that, Brazil slips to last positions in terms of citations received per article and its low performance negatively affects perceptions of the relevance of journals both internationally and nationally).[6,10]

Publishing in English is one of the most efficient aspects of internationalization. In this scenario, publication in languages other than English is mentioned as one of the aspects that contributes to making the scientific outputs of non-English speaking countries internationally invisible as an effect of the "publish (in English) or perish".[3]

In this sense, Latin American journals have made several internationalization efforts, among them, increasing the proportion of English publications, a trend also followed by journals in the Public and Collective Health field, which has made a special effort in the adoption of multilingualism (by publishing bilingual or trilingual editions)[6,10] as a tool and strategy to achieve broader internationalization and overcome linguistic barriers.

Based on the above considerations, the purpose of this study was to analyze the distribution of the language of scientific publications in Public and Collective Health from the most prolific countries in SciELO Citation Index (SciELO CI). Citation analysis was performed to measure the relationship between percentages of publication in the national language and percentage of impact from: a) articles in the national language; and b) in SciELO Citation Index, regarding the total citations received in the Web of Science platform.

LITERATURE REVIEW

As observed in previous studies,[11] the main articles published in one of the most prestigious international scientific databases (Web of Science) are in English. These scientific articles represent 95.9%, followed by articles in Spanish, with 3%. If we compare this scenario with the countries of the European Union, the documents written in English represent 93%, followed by German (2.6%), French (1.71%) and Spanish (1.4%). In the case of Latin American and Caribbean (LAC) countries, 86.0% of records are published in English, followed by Portuguese (7.4%), and Spanish (6.8%). Considering these figures, it can be stated that English is the language of collaboration between the European Union and Latin American and Caribbean (EU-LAC) countries and, while at the same time, Portuguese and Spanish maintain a relatively significant weight. This is undoubtedly on account of the important role played by Portugal and Spain in scientific collaboration with Latin America and Caribe and also because in recent years there has been an increase in the number of scientific journals in Portuguese and Spanish that are indexed in the Web of Science (WoS), especially those published by Latin American and Caribbean countries. This increase was a direct result of the expansion of WoS coverage between 2005 and 2010,[12] which in the case of Brazil led to an increase in the number of journals from 27 to 132, while Argentina, Chile and Mexico had 15, 35 and 26 indexed journals, respectively. According to Collazo-Reyes,[13] the greater presence of Brazilian journals meant that, for the first time, Portuguese became the second language, when the whole LAC output in journals in the region was taken into account.

In terms of a wider domain (SciELO and WoS), Lucio-Arias, Vélez-Cuartas and Leydesdorff[14] gathered 79,924 documents to analyze international cooperation and noted that LAC documents in Spanish and Portuguese remained the main languages for communication. According to the authors, this is one of the reasons why researchers from LAC countries may have limited participation in collaborative networks. It should be recalled what Luukkonen, Persson and Sivertsen[15] stated in the early '90s, that the stronger the scientific infrastructure of a country, the lower the level of international collaboration. For this reason, the rate of international collaboration in Haiti, Honduras, Nicaragua and the Dominican Republic is above 80%, while in the case of the more prolific Argentina, Brazil,

Chile and Mexico, it is between 30 and 40%.[16] In recent years, scientific collaboration was closely related to the countries' scientific capability, measured through different indicators. International collaboration increases impact and enhances scientific reputation.[17] For example, the availability of technologies and communication infrastructure plays a relevant role in the scientific collaboration scenario.[18] Countries with low infrastructure and development index can be considered important partners due to the availability of unique natural resources for research. International collaboration also led to an increase in citation impact, reaching almost five times the world average.[17] Collaboration of Brazilian researchers with foreign partners brings benefits for both sides, with Brazilian authors having access to financing from international agencies and foreign partners benefiting from higher impact on research.[19,20]

In a recent study, Vélez–Cuartas, Lucio–Arias and Leydesdorff[18] compared the degree of visibility of LAC publications in the WoS and SciELO to determine the dissemination of scientific knowledge produced in the region. They found out that there was a prevalence of Spanish and Portuguese, the main languages in the region, according to the geographical distribution of the collaborators. Collaborations in LAC are generally mediated by research groups from developed countries, particularly in Europe. Conversely, researchers from LAC countries play a leading role as first authors in two–thirds of multi–authored articles, i.e., LAC researchers are well integrated in the global dynamics of science. Open Access (OA) articles from Latin American countries indexed in the WoS have been found to be published mainly in English (70.4%), followed by Portuguese (17.9%) and Spanish (11.6%).[18] Other languages, such as French, Italian and German, represent only 0.2%. Compared to the distribution of languages within the total number of LAC publications indexed by the WoS, the presence of Spanish, and especially Portuguese, is significantly higher in the OA sample (11.6% and 17.9% against 7.4% and 7.4%, respectively), while, as a result, the dominance of English is low (70.4% against 85.0% in the total sample). In contrast, articles indexed in the SciELO Citation Index database (SciELO CI) are mainly published in Spanish (39.6%), followed by Portuguese (33.3%), and English (27.1%). Other languages account for only 0.03%.[21]

The data, thus, shows that the predominance of English-language articles is limited to the WoS. Within the WoS, Portuguese is the second most important language of scientific publications, while Spanish has a smaller, but still significant, importance. Hence, it is worth noting that the distribution of the language of scientific publication is completely different within SciELO CI, where Spanish dominates and represents approximately 40% of the articles, while the presence of Portuguese is more prominent than English. In general terms,

science is a global undertaking, and scientific knowledge is of global significance, that is, in theory, it should have a worldwide audience.[21]

According to Adams,[22] we are currently in the '*fourth age of research*', driven by international collaborations between elite research groups where institutions that do not form international collaborations "*risk progressive disenfranchisement, and countries that do not nurture their talent will lose out entirely*" considering that research has moved on from individual, institutional and national levels to international collaboration.[22,23] International collaboration can assist in overcoming global challenges and provide access to a suitable infrastructure and funding.[23] Moreover, international collaboration can lead to a greater research impact,[23] for example, in terms of citations.[24] Researcher mobility based on bibliographic data has also been investigated in several studies,[25,26] and international collaboration networks have been studied in depth.[27,28] These networks are also strongly supported by research funding.[29] Finally, research from one region or country can be of great importance for other geographical regions, especially in areas such as Public and Collective Health. As mentioned above, English as the lingua franca of science can reach out to an audience beyond national borders. This is essential for communicating with researchers and society abroad, although research must also be conducted in native languages, in order to have an impact and reach out to the local community.

The importance of a diverse landscape of languages that are used in research was emphasized again in the Helsinki Initiative on Multilingualism in Scholarly Communication.[30] Furthermore, bibliometrics is not neutral,[31] because the selected indicators also define the real circumstances in which the research is carried out. As far as the internationalization of research is concerned, English is often used as a means of communication. This has also been reflected in the internationalizing strategies adopted by several countries in the world,[32] and also by universities.[33] One of its objectives is to attract international talents,[32] while the international level of research also makes it an attractive profession for students.[34] It is argued that the university of the future will be international,[32] and funding Programmes have been launched to support these kind of universities, such as the *African University of Science and Technology* and the so-called *European Universities*.[35]

One means of estimating the size of the research audience that stretches across national boundaries, is to investigate the citations from a particular country.[28] This method allows us to know what kind of research is being noticed in other geographical areas, and this has also been reflected in some university classifications, for example, the *Leiden Ranking*.[36] Citations can be positive or negative, because

certain statements and results by other authors can be supported and/or criticized,[37] and how the effects of a country's publications are generalized has also been criticized.[38] International publications are also often the target of publicly funded research projects, for example in the European Union.[39] While this citation impact can be estimated for international publications, it can have the same effect on the country in which the research findings are published. The number of publications in specific languages has been investigated in several studies of countries,[34] so this study gives an insight into how publications are cited in various languages.

METHODS

A bibliometric study of an exploratory nature has been carried out by adopting a descriptive approach. A descriptive statistical analysis was conducted to characterize the sample. Simple Linear Regression was employed for the analysis of the relationship between variables.

SciELO Citation Index served as a source of data of the scientific output from countries, in the areas of Public and Collective Health. The SciELO Citation Index (SciELO CI) came into operation in January 2014, as a partnership between Thomson Reuters and the SciELO Program[1], through the Web of Science (WoS) platform. One of the objectives of this integration was to operate the indexing of SciELO journals, in particular the citation count in a wide universe of journals, including those indexed in the SciELO Network and the WoS platform.[37] The emergence of this type of product tries to compensate for one of the main deficiencies of WoS in relation to its rival Scopus, that is the low coverage of regional journals beyond the Anglo-Saxon world.

SciELO CI was accessed through the Web of Science platform, in March 2021. The following search limits were applied, resulting in 17,527 results:

Year Published: PY=(2011 OR 2012 OR … OR 2018)

Research Area: SU=”Public, Environmental and Occupational Health”

Document Type: DT=(Research–Article OR Review–Article)

The choice of these areas was based on two factors: its local character and the fact that the coverage area for SciELO CI is more extensive in Health areas, since in the case of SciELO Spain, it is restricted to the Health areas. The use of SciELO CI is important because it is a relatively new citation index, as well as being one of the few regional indexes, that are still

---

1 SciELO is an international open-access program that indexes and provides visibility to Latin American journals with important spillovers to improve the journal's quality by promoting the adoption of international standards and Open Science best practices among the region's editors.

not widely used in bibliometric studies. Furthermore, the study of the citation flow between languages is important if undertaken in a regional citation index.

To have robustness in the calculations of percentages, the production of the most prolific countries was considered, based on the author's country of affiliation, being included those who presented more than 200 articles (reducing the number of articles to 15,478). Then the articles from each country were distributed in groups according to the language of the scientific publication. Finally, the citation analysis was performed, to each group of articles (country x language), with the condition, it has at least 50 articles. Two impact indicators were proposed, to measure: a) the audience of articles from each country, in the various languages, based on the percentage of citing articles published in the national language of the country; b) the representativeness of citations counted in SciELO CI, concerning the total number of citations in all the WoS citation sources – among which SciELO CI is one. The citation context considered to each indicator is extremely different in size to each indicator, since the former is restricted to SciELO CI and the latter the WoS platform.

The procedure for obtaining the first impact indicator above, was the use of the WoS platform interface, selecting specifically the database SciELO CI. In the case of Brazil, for example, the data was obtained from the selection of its articles published in Portuguese, then accessing the link "Citation Report" to calculate the percentage of citing articles ("Citing Articles without Self-citations" link) in Portuguese. After that, the same was done for the Spanish articles from Brazilian authors, to equally calculate the percentage of citing articles in the national language (Portuguese). Lastly, the same was done for the articles in English. Successively, the procedure was repeated to each other nine countries of the sample.

The second impact indicator was calculated using the number of citations set out in each of the registers of articles, that were downloaded in 500 at a time. The percentage consisted in the division of the sum of citations received by the articles of each group (country x language) in SciELO CI (*TC-SciELO Citation Index Times Cited Count*) by the sum of the total citations (*Z9-Total Times Cited Count*[2]) received by the same group.

Finally, both impact indicators were used as dependent variables in a Simple Regression Analysis, being the percentage of publications in the national language the independent variable. We used the Microsoft Excel software to adjust the regression line and calculation of the R squared.

---

2 Web of Science Core Collection, BIOSIS Citation Index, Chinese Science Citation Database, Data Citation Index, Russian Science Citation Index, SciELO Citation Index.

We understand that some limitations of the study need to be stated. First, we have used the accumulated number of citations to each article, so we decided not to carry out a temporal analysis. In this sense, we lost the possible temporal effects that could be observed using a citation window. The main reason for this limitation is the kind of access we have to the Web of Science, through the online web interface, and not an in-house version for bibliometric purposes. Another limitation is the attribution of each article to a country independent of the number of countries collaborating. To complement this limitation, we analyzed the percentage of collaboration related to each group of articles (country x language).

RESULTS

The distribution of the production of authors from institutions in the countries that publish most in SciELO CI journals is displayed in Table 1. The publication rate in the national language is at least 79.4%, for countries that belong to the SciELO Network. Exceptions are Portugal, with 70.8% and the United States (USA), with 68.0%. It is important to note that the latter, despite being the only one that does not have its own journals in SciELO CI collection, appears as the country with the highest number of co-authored publications with authors from Latin American countries in most of the articles.

On the other hand, publications in other languages are more evident for Portugal (where 25.9% are in English), followed by Brazil (with 18,3% in English), the USA (with 17.3% in Spanish and 14,7% in Portuguese) – the US-Americans stand out as important contributors – Chile (with 17.1% in English), Mexico (with 15.7% in English) and Spain (with 13.0% in English). Brazil has been making a significant effort to internationalize its journals, which seems to be reflected in its current output. Peru and Colombia are countries whose journals publish articles mainly in the national language.

Table 1: Distribution of scientific output on Public and Collective Health, according to country and language of scientific publication – source: SciELO CI, 2011-2018.

Author´s country	Language							
	PT	SP	EN	Other	Total	%PT	%SP	%EN
Brazil	7,209	202	1,665	1	9,077	79.4	2.2	18.3
Colombia	11	1,948	148	-	2,107	0.5	92.5	7.0
Spain	36	1,336	205	-	1577	2.3	84.7	13.0
Peru	-	727	24	-	751	0.0	96.8	3.2
Cuba	3	694	46	-	743	0.4	93.4	6.2
USA	77	91	357	-	525	14.7	17.3	68.0
Mexico	7	363	69	-	439	1.6	82.7	15.7
Portugal	301	14	110	-	425	70.8	3.3	25.9
Argentina	13	349	41	-	403	3.2	86.6	10.2
Chile	3	288	60	-	351	0.9	82.1	17.1

Abbreviations: PT=Portuguese; SP=Spanish; EN=English.

The percentages in Figures 1 and 2 were only calculated in cases where the country had published at least 50 articles in that language.
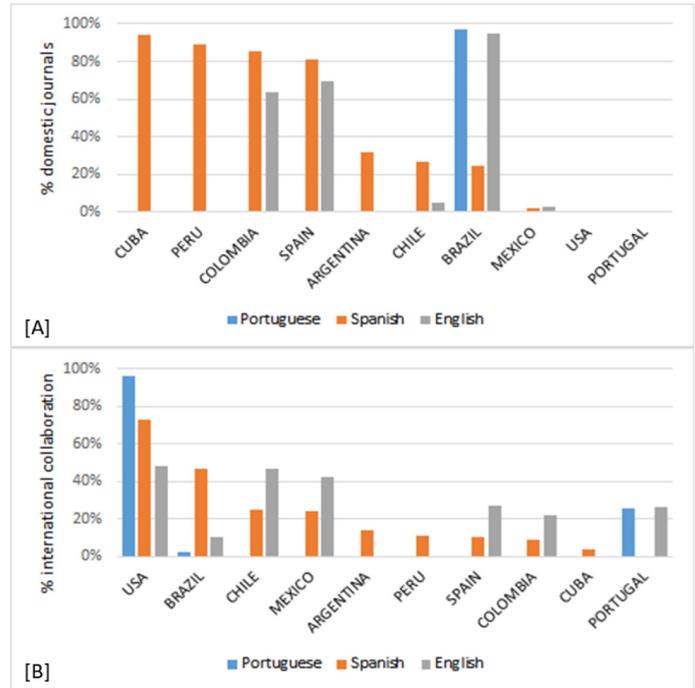

[A]

[B]

Figure 1: Comparison among group of articles (country x language), according to the percentage of articles: [A] in collaboration with another of the ten countries (IES); [B] published in domestic journals – source: SciELO CI, 2011-2018.
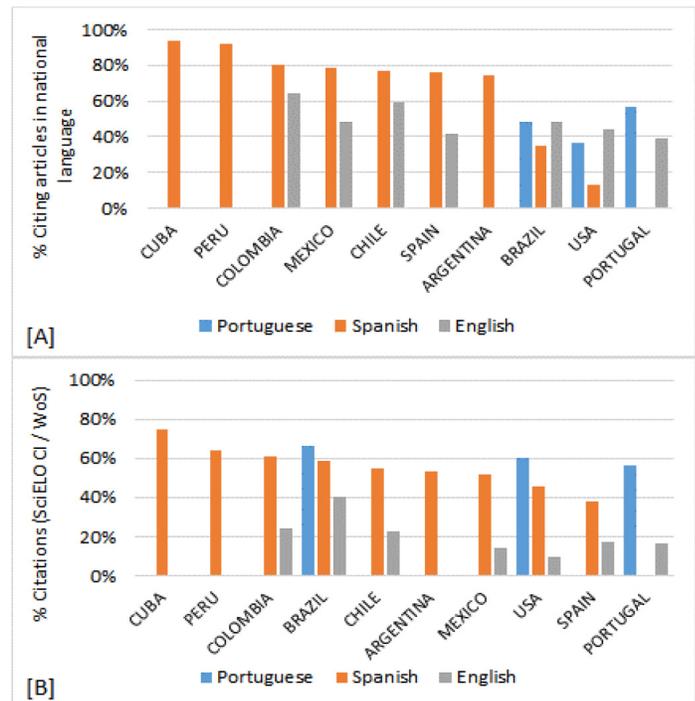

[A]

[B]

Figure 2: Impact of articles according to country and language of scientific publication, considering the percentage of: [A] citing articles in the national language with regard to all citing articles; [B] citations in SciELO CI with regard to all the WoS citation sources – source: SciELO CI, 2011-2018.

Collaboration with other countries is a factor that certainly affects the decision about the language of scientific publication of an article. Figure 1A confirms what was said earlier about the USA being the main collaborator, regardless of the language of scientific publication. Another aspect that deserves attention is that a great part of the countries uses to collaborate more in other languages, different from the national one. It is observed about all Spanish-speaking countries that have published in English, as well as to Brazil and Portugal. In the case of Brazil, the greatest rate of collaboration is observed in Spanish articles, while Portugal presents similar percentages to Portuguese and English. Additionally, it is worth highlighting that Brazil is the main collaborator of Portugal's articles in Portuguese, representing more than 20%, although for Brazil it represents less than 1%.

Finally, it is important to observe that Brazil is an important collaborator to the countries of the sample, regardless of the language of the scientific publication (data not shown), representing: 27.0% to US-American authors and 19.1% to Portuguese ones; likewise, Spain is an important contributor to Chilean authors (9.4%), while Colombia represents 9.3% of Mexican authors, collaborations.

Another aspect that deserves attention is the publication, by each country, in domestic or foreign journals. It can be an additional factor to possibly affect the decision about the language of scientific publication. Figure 1B presents a diverse scenario, revealing on the left the countries with the highest percentage of publications in domestic journals, when publishing in Spanish. Cuba is totally on the opposite side from Figure 1A, since it publishes mainly in domestic journals and does not collaborate internationally. Peru, Colombia and Spain present a similar profile, with the difference that Colombia and Spain collaborate more, as well as tend to opt for domestic journals, when publishing in English. On the right side of Figure 1B, Portugal, the USA and Mexico reveal the lowest rate, or absence of publications in domestic journals – the reason is that there are no journals on the subject in their collection, except for the USA, where the *Revista Panamericana de Salud Pública* is edited.

Finally, Brazil is the exception, presenting the highest rates of publication domestically, when publishing both in Portuguese and English. Comparing its profile with Figure 1A, we conclude that when publishing in Spanish, Brazilian authors tend to do it in collaboration, and in international journals.

Both graphs in Figure 1 reveal that in the regional context of SciELO CI the variables presented in each one are complementary: a) countries that publish domestic journals in Spanish and do not tend to collaborate (those on the left side of Figure 1B); b) other Spanish-speaking countries with few domestic journals and a moderate rate of collaboration (in the

middle of Figures 1A and 1B; Brazil, with the highest number of domestic journals, also in English, tending to collaborate especially when publishing in Spanish; and the USA and Portugal, that beside of Mexico, do not publish domestically, but present very different profiles of collaboration among the language of scientific publication.

When analyzing the citations obtained, two impact indicators allowed us to assess: a) the percentage of citing audience in the national language, to each group of articles (country x language) (Figure 2A); and b) the percentage of citations from the SciELO CI database regarding all the citations from WoS platform (Figure 2B).

Figure 2A shows that Portugal achieves the greatest impact when publishing in Portuguese (56.6%), but when publishing in English, the citing articles in its national language drop to 38.7%. This kind of observation highlights the relevance of a regional citation index, since it captures the impact received from national audiences in their own languages. Moreover, when publishing in English, the Portuguese authors show a higher audience in English, suggesting its insertion in the conversation established. When looking to a more general citation context (Figure 2B), its articles in English shows that the citations received in SciELO CI represents only 16.3% of the citations from the WoS platform, reinforcing that the impact of its articles in English reaches extra-regional boundaries.

In the case of Brazil, publishing in Portuguese or in English results in almost the same percentage of the citing articles in Portuguese (respectively, 48.7% and 48.5%), followed by 34.5% when publishing in Spanish. Finally, the last country to account its impact to Portuguese publishing was the USA (with 37.0% of citing articles in English). Following the trend of the previous, the USA shows also high impact from citing articles in its national language (43.9%), and the lowest rate of citing articles in English when publishing in Spanish.

The impact on citing articles in Spanish is prevalent in all Spanish-speaking countries, exclusively when publishing in the native language (from 74.6%, in the case of Argentina, to 93.4, presented by Cuba). Among the countries that have calculated the indicator also to English articles, Spain and Mexico showed almost half of the percentage of citing articles in Spanish when publishing in its own language, then their articles in English. While in the case of Chile and Colombia, the Spanish-speaking audience is higher (around 60%) when publishing in English. Finally, Cuba and Peru have more than 90 percent of their impact in their own language. It is clear from Figure 2A that the language of the scientific publication influences on the language of the target audience. Besides, a comparison with the graphs from Figure 1 shows that Cuba

and Peru do not have any publication and collaboration in English articles.

We now turn to Figure 2B, where the audience provided by SciELO CI is analyzed in a broader citation context. The case of Brazil is interesting, in the sense it is the only one (besides the USA) that presents similar and higher percentages of citations in SciELO CI to regional languages (Portuguese and Spanish). When publishing in English, the share of citations in other WoS databases or more prevalent, mainly in the case of the USA (more than 90% of its impact comes from WoS databases. In the case of Brazil, publishing in English still guarantees a significant impact in SciELO CI (40.0%), followed by Colombia (24.5%) and Chile (22.5%).

All the other countries have a more representative impact on SciELO CI, when they publish in their national language, that is the case of all Spanish speaking countries (more than 50%, being Spain the exception, with 38.3%) and Portugal (56.0%).

The SciELO CI domain is important for these ten countries' audience, representing about two-thirds of the citations from WoS platform, when publishing in Portuguese. Next comes the Spanish language, where the impact from SciELO CI is 52.8%. The effects of the English language are less represented in the SciELO CI domain (28.3%).

Each of the two previous variables are now analyzed together with the percentage of publications in the national language as shown in Figure 3. The linear relationship of the percentage of publications in the native language is more representative when the analysis considers the impact among citing articles in the national language (Figure 3A). This suggests that countries that publish less in their own language have a more representative audience in other languages.

The non-linearity of the relationship with the representativeness of the SciELO CI domain seems, to some extent, to be caused by the countries represented in it (Figure 3B), which are Brazil, Portugal, Spain and the USA. Brazil is the one that has more journals, and has a relatively significant impact on SciELO CI, even though it does not publish a proportional percentage in its own language. Obviously, this universe is not significant for the USA, as already mentioned. In the case of Spain, the non-representation is caused more by factors related to production than to the question of its impact. On the other hand, it should be pointed out that, although SciELO CI is a recognized important database in the analysis of the impact of the research output of the LAC countries on Public and Collective Health, they all also have some impact within the wider context of the WoS platform.
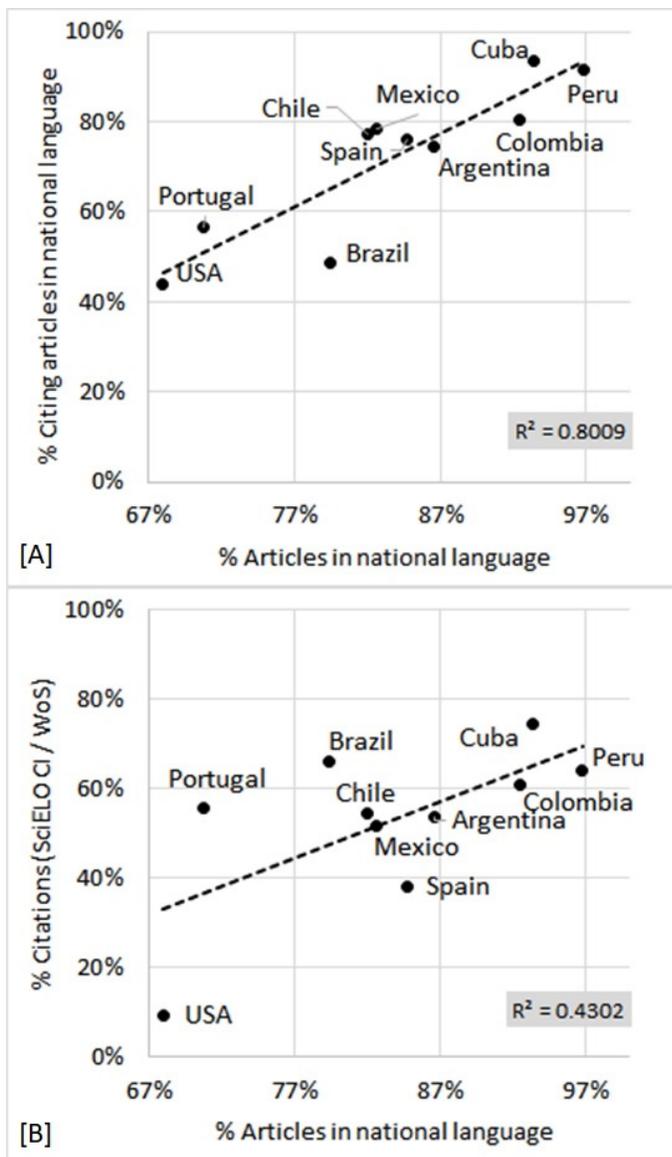

Figure 3: Simple Linear Regression for adjustments between the percentage of publications of countries in the national language and percentage of: [A] citing articles in the national language of the country with regard to all citing articles; [B] citations in SciELO CI with regard to all the WoS citation sources – source: SciELO CI, 2011-2018.

DISCUSSION AND CONCLUSION

We have explored the relationship between language and its impact to generate scientific knowledge in Public and Collective Health. We have also observed the degree of exposure of certain publications in different databases for evaluating local scientific output. The differences in language often facilitate local or national scholarly exchange of knowledge.[40,41] Previous studies observed that Brazilian scientists annually publish approximately 50,000 articles, 60% of these are in Portuguese.[2] For comparison purposes, 35% of Japanese articles are written in Japanese.[42] The case of Portuguese in Brazilian publications deserves attention, due to their greatest impact as compared to publications in

English. The case of Mexican and Spanish publications is also significant, to whom the percentage of citing articles in Spanish, for documents cited in the same language, is respectively 1.6 and 1.8 higher than for documents cited in English.

The selected language affects also the type of collaboration.[43,44] Although institutional collaboration is more important than domestic collaboration, international collaboration is even more so in the sense that it increases the citation rates far above those of domestic national collaboration.[45-47]

The creation of the SciELO Citation Index is a recent undertaking that has replaced its older global reference – the WoS Core Collection. For a long time, this endeavor was justified simply on the grounds that it could make scientific output visible and available in open access – which is a very important fact, insofar as it provides greater exposure to nationally published journals. Furthermore, one feature that adds to the credibility of the SciELO collection is its selectivity, which is the result of the stricter assessment process and criteria applied by the scientific advisory committee of each country that belongs to the SciELO Network. Unfortunately, many previous studies have shown that research published, by despite 'high-quality', nationally published journals receives fewer citations than commercially published 'high impact' journals.

After more than 20 years of SciELO Network's existence, the analysis of the scientific output impact in the Public and Collective Health area made it possible to identify the most productive countries in the region. It can be determined by the citation flow and takes into account: a) the cited and citing articles' language; and b) the representativeness of SciELO CI, regarding the set of databases in the WoS platform. This approach underlines the importance of a database such as SciELO CI, in an area being classified as research-oriented to local problems, which are better investigated by nationally affiliated researchers, whose impact is evident on regional or local communities, and whose native languages play an important role.

The analysis of the citations received from articles in native language allowed us to identify that there was a linear relationship between the percentage of publications and their impact on their native language. Brazil and the USA publish more than 90 articles in each of the languages – and beside Portugal, the least amount in their own language – presenting a more substantial citing audience when publishing in other languages. In this sense, these two countries contrast to those which publish some part of their production in a second language (English) and present impact from these publications in native language – Argentina, Chile, Colombia, Mexico, Portugal and Spain – and especially with those who practically

only publish in their own language and consequently concentrate their citing audience in the same language.

We have also noticed that in SciELO CI, Spain and the USA have the smallest part of their impact portrayed, which indicates that their citations come from outside the SciELO CI domain. On the other hand, the Latin American countries were notable for the extent of their impact in this research context.

The hegemony of English in science has imposed the use of this language in scientific communication. However, the use of native languages is fundamental to communicate the research results in some local and regional communities, such as in the case of Public and Collective Health. In recent years, there has been a growing interest from the scientific community in discussing and recommending the responsible use of indicators in evaluation , for instance, the most recent one was named the Hong Kong Principles for assessing researchers.[51] As part of the efforts of this movement, a series of documents with high-level positioning[31,47-51] were published and some of them highlight the importance of publishing in the native language.

In many parts of the world, excellence in research is associated with publication in the English language. This can be particularly problematic for areas such as Social Sciences and Humanities or Public and Collective Health, which are more oriented towards research on national or regional interests. We have seen that publishing in English and in prestige journals is considered a factor that attracts more citations and has a 'higher impact' than in any other native language.[3] Our findings at SciELO CI show that Brazil receives a considerable number of citations, even published in the native language (Figure 3B) or in Spanish (Figure 2B). This 'regional impact' should not be ignored by science policy. Understanding the dynamics of the citation and the scope of publication in the native language can assist in the definition of fairer and more responsible science policy strategies that do not make them invisible.

It is worth recognizing and rewarding the contribution of research communities with scientific output of local or regional interest, making it possible to achieve a better balance between publication in English and native languages. Considering the presence of regional citation indexes in broader platforms such as the Web of Science, it seems possible to promote visibility and inclusion of this knowledge in the global flow of scientific communication, as well as to guarantee 'linguistic accessibility' to local or regional communities.

Financing Source

ACKNOWLEDGEMENT

CONFLICT OF INTEREST

Solange Maria dos Santos is SciELO Brazil publishing coordinator.

REFERENCES

1. Nassi-Calò L. Estudo aponta que artigos em inglês atraem mais citações. São Paulo: SciELO em Perspectiva; 2016. Available from: https://blog.scielo.org/blog/2016/11/04/estudo-aponta-que-artigos-publicados-em-ingles-atraem-mais-citacoes/.

2. Meneghini R, Packer AL. 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. 2007;8(2):112-6.

3. Di Bitetti MS, Ferreras JA. Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. Ambio. 2016;46(1):121-7.

4. Fiorin JL. Internacionalização da produção científica: A publicação de trabalhos de Ciências Humanas e Sociais em periódicos internacionais. Revista Brasileira de Pós-Graduação. 2007;4(11):263-81.

5. Vargas RA, Vanz SAAS, Stumpf IRC. Brazilian agricultural research in the web of science: Bibliometric study of scientific output and collaboration (2000-2011). Em Questão. 2015;21(3):296-318.

6. Packer AL. Indicadores de centralidade nacional da pesquisa comunicada pelos periódicos de saúde coletiva editados no Brasil. Ciênc Saúde Coletiva. 2015;20(7):1983-95.

7. Gazni A. Globalization of national journals: Investigating the growth of interna¬tional authorship. Learned Publishing. 2015;28(3):195-204.

8. Sivertsen G. Balanced multilingualism in science. BiD: Textos universitaris de biblioteconomia i documentació. 2018;40.

9. Albagli S, Clinio A, Raychtock S. Ciência aberta: Correntes interpretativas e tipos de ação. Liinc em Revista. 2014;10(2):434-50.

10. Antunes JLP, Barros AJD, Minayo MCS. Caminhos da internacionalização dos periódicos de saúde coletiva. Saúde Debate. 2019;43(122):875-82.

11. Belli S, Baltà J. Stocktaking scientific publication on bi-regional collaboration between Europe 28 and Latin America and the Caribbean. Scientometrics. 2019;121(3):1447-80.

12. Testa J. The globalization of Web of Science. Philadelphia: Thomson Reuters; 2011. Available from: http://wokinfo.com/media/pdf/globalwos-essay.

13. Collazo-Reyes F. Growth of the number of indexed journals of Latin America and the Caribbean: The effect on the impact of each country. Scientometrics. 2013;89(1):197-209.

14. Lucio-Arias D, Vélez-Cuartas G, Leydesdorff L. SciELO Citation Index and Web of Science: distinctions in the visibility of regional science. 15th International Conference of the International Society of Scientometrics and Informetrics; Istanbul: Bogaziçi Universitesi; 2015;1152-60. Available from: http://issi-society.org/proceedings/issi_2015/1152.pdf.

15. Luukkonen T, Persson O, Sivertsen G. Understanding patterns of international scientific collaboration. Science, Technology and Human Values. 1992;17(1):101-26.

16. Sancho R, Morillo F, De Filippo D, Gómez I, Fernández MT. Indicadores de colaboración científica intercentros en los países de América Latina. Interciencia. 2006;31(4):284-92.

17. Leydesdorff L, Wagner C, Park HW, Adams J. International Collaboration in Science: The Global Map and the Network. arXiv. 2013. 10.3145/epi.2013.ene.12.

18. Vélez-Cuartas G, Lucio-Arias D, Leydesdorff L. Regional and global science: Publications from Latin America and the Caribbean in the SciELO Citation Index and the Web of Science. El Profesional de la Información. 2016;25(1):1699-2407.

19. McManus C, Neves AA. Production Profiles in Brazilian Science, with special attention to Social Sciences and Humanities. Scientometrics. 2020;1-23. DOI: 10.1007/s11192-020-03452-2. 21.

20. McManus C, Neves AAB, Maranhão AQ, Filho S, Santana AGJM. International

21. Minniti S, Santoro V, Belli S. Mapping the development of open access in Latin America and Caribbean countries: An analysis of web of science core collection and SciELO Citation Index (2005-2017). Scientometrics. 2018;117(3):1905-30.

22. Adams J. The fourth age of research. Nature. 2013;497(7451):557-60.

23. Hacker J. Deutsche Wissenschaft im europäischen und internationalen Kontext – Effekte? Aus Deutschland in die Welt – und umgekehrt (Deutsche Akademie der Naturforscher Leopoldina, Nationale Akademie der Wissenschaften). Pakte, nichts als Pakte – Was braucht das deutsche Wissenschaftssystem in den 2020er Jahren? Hannover, Germany: Volkswagen Stiftung. 2019. Available from: https://www.volkswagenstiftung.de/sites/default/files/downloads/Programm_Pakte.pdf

24. Kosmützky A. A two-sided medal: On the complexity of international. Comparative and collaborative team research. Higher Educ Q. 2018;72(4):314-31.

25. Katz JS, Hicks D. How much is a collaboration worth? A calibrated bibliometric model. Scientometrics. 1997;40(3):541-54.

26. Aref S, Zagheni E, West J. The demography of the peripatetic researcher: Evidence on highly mobile scholars from the web of science. Social informatics. Cham: Springer; 2019;50-65.

27. Robinson-Garcia N, Sugimoto CR, Murray DS, Yegros-Yegros A, Larivière V, Costas R. Scientific mobility indicators in practice: International mobility profiles at the country level. Profesional De la Información. 2018;27(3):511-20.

28. Wagner CS, Leydesdorff L. Network structure, self-organization, and the growth of international collaboration in science. Research Policy. 2005;34(10):1608-18.

29. Leite P, Mugnaini R, Leta J. A new indicator for international visibility: Exploring Brazilian scientific community. Scientometrics. 2011;88(1):311-9.

30. Helsinki Initiative. Helsinki Initiative on Multilingualism in Scholarly Communication. Helsinki: Federation of Finnish Learned Societies; 2019. DOI: 10.6084/m9.figshare.7887059.

31. Aufderheide E. Deutsche Wissenschaft im europäischen und internationalen Kontext – Effekte? Aus Deutschland in die Welt – und umgekehrt (Alexander von Humboldt-Stiftung). Pakte, nichts als Pakte – Was braucht das deutsche Wissenschaftssystem in den 2020er Jahren? Hannover, Germany: VolkswagenStiftung; 2019. Available From: https://www.volkswagen-stiftung.de/sites/default/files/downloads/Programm_Pakte.pdf.

32. Ruland D. Deutsche Wissenschaft im europäischen und internationalen Kontext – Effekte? Aus Deutschland in die Welt – und umgekehrt (Deutscher Akade-mischer Austauschdienst). Pakte, nichts als Pakte – Was braucht das deutsche Wissenschaftssystem in den 2020er Jahren? Hannover, Germany: VolkswagenStiftung; 2019. Available From: https://www. volkswagenstiftung.de/sites/default/files/downloads/Programm_Pakte.pdf.

33. Fraumann G, Güney I. Human resources management, internationalization and academic staff mobility. In: Pausits A, editor. Human resource management in higher education: Case studies and future scenarios. Krems an der Donau: Danube University Krems; 2015.

34. Kulczycki E. Insights regarding social sciences and humanities obtained from studies of national bibliographic databases. Poznań: 3rd ENRESSH Training School; 2019. Available from: https://enressh.eu/wp-content/uploads/2019/06/10-Emanuel-Kulczycki-%E2%80%93-SSH.pdf.

35. Education, Audiovisual and Culture Executive Agency. European universities. Brussels: Education, Audiovisual and Culture Executive Agency; 2018. Available from: https://eacea.ec.europa.eu/erasmus-plus/actions/key-action-2-european-universities_ en.

36. Waltman L, Calero-Medina C, Kosten J, Noyons ECM, Tijssen RJW, van Eck NJ, et al. The Leiden ranking 2011/2012: Data collection, indicators, and interpretation. Journal of the Association for Information Science and Technology. 2012;63(12):2419-32.

37. Small H. On the shoulders of Robert Merton: Towards a normative theory of citation. Scientometrics. 2004;60(1):71-9.

38. Sørensen MP, Schneider JW. Studies of national research performance: A case of 'methodological nationalism' and 'zombie science'?. Science and Public Policy. 2017;44(1):132-45.

39. Wagner CS, Park HW, Leydesdorff L. The Continuing Growth of Global Cooperation Networks in Research: A Conundrum for National Governments. PLoS One. 2015;10(7):e0131816.

40. Packer AL. SciELO Citation Index no Web of Science. São Paulo: SciELO in Perspectiva; 2014. Available from: https://blog.scielo.org/blog/2014/02/28/scie-locitation-index-no-web-of-science/.

41. Salager-Meyer F. Writing and publishing in peripheral scholarly journals: How to enhance the global influence of multilingual scholars?. Journal of English for Academic Purposes. 2014;13:78-82.

42. Chikazawa Y, Katsurai M, Ohmukai I. Multilingual author matching across different academic databases: A case study on KAKEN, DBLP, and PubMed. Scientometrics. 2021;1-17.

43. Amano T, González-Varo JP, Sutherland WJ. Languages are still a major barrier to global science. PLoS Biol. 2016;14(12):e2000933.

44. Leimu R, Koricheva J. Does scientific collaboration increase the impact of

ecological articles? Bio Science, 2005;(55): 438-43.

45. Katz J, Hicks D. How much is a collaboration worth? A calibrated bibliometric model. Scientometrics. 1997;40(3):541-54.

46. Barrantes LBS, Bote GVP, Rodríguez ZC, De Anegón MF. Citation flows in the zones of influence of scientific collaborations. Journal of the American Society for Information Science and Technology. 2021;63(3):481-9.

47. McManus C, Neves AAB, Maranhão AQ, Filho SAG, Santana JM. International collaboration in Brazilian science: Financing and impact. Scientometrics. 2020;125(3):2745-72.

48. DORA. San Francisco Declaration on Research Assessment. San Francisco: DORA; 2012. Available from: https://sfdora.org/read/.

49. Hicks D, Wouters P, Waltman L, De Rijcke S, Rafols I. Bibliometrics: The Leiden Manifesto for research metrics. Nature. 2015;520(7548):429-31.

50. Wilsdon J, Allen L, Belfiore E, Campbell P, Curry S, Hill S, *et al*. The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. Bristol: Higher Education Funding Council for England; 2015. DOI 10.13140/RG.2.1.4929.1363.

51. Moher D, Bouter L, Kleinert S, Glasziou P, Sham MH, Barbour V, et al. The Hong Kong Principles for assessing researchers: Fostering research integrity. PLoS Biol. 2020;18(7):e3000737.