https://journals.scientia.international/SIJLLA

Scientia International Journal for Linguistics, Letters and Arts

Vol. 1, 1 (2026)
Type: Research Article | DOI: 10.56365/ws7maa38

Language barriers and the circulation of science: Global evidence and editorial strategies for Latin American scholarly output†

Jéssica Biltches Costa[1*], Bruno César Alves Marcelino[2]

[1]Master's student in Language and Literature Studies; teaching degree in Portuguese and Spanish Language and Literature; Universidade Federal de São Paulo, São Paulo, Brazil; ORCID: https://orcid.org/0009-0002-9781-4141

[2]Doctoral student in Interdisciplinary Social Studies of Europe and Latin America; master's degree in Contemporary Latin American Integration; Scientia International, Murcia, Spain; ORCID: https://orcid.org/0000-0002-0638-7889

* Corresponding author: jessybilt4@gmail.com

Abstract

The consolidation of English as the lingua franca of science has transformed a means of communication into a selection mechanism that unequally affects researchers whose first language is not English. This article examines how language barriers contribute to what the literature calls scientific invisibility and discusses editorial strategies capable of expanding the circulation of Latin American scholarly output. Through a critical literature review, the study brings together two distinct bodies of evidence: global studies on the disadvantages faced by non-English-speaking researchers and bibliometric studies of the Latin American publishing circuit. The former indicate that non-native English-speaking researchers spend more time reading and writing in English and report language-related rejections and revision requests at frequencies between 2.5 and 12.5 times those of native speakers; among Colombian doctoral students, the cost of translation services per article amounts to between one-quarter and one-half of a doctoral student's monthly salary. The latter show that 91% of articles in Scopus and 96% of those in the Web of Science Core Collection are in English, whereas the regional corpus indexed in SciELO and Redalyc comprises 43.7% of articles in Spanish, 32.1% in Portuguese, and 23.9% in English. The article discusses the debate on linguistic injustice, the epistemic dimension of invisibility, and the regional open access infrastructure. It argues that shifting the burden of translation from individual researchers to the editorial infrastructure is a necessary, although insufficient, condition for changing this situation.

Keywords: language barriers; circulation of science; Latin America; open access; multilingualism; linguistic injustice.

Article details | Open Peer Review
Edited by:
William Jônatas Vidal Coutinho
Reviewed by:
Fabio do Vale
William Jônatas Vidal Coutinho
Citation:
Costa, J. B., & Marcelino, B. C. A. (2026). Language barriers and the circulation of science: Global evidence and editorial strategies for Latin American scholarly output. *Scientia International Journal for Linguistics, Letters and Arts, 1*(1). https://doi.org/10.56365/ws7maa38

Article history
Received: 28 June 2026
Revised: 3 September 2026
Accepted: 3 September 2026
Published: 15 September 2026

† This article originated from the research period undertaken by Jéssica Biltches Costa at Scientia International (Murcia, Spain), as part of the project "Implementation of Scientia International's Editorial Line for the Latin American and Portuguese-Speaking Market." The work was supervised by Bruno César Alves Marcelino in 2024.

ISSNe 3127-6466 © 2026 The Authors. Published by Scientia.International S.L. (Spain).
This is an open access publication under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Contemporary science presupposes communication in English as an almost tacit condition of its existence. Throughout the twentieth century, international scientific communication shifted from a multilingual regime, in which French, German, Russian, and Spanish coexisted, to a position of English preeminence unprecedented in the modern history of science (Hamel, 2007). According to Hamel (2007), this is the first time that a single language has occupied such a dominant position in global scientific communication. This shift is not neutral: by becoming a condition for access to the so-called mainstream of science, language ceases to function solely as a vehicle of communication and also becomes a filter, distributing opportunities for circulation and recognition unequally among researchers worldwide.

For Latin American researchers, whose first language is predominantly Portuguese or Spanish, this condition imposes a contradictory dual movement. On the one hand, publishing in English increases the likelihood of indexing in dominant databases and, with it, of citation within the circuits those databases measure. On the other hand, work written in a national language—often more relevant to local problems and regional audiences—tends to remain beyond the reach of global indexing and assessment systems. This paradox, in which visibility in one circuit is obtained at the expense of visibility in the other, lies at the heart of the problem this article seeks to examine.

In this study, the notion of scientific invisibility refers to the set of mechanisms through which a body of knowledge fails to circulate, be cited, and influence the research agenda for reasons unrelated to its intrinsic quality—among which the language barrier occupies a central position. Distinguishing the quality of research output from its circulation is therefore a prerequisite for the argument. The hypothesis explored here is that the more limited international circulation of Latin American science is, to a significant extent, an effect of asymmetries in the global publishing and indexing system rather than a direct indicator of lesser scientific merit.

Because the vocabulary of this debate is often used interchangeably, the terms employed below should be defined. Indexing means the inclusion of a journal or document in a particular bibliographic database according to that database's own criteria; reach refers to the breadth of audiences able to read a text, as conditioned by language and access; circulation is the actual movement of the text among those audiences, observable through downloads, republication, and uses; citation is the formal record of such use in the literature indexed in a specific database; visibility is the likelihood that a work will be located and retrieved within a given circuit; and impact is any aggregate measure derived from the preceding dimensions. Two consequences follow from this distinction. First, citation and visibility are always relative to the database that measures them: an article rarely cited in the Web of Science may circulate extensively within the regional circuit, and vice versa. Second, the impact factor is a journal-level indicator, not an article- or author-level indicator, and cannot automatically be interpreted as a measure of individual merit—a warning repeatedly

emphasized in the Latin American bibliometric literature itself (Salatino, 2022). For this reason, this article addresses circulation and visibility rather than scientific impact in the strict sense, which this review does not measure.

It should be noted at the outset that the thesis advanced here is not universally accepted. The literature contains an active debate on the existence and extent of so-called linguistic injustice, to which this article will devote specific attention, in order to avoid oversimplifying a problem that is simultaneously linguistic, economic, epistemic, and political. It should also be noted that the hypothesis formulated above does not posit a causal opposition between quality and structural asymmetries or assign a specific causal weight to either: the review does not directly measure the scientific quality of the outputs being compared, and what can be supported is the existence of documented barriers whose effect on circulation is independent of the merit of the works concerned. The argument will be built on quantitative empirical evidence rather than intuitions, precisely because the controversy demands rigor.

The article comprises six sections in addition to this introduction. The methodology explains how the corpus was constructed. The section on the hegemony of English discusses the measurable linguistic costs it imposes on non-English-speaking researchers. The section on linguistic injustice examines the theoretical debate and its epistemic dimension. The section on the circulation of Latin American output addresses the regional open access infrastructure and the tension between internationalization and multilingualism. The section on editorial strategies analyzes structural multilingualism, artificial intelligence-assisted translation, and the legal and technical conditions for circulation. The concluding remarks synthesize the argument and outline its implications.

2. Methodology

This study is a critical literature review, qualitative in nature and analytical-interpretive in purpose. Unlike a systematic review, which seeks to map a field exhaustively according to fixed search protocols and addresses a narrowly defined, replicable question, a critical review selects and brings together key works to construct a theoretical argument and identify tensions in the state of the art (Grant & Booth, 2009). This choice has two consequences that should be made explicit. First, the findings presented here make no claim to exhaustiveness and do not support statistical inference about the literature in the field. Second, the process must nevertheless remain traceable: this section describes the sources consulted, search terms, time frame, inclusion and exclusion criteria, and how the analytical axes were constructed, enabling readers to assess—and, where appropriate, challenge—the composition of the corpus.

The search was conducted between January and May 2026 across five complementary sources. Two mainstream databases, Scopus and Web of Science (Core Collection), were used to locate international literature on language barriers and scientific communication. Two regional full-text open access databases,

SciELO and Redalyc, were used to locate Latin American literature on publishing circuits, indexing, and bibliodiversity—precisely because, as argued in Section 5, this output is underrepresented in the first two databases. Google Scholar was used as a complementary source to retrieve gray literature and works not indexed in the preceding databases. These searches were supplemented by targeted searches in preprint repositories (arXiv, SciELO Preprints) and institutional websites (SciELO, Redalyc, LA Referencia, Crossref, Helsinki Initiative), necessary to document editorial standards, indexing policies, and technical specifications that do not circulate in article form.

Search terms were combined in three languages, reflecting the nature of the subject. In English: language barrier, linguistic injustice, non-native English speaker, scientific publishing, multilingualism, scholarly communication, machine translation. In Spanish: barreras lingüísticas, hegemonía del inglés, circuitos de publicación, multilingüismo, acceso abierto, bibliodiversidad. In Portuguese: barreiras linguísticas, invisibilidade científica, comunicação científica, acesso aberto, multilinguismo. Terms were combined using Boolean operators and the qualifiers América Latina, Latin America, and latinoamericano when the aim was to isolate regional literature. The preferred time frame was 2007–2026, taking Hamel's (2007) systematic account of the dominance of English in international periodical literature as the starting point; earlier works were included when their status as key references in the field justified doing so.

The inclusion criteria were: (a) directly addressing the relationship between language, scientific publishing, and inequalities in circulation; (b) presenting primary data, a bibliometric analysis of a defined corpus, an experimental or quasi-experimental design, or constituting an authoritative institutional policy or technical document; and (c) being published in Portuguese, Spanish, or English. The exclusion criteria were: opinion essays without an empirical basis or original conceptual development; studies restricted to teaching English for academic purposes without connecting it to the publishing system; and duplicates across databases. Preprints were included but identified as such in the text so that readers could weigh the status of the evidence.

Screening proceeded in three stages. Approximately 340 records were identified in the searches; after duplicate removal and the screening of titles and abstracts, 96 were retained for examination; of these, 41 were read in full, and 21 constitute the corpus analyzed in depth and cited in this article. This substantial reduction results from criterion (b): most of the retrieved literature consists of programmatic essays that reiterate the diagnosis without adding evidence, or single-journal case studies whose generalization would be unwarranted. The final corpus combines established key works—Hamel (2007), Lillis and Curry (2010), and Santos (2014)—with recent large-scale empirical studies, including Amano et al. (2023), Beigel et al. (2024a, 2024b), and Salatino (2022).

The four analytical axes structuring the subsequent analytical sections were not defined a priori but emerged from the thematic grouping of the 41 works read in full, according to the question each addressed: (a)

the measurement of individual linguistic costs; (b) the theoretical debate on the nature and extent of disadvantage; (c) the actual linguistic composition of publishing circuits and the infrastructure that sustains them; and (d) editorial responses proposed or already in operation. Works that did not fit any of these four axes—especially those concerning national research assessment policies—were excluded from the final corpus, although they inform the background to the discussion.

Two synthesis decisions warrant mention. The first concerns the integration of quantitative evidence into a qualitative review: numerical data are reported with the qualifications of the original study—population surveyed, country, field, period, and nature of the measure—and are not pooled across studies because the designs are not commensurable. The second concerns the treatment of contrary evidence: Hyland's (2016) critical position, which questions the very existence of linguistic injustice, was deliberately incorporated into the corpus, as were the null findings of Di Bitetti and Ferreras (2017) regarding differences in quality between articles published in English and Spanish in the same journals. The inclusion of these works is not merely decorative: they limit the scope of the conclusions that Section 7 can support.

Three limitations must be acknowledged. First, the corpus is selective rather than exhaustive, with the implications already noted. Second, a substantial share of the available quantitative evidence on individual costs comes from the natural and environmental sciences—Amano et al.'s (2023) sample, for example, is restricted to environmental scientists—which calls for caution when generalizing to the humanities and social sciences, whose publishing patterns and languages of circulation differ. Third, and crucial to the scope of this article, the evidence on individual costs is predominantly global and includes only two substantial Latin American cases—Bolivia, in Amano et al.'s (2023) sample, and Colombia, in Ramírez-Castañeda (2020)—whereas the specifically regional evidence is bibliometric and concerns circuits rather than individuals. For this reason, throughout the text the article systematically distinguishes between global evidence on non-English-speaking researchers and evidence on the Latin American publishing circuit, avoiding the transfer of conclusions from one level to the other.

3. The hegemony of English and the measurable costs of scientific publishing

3.1. The consolidation of a lingua franca

The predominance of English in international scientific literature is a historically recent and statistically substantial phenomenon. According to Hamel (2007), by the beginning of the century, more than 75% of articles in the social sciences and humanities and more than 90% in the natural sciences published in international journals were already written in English. The author demonstrates that this shift entailed the decline of French, German, Russian, Spanish, and Japanese as languages of science, leading many observers to conclude that publishing in English—if not exclusively in English—was inevitable.

Subsequent surveys confirm the magnitude of the phenomenon, provided that the universe to which each percentage refers is made explicit. Drawing on data collected in September 2022 from the official Scopus and Clarivate websites, Salatino (2022) reports that 90.8% of the 62.4 million articles in Scopus are in English, compared with 0.9% in Spanish and 0.4% in Portuguese, and that the concentration is even greater in the Web of Science Core Collection, where 96% of the 34.4 million articles are in English, compared with 0.75% in Spanish and 0.32% in Portuguese. Drawing on Hamel (2007) and earlier surveys, Di Bitetti and Ferreras (2017) estimate that more than 90% of indexed articles in the natural sciences are in English. It is worth noting that the figure of 98% frequently cited in the literature—including in Ramírez-Castañeda's (2020) initial framing—is not accompanied by a database, field, or period; this is why the present article instead uses the percentages documented by Salatino (2022), whose source, date, and universe are explicit.

These figures must be qualified to avoid turning data about databases into data about world science. The percentages above describe the linguistic composition of Scopus and Web of Science, not the totality of what is produced and published worldwide. As will be seen in the discussion of regional circulation, there is a vast publishing circuit in other languages that these databases do not capture—itself one aspect of the problem examined here. The decisive point, however, is not statistical prevalence in itself, but the fact that research assessment regimes use these databases as benchmarks, turning linguistic competence into a condition for access to the institutionally rewarded circuit (Salatino, 2022; Ramírez-Castañeda, 2020).

3.2. The unequal burden of second-language writing

The specialist literature consistently documents that writing science in a second language imposes an additional burden that falls unequally on non-native speakers. In a comparative study involving 148 Mexican and 236 Taiwanese researchers, Hanauer, Sheridan, and Englander (2019) quantified this burden using a validated instrument for measuring writing effort: compared with writing in the first language, scientific writing in English entailed average increases of 24% in perceived difficulty, 10% in dissatisfaction, and 22% in anxiety. This is a self-reported measure of perception rather than observed performance; its contribution lies in replacing anecdotal impressions with an instrument that allows comparisons across groups.

The broadest and most influential recent study is that of Amano et al. (2023), published in PLOS Biology, which administered a questionnaire between June and October 2021 to 908 environmental science researchers of eight nationalities, stratified by national English proficiency and country income level: Bangladeshi (n = 106), Bolivian (100), British (112), Japanese (294), Nepali (82), Nigerian (40), Spanish (108), and Ukrainian (66). The authors summarize the problem as follows:

> English serves as a convenient common language for science. However, this practice imposes insurmountable barriers on those whose first language is not English—the majority of people in the world. (Amano et al., 2023, p. 1, authors' translation in the Portuguese original)

The quantitative findings warrant careful elaboration because their common interpretation in the secondary literature often overstates them. Regarding time, among researchers who had published only one article in English, those from low-proficiency countries spent a median of 90.8% more time reading an article in English, and those from moderate-proficiency countries spent 46.6% more time; the same comparison in the first language reveals no disadvantage, locating the effect in language rather than reading ability. Regarding language editing, non-native speakers ask someone to check their English in 75% or more of their texts, whereas most native speakers do so in fewer than half. Regarding the editorial process, the most frequently cited figures require precision: among respondents with one published article, 38.1% of non-native speakers from moderate-proficiency countries and 35.9% of those from low-proficiency countries reported having had an article rejected because of its English writing, compared with 14.4% of native speakers—hence a 2.5- to 2.6-fold higher frequency of language-related rejection. Meanwhile, 42.5% and 42.6% of these two groups, compared with only 3.4% of native speakers, reported being asked frequently, most of the time, or always to improve their English writing during review—hence the 12.5-fold higher frequency. In both cases, these are frequencies of self-reported experiences in a specific sample, not the probability of any given article being rejected or returned, an interpretation the data themselves do not support.

The authors further observe that a substantial proportion of early-career researchers refrain from attending or presenting at international conferences because of the language barrier, and argue that, all else being equal, the accumulation of these costs would tend to reduce the apparent scientific productivity of non-native speakers. It should be kept in mind that this last proposition is an inference the authors draw from the combined costs measured, not a survey finding. For the purposes of this article, what the sample does and does not allow us to say about Latin America also matters: the only Latin American nationality represented is Bolivian, which calls for treating these findings as global evidence on non-English-speaking researchers rather than as a regional portrait.

As their central recommendation, Amano et al. (2023) argue that journals, as gatekeepers of the system, should assume part of the effort by offering free language-editing support and actively promoting multilingualism in science—a proposition that anchors the discussion of editorial strategies below.

3.3. The barrier as an economic and class filter

Language barriers are not merely distributed unequally between native and non-native speakers; they overlap with socioeconomic inequalities, becoming a class filter. In a study of Colombian doctoral students in the biological sciences, Ramírez-Castañeda (2020) reports that more than 90% of articles published by Colombian researchers were in English and that 43.5% of respondents reported at least one rejection or revision request because of English grammar; 33% stated that they had decided not to participate in international meetings because English was mandatory for oral presentations. The financial cost is even more revealing: among the six translation and editing service providers analyzed by the author, the price per article

ranged from one-quarter to one-half of a doctoral student's monthly salary in Colombia, using the country's average doctoral salary as a benchmark (approximately US$947 at the time of the survey). This finding is specific to that national context, field, and pricing structure and should not be generalized to all Latin American researchers—whose salary conditions and access to editorial services vary considerably across countries.

This creates what might be called a pay-to-pass filter: in the context studied, only researchers with sufficient resources to pay for professional editing services can overcome the linguistic obstacle on competitive terms. Ramírez-Castañeda (2020) additionally identifies a positive correlation between English proficiency and higher socioeconomic background, suggesting that the language barrier is not only a consequence of, but also a mechanism for reproducing, preexisting inequalities. In this setting, language and economic capital operate jointly to exclude researchers from less well-resourced contexts—and to exclude them even before their work is assessed on its merits. It should be noted that the identified correlation is associative: the study design does not establish a causal ordering between proficiency and social background.

4. Linguistic injustice, theoretical controversy, and the epistemic dimension

4.1. The debate over the existence of linguistic injustice

It would be a mistake to present the language-barrier thesis as uncontroversial. The controversy has a clear landmark: in a provocatively titled article on the myth of linguistic injustice, Hyland (2016) argues that the notion of disadvantage among authors who write in English as an additional language has acquired the status of an unquestioned orthodoxy. For Hyland, academic English is no one's mother tongue and requires deliberate learning by native and non-native speakers alike; attributing publishing difficulties exclusively to non-native-speaker status would therefore oversimplify a phenomenon whose more decisive causes are the physical, academic, and financial isolation of researchers on the periphery.

Hyland's (2016) critique has the merit of demanding empirical evidence rather than introspection and of warning against the risk that automatically invoking the non-native label may turn it into a disabling stereotype. However, the response by Politzer-Ahles et al. (2016) argues that this position underestimates the role of linguistic privilege. The authors observe that the fact that many non-native speakers publish successfully is not evidence against disadvantage: only a caricatured understanding of linguistic injustice would maintain that articles by non-native speakers are rarely published. The disadvantage manifests itself not in an absolute inability to publish but in the additional effort, time, cost, and differential frequency of rejection and revision requests—precisely the variables that subsequent studies by Hanauer et al. (2019) and Amano et al. (2023) would measure.

In light of this evidence, the most defensible position is neither that language explains everything nor that it explains nothing, but that the language barrier constitutes one of several dimensions of disadvantage—

measurable and real—intertwined with economic and structural factors that the available designs do not allow us to separate fully. This is the nuanced position adopted in this article.

4.2. From individual disadvantage to epistemic injustice

Beyond the individual level of writing, the language barrier is frequently associated with effects at the collective level of knowledge. The literature on epistemic injustice and the coloniality of knowledge argues that the hegemony of a language not only hampers communication but also establishes hierarchies of what counts as legitimate knowledge. In formulating the notion of epistemicide, Santos (2014) describes the process through which forms of knowledge produced outside the hegemonic center are systematically disqualified, ignored, or erased. Language would be one of the vectors of this process: by defining English as the language of the most rewarded circuit, the system would relegate to the periphery not only researchers but also ways of producing and framing problems that do not circulate in the dominant languages.

The status of this argument must, however, be clearly identified. Moving from measurable individual disadvantage—time, cost, frequency of rejection—to collective claims about hierarchies of knowledge, the imposition of the center's topics and methods, and disconnection from local priorities is an interpretive move, not an empirical demonstration. The studies discussed above measure effort and the frequency of editorial events; they do not measure the formation of research agendas or the erasure of intellectual traditions. The notion of epistemicide therefore operates here as a theoretical interpretive lens that makes the findings intelligible, not as a result deduced from them. There are nevertheless empirical indications consistent with this interpretation, and Salatino (2022) provides the most direct: according to the author, the historical constitution of the regional space of scientific communication in Spanish and Portuguese is not a matter of translation or learning English but the outcome of intellectual and sociocultural roots extending back more than two hundred years—meaning that linguistic assimilation into the hegemonic circuit would entail costs not reducible to technical matters.

This distinction helps explain why invisibility cannot be resolved through translation alone. When output from the periphery must conform to the topics, methods, and frames of reference valued at the center in order to be accepted, translation provides access without changing the criteria of valuation. Acknowledging this is important both to avoid reducing the solution to a technical operation that ignores the political and epistemic dimensions of the problem and, conversely, to avoid turning a plausible theoretical interpretation into an empirical finding that the available literature does not yet provide.

5. The circulation of Latin American output and the regional infrastructure

5.1. Knowledge that does not circulate

The counterpart of the hegemony of English is the more limited circulation of knowledge produced in other languages, and this asymmetry has material effects on science itself. In a survey conducted through Google Scholar in sixteen languages, Amano, González-Varo, and Sutherland (2016) found that 35.6% of the 75,513 scientific documents on biodiversity conservation published in 2014 were not in English—with Spanish (12.6%) and Portuguese (10.3%) the most frequent non-English languages, followed by Simplified Chinese (6%) and French (3%). The authors demonstrate that disregarding this non-English-language knowledge creates biases in the understanding of the subjects of study themselves and hampers the use of science by local managers and policymakers: 54% of protected-area directors in Spain identified language as a barrier to using the latest science.

Two qualifications are necessary before applying this finding to the regional argument. First, the survey identifies documents by publication language, not geographical origin: texts in Spanish and Portuguese do not automatically equate to Latin American output, since Spain and Portugal belong to the same linguistic circuits. Second, the finding on managers refers specifically to Spain, not Latin America; its relevance here is to illustrate the mechanism—the unavailability of science in the language of those who would apply it—rather than to quantify its regional incidence. With these qualifications, the finding remains fundamental to the thesis of this article because it reveals that the asymmetry is bidirectional. It is not only that science produced in Spanish and Portuguese struggles to reach international audiences; science produced in English also fails to reach local audiences that could benefit from it. The language barrier fragments the global knowledge system into circuits that communicate poorly with one another, to the detriment of all sides.

5.2. The linguistic circuits of regional publishing

Analysis of indexed Latin American output confirms the existence of distinct linguistic circuits and makes it possible to gauge their scale. The OLIVA project, conducted by Beigel, Packer, Gallardo, and Salatino (2024a), consolidated a bibliographic metadata database from the complete SciELO and Redalyc collections, provided by the platforms themselves and subjected to data cleaning and overlap detection among journals shared by both; document-level data were collected through June 2019, and primary metadata collection for each journal took place in May 2021. The corpus comprises 1,720 journals and 908,982 documents published between 1909 and May 2019, of which 790,304 are articles, published in fifteen countries (Salatino, 2022). Its disciplinary composition is dominated by the social sciences (33%), followed by the medical and health sciences (19%), the humanities (12%), and the natural and exact sciences (10%).

The linguistic composition of this corpus is central to the argument. According to Salatino (2022), 43.7% of articles are published in Spanish, 32.1% in Portuguese, 23.9% in English, and 0.2% in French, while 0.12%

have no language information—with substantial intraregional variation: Brazil has the highest proportion of output in English (33.9% of its total publications), and Argentina the lowest (10%). By cross-tabulating publishing institutions, languages, countries of publication, disciplinary fields, and indexing, the author identifies at least four regional linguistic circuits: one that has migrated almost entirely to English to reach the mainstream (39 journals, rooted in medicine and the biological and agricultural sciences); one strongly oriented toward Spanish (211 journals); one predominantly in Portuguese (12 journals); and, the largest and with the highest volume of articles per journal, a multilingual circuit comprising more than one thousand journals publishing in at least three languages. The regional circuit is therefore neither monolingual nor uniformly resistant to English: it is multilingual, and this is what distinguishes it from the monolingualism of the mainstream.

The contrast with dominant databases can be measured at two levels. At the article level, 90.8% of documents in Scopus and 96% of those in the Web of Science Core Collection are in English, compared with 0.9% and 0.75% in Spanish and 0.4% and 0.32% in Portuguese, respectively (Salatino, 2022). At the journal-title level, the overlap between the two universes is quantifiable and is the most direct indicator that points of contact are limited: of the 27,339 journals indexed in Scopus, 932 are Latin American (3.4%); of the 24,769 in the Web of Science Core Collection, 1,287 are regional (5.2%); and of the 13,802 journals with an Impact Factor in the Journal Citation Reports, only 270 are from the region (0.2%), originating from just nine countries (Salatino, 2022). These figures should be read against the regional universe actually indexed: the primary survey by Beigel et al. (2024b) counted 4,077 active journals published and indexed in Latin America in May 2023, distributed across 22 countries. In other words, between one-quarter and one-third of the indexed regional universe reaches the two major international databases, and a much smaller fraction reaches the subset of those databases with an Impact Factor.

This establishes a hypercentrality of English that operates not only at the level of individual writing but also within the very architecture of the indexing systems that define what counts as visible science in each circuit. In light of the data above, it is legitimate to speak of two partially overlapping circuits: one regional, vast, multilingual, and predominantly diamond open access; the other international, English-language, and more institutionally rewarded. Their separation should not, however, be overstated. The available data measure indexing overlap and linguistic composition; they do not measure cross-citation or the actual geographical reach of texts. Claiming that the two circuits hardly communicate would exceed the evidence; what it supports is that communication between them is asymmetrical and institutionally discouraged.

5.3. Open access infrastructure as a historical response

The Latin American response to this situation predates, in many respects, the international vocabulary of open access itself. SciELO (Scientific Electronic Library Online) was proposed in 1996, developed in 1997 as a pilot project with ten journals, and launched in 1998 with support from the São Paulo Research Foundation

(FAPESP) and in collaboration with BIREME/PAHO/WHO—four years before the Budapest Declaration would formalize open access. Redalyc, in turn, was established in 2002 at the Universidad Autónoma del Estado de México and opened its portal to the public in 2003. Alongside Latindex and Biblat, the two platforms constitute a space for validating editorial quality grounded in good practices and peer review (Beigel et al., 2024b).

Functions frequently conflated in non-specialist accounts must, however, be distinguished. SciELO and Redalyc are full-text open access journal indexing services that apply editorial criteria for admission and continued inclusion and, to that extent, perform a certifying function. LA Referencia, by contrast, is a network of repositories: it harvests research outputs, reports, and theses from more than 800 repositories in ten Ibero-American countries, currently bringing together approximately 2.95 million articles, 1.02 million master's dissertations, 470,000 doctoral theses, and 9,600 datasets (Beigel et al., 2024b). Its contribution to multilingualism and bibliodiversity is substantial, but it operates through metadata aggregation rather than editorial certification—a distinction that matters when discussing editorial policy strategies, because the available levers differ in each case.

This infrastructure sustains what has become known as the diamond open access model—free both to readers and to authors, with no article processing charges (APCs)—predominantly operated by universities and public institutions. Its scale is internationally significant: the OA Diamond Journals Study attributes approximately 25% of the world's journals in this category to Latin America (Bosman et al., 2021). The model nevertheless faces financial sustainability challenges. The regional literature identifies the main pressure as the growing commodification of scientific publishing and the spread of APC-funded open access, which introduces into the ecosystem a logic different from the one that historically sustained this circuit (Beigel et al., 2024b). It should be noted that the extent of this adoption within regional journals themselves has not been systematically quantified in the available literature, which is why no precise magnitudes are assigned to it here. Even so, accumulated experience demonstrates that it is possible to build high-quality publishing circuits founded on the principle of scientific knowledge as a public good.

5.4. The tension between internationalization and multilingualism

It would be simplistic, however, to present the regional infrastructure as an unconditional defense of multilingualism. SciELO itself has experienced a revealing tension over the past decade. In its internationalization policy, announced in January 2014, the program began strongly encouraging publication in English as a means of expanding international circulation, setting minimum percentages of articles in that language as indicators of journal internationalization: 80% for the broad field of health and 25% for the humanities and applied social sciences, accompanied by minimum percentages of authors with foreign affiliations of 25% and 20%, respectively (FAPESP, 2014).

The target should be clearly distinguished from the outcome. As a stated editorial policy target, the percentage does not describe the actual composition of the collection: in 2019, the SciELO network comprised more than 1,200 journals publishing approximately 50,000 articles per year, only 42% of which were in English. There is, however, evidence that the policy had an effect where it was most strongly applied. For Brazilian journals indexed in SciELO and Redalyc, Beigel et al. (2024b) report an established trend toward migration to English that they explicitly attribute to SciELO Brazil's indexing policies: in 2016, articles in English surpassed those written in Portuguese within this subset, and the gap continued to widen. Salatino's (2022) finding that Brazil has the highest proportion of regional output in English (33.9%, compared with 10% for Argentina) is consistent with this interpretation, although it does not establish it independently, since other factors—disciplinary profile, publishing tradition, and internationalization capacity—operate simultaneously.

This policy exposes the dilemma at the heart of the problem. On the one hand, requiring English reproduces within the regional infrastructure itself the hegemony it was intended to counterbalance. On the other, SciELO's editorial model preserved a decisive feature: the possibility of simultaneous publication in two or more languages and the requirement that the title, abstract, and keywords always also be available in the text's original language when it is not English (SciELO, 2020). The resulting path was therefore not the wholesale replacement of national languages by English but an arrangement for their coexistence—an arrangement under strain, as the Brazilian data show, whose balance depends on editorial policy decisions that can be revised. The next section addresses this scope for decision-making.

6. Editorial strategies to expand circulation

If the more limited international circulation of Latin American output stems to a significant extent from structural asymmetries in the publishing system, overcoming it depends less on individual researchers' efforts and more on deliberate editorial policy decisions. This section discusses three complementary strategies. It should be emphasized at the outset that all three are normative proposals grounded in argumentative plausibility and partial experiences, not interventions whose effects on circulation or citation have been measured; empirical assessment of these effects is precisely the outstanding agenda identified in the concluding remarks.

6.1. Multilingualism as a deliberate editorial policy

The first and most comprehensive strategy is to shift the burden of translation from individual authors to the editorial infrastructure, adopting multilingualism as a deliberate policy rather than an exception. Instead of requiring researchers to bear the linguistic costs of access to the international circuit, the journal or publisher assumes responsibility for translation and for making content available in multiple languages. This reversal of responsibility is precisely the central recommendation of Amano et al. (2023), who argue that journals, as

gatekeepers of the system, should actively promote multilingualism in science. Salatino (2022) formulates a convergent proposal based on the regional case: since Latin American journals already receive articles in several languages, the next step would be to enable the publication of the same article in at least Spanish, Portuguese, and English, making it possible to cross linguistic boundaries within the region itself.

The simultaneous publication of the same work in more than one language—a practice that Lillis and Curry (2010) describe in their analysis of multilingual scholars' strategies as a form of equivalent publication—makes it possible to reconcile the two poles of the paradox identified in the introduction: the work reaches international audiences without relinquishing its local and regional accessibility. When incorporated into the editorial structure rather than relegated to isolated author initiatives, this practice turns multilingualism from an obstacle into an asset for circulation. The Helsinki Initiative on Multilingualism in Scholarly Communication, launched in April 2019, formalized this principle internationally through three recommendations: support the dissemination of research results for the benefit of society, protect national infrastructures for publishing locally relevant research, and promote linguistic diversity in research assessment and funding systems (Helsinki Initiative, 2019). The status of this support should be qualified: this is a normative document, endorsed by hundreds of institutions, that establishes multilingualism as a principle of science policy. The Initiative does not present—nor does it seek to present—empirical evidence that multilingual publication increases impact or citation. Its argumentative value lies in institutional legitimacy, not in demonstrating an effect.

6.2. Artificial intelligence-assisted translation: Opportunity and caution

The second strategy concerns the emerging role of machine translation technologies. Historically, the cost of professional translation was the principal obstacle to adopting multilingual policies on a large scale. Recent advances in artificial intelligence-assisted translation tools substantially change this equation by markedly reducing the cost and time required to make content available in multiple languages.

The evidence of adoption is striking and, significantly, is strongest precisely where the barriers are greatest. In a preprint analyzing more than two million full-text biomedical publications in PubMed Central between 2021 and 2024, Liu, He, Zheng, Bu, and Ni (2025) estimate the proportion of AI-generated content and document, following the launch of generative AI tools, growth in assisted writing of approximately 400% in non-English-speaking countries, compared with approximately 183% in English-speaking countries, with more intensive adoption among less established, early-career researchers and those at less prestigious institutions. The authors also associate more intensive AI use with a modest increase in productivity, alongside a narrowing of the publication gap between scientists in English-speaking and non-English-speaking countries. Three qualifications are essential. The first concerns status: this is a preprint that has not yet undergone peer review. The second concerns scope: the corpus is biomedical and English-language, and what is measured is assisted writing, not the translation of articles between languages—extrapolation to

editorial translation policies in the humanities and social sciences is therefore indirect. The third concerns inference: the observed narrowing is an association between adoption and productivity, and the authors themselves frame it as a potential opportunity to mitigate linguistic inequalities, not as a demonstrated reduction in those inequalities.

This promise is indeed not unconditional. In an essay published in PLOS Biology, Amano, Bowker, and Burton-Jones (2025) argue that AI-mediated translation opens up two possible futures for academic publishing. In the first, English remains the lingua franca, but researchers with limited proficiency write in their own language and use AI to translate before submission and to read and assess others' texts—knowledge remains centered on English, although more accessible. In the second, peer review itself becomes denationalized: each participant in the editorial process can read, write, and review in their own language. The authors maintain that the choice between these futures is not technological but political and institutional, depending on how infrastructure, training data, and governance are distributed. A technical caution should be added: machine translation does not eliminate the need for qualified human review, especially regarding terminological precision and cultural appropriateness, lest new errors be introduced. The most robust strategy combines the speed and low cost of machine translation with specialized human review as a quality-control layer.

6.3. Legal and technical conditions for circulation

The third strategy concerns the legal and technical conditions that make the first two feasible. In legal terms, translating means producing a derivative work, which requires authorization from the rights holder. Creative Commons licenses that do not include the NoDerivatives restriction—including Attribution (CC BY), Attribution-ShareAlike (CC BY-SA), Attribution-NonCommercial (CC BY-NC), and Attribution-NonCommercial-ShareAlike (CC BY-NC-SA)—provide advance authorization for translation and redistribution, each subject to its own conditions. CC BY 4.0 is the most permissive and therefore creates the least friction for multilingual policies, but it is not an indispensable condition: the actual requirement is the absence of the ND clause, not the adoption of a specific license. It should be noted that most of the Latin American circuit operates under noncommercial Creative Commons licenses, which authorize nonprofit translations and are compatible with the regional diamond model.

At the technical level, assigning persistent digital identifiers (DOIs) serves a complementary function, but it must be described precisely. A DOI is a persistent identifier that ensures stable address resolution and facilitates the discovery and correct citation of a work; it does not, by itself, guarantee indexing in any database or generate citations, which depend on editorial criteria and reading practices beyond the identifier. For translated versions, Crossref recommends registering a separate DOI for each translation and linking them through relationship metadata—the hasTranslation and isTranslationOf types—a procedure that entails no additional registration cost; the agency itself acknowledges that its support for multilingual content

remains limited (Crossref, 2025). Without these relationship metadata, versions of the same work in different languages tend to be treated as independent objects, fragmenting the citation record rather than consolidating it—thus turning a strategy for expanding reach into a traceability problem.

The combination of these three strategies—structural multilingualism, AI-assisted translation with human review, and open licensing with appropriate relationship metadata—outlines an editorial model capable of addressing both sides of the problem described throughout this article: expanding the international reach of Latin American science without sacrificing its local and regional accessibility. In its architecture, this model internalizes the principle that responsibility for overcoming language barriers should be collective and institutional rather than individual. Whether this model actually produces greater circulation, and to what extent, remains an open empirical question.

7. Concluding remarks

This article has sought to bring together and qualify the available evidence on the relationship between language barriers and the circulation of Latin American scientific output. What this evidence supports can be stated precisely. The hegemony of English, consolidated over the twentieth century (Hamel, 2007), transformed language from a vehicle of communication into a condition for access to the institutionally most rewarded publishing circuit. This condition imposes documented additional costs on non-English-speaking researchers: among environmental scientists from eight countries, non-native speakers of nationalities with low proficiency spend a median of 90.8% more time reading an article in English and report language-related rejection at a frequency 2.5 to 2.6 times higher and language revision requests at a frequency 12.5 times higher than native speakers (Amano et al., 2023); among Mexican and Taiwanese researchers, writing in English increases perceived difficulty by 24% and anxiety by 22% compared with writing in the first language (Hanauer et al., 2019); and among Colombian doctoral students in the biological sciences, the cost of translation services per article equals between one-quarter and one-half of a doctoral student's monthly salary in the country (Ramírez-Castañeda, 2020). Each of these figures is inseparable from the design that produced it, and none independently warrants a claim about Latin American researchers as a whole.

For the regional circuit, the evidence is different in nature and equally robust. The OLIVA corpus shows a vast, multilingual publishing space—43.7% of articles in Spanish, 32.1% in Portuguese, and 23.9% in English—whereas 90.8% of articles in Scopus and 96% of those in the Web of Science Core Collection are in English, and only 3.4% and 5.2% of journals in those databases are Latin American, compared with 4,077 active journals indexed in the region (Salatino, 2022; Beigel et al., 2024a, 2024b). These figures document limited representation and partial overlap between circuits; they do not, however, measure the scientific quality of the outputs being compared or the relative causal weight of asymmetries versus other factors. The formulation supported by the evidence is therefore more restrained than that usually encountered: the more

limited presence of Latin American output in dominant databases is consistent with documented linguistic, economic, and coverage barriers and cannot be read as an indicator of lesser merit—but claiming that quality has demonstrably been ruled out as a factor would exceed what this review measured. The findings of Di Bitetti and Ferreras (2017), whose design comes closest to a direct test, point in this direction: comparing 1,328 articles from six journals, three of them Latin American, the authors found a higher likelihood of citation for texts in English, but blinded assessors did not consistently identify greater quality, interest, or originality in English abstracts, and data supplied by editors revealed no difference in rejection rates between submissions in English and other languages in two of those journals.

The article has also sought to avoid oversimplification. The debate between Hyland (2016) and Politzer-Ahles et al. (2016) shows that the language barrier does not explain everything but constitutes a real and quantifiable dimension of disadvantage, intertwined with economic and structural factors that the available designs do not fully isolate. Reading the evidence through the notion of epistemicide (Santos, 2014) offers an interpretive lens for understanding why invisibility cannot be resolved through translation alone—a lens whose status, as noted above, is theoretical rather than demonstrative.

The region, however, has not simply been a passive recipient of this order. The construction of a globally pioneering public and multilingual open access infrastructure, comprising SciELO, Redalyc, Latindex, Biblat, and LA Referencia (Beigel et al., 2024a, 2024b), demonstrates that it is feasible to sustain high-quality publishing circuits anchored in national languages—and on an internationally significant scale, given that approximately 25% of the world's diamond journals are published in Latin America (Bosman et al., 2021). SciELO's own trajectory, with its internationalization policy encouraging English (FAPESP, 2014) and its visible effects within the Brazilian subset (Beigel et al., 2024b), reveals that the challenge is not to choose between international circulation and local relevance but to construct arrangements accommodating both through multilingualism rather than language replacement.

The strategies discussed—multilingualism as a deliberate policy, artificial intelligence-assisted translation with human review, and open licensing with appropriate relationship metadata—point toward a fundamental shift: transferring the burden of overcoming language barriers from individual researchers to the editorial infrastructure. This shift is a necessary, although insufficient, condition for changing the situation described. It is insufficient because full transformation also depends on changes in research assessment regimes, which still disproportionately reward publication in dominant databases, and because, as Amano, Bowker, and Burton-Jones (2025) warn, AI itself may reinforce asymmetries if its governance remains concentrated. Yet it is necessary because, without an editorial infrastructure that collectively undertakes the task of multilingualism, the promise of an open and global science will remain, for most, an unfulfilled promise.

Two research agendas follow directly from what this review could not establish. The first concerns scope: the body of studies on individual linguistic costs in Latin American contexts needs to be expanded beyond the

Bolivian and Colombian cases and into the humanities and social sciences, where the empirical literature on the topic is comparatively limited. The second concerns effects: the impact of the editorial strategies discussed here should be evaluated empirically, measuring the effects of multilingual publication on citation rates, geographical reach, and use by non-academic audiences, as well as the actual quality of AI-assisted translations across fields. It would be equally valuable to measure overlap between the regional and international circuits directly in terms of cross-citation rather than indexing alone. This review provides the conceptual foundation and the state of the evidence on which such investigations can build.

Declarations

Author contributions: Conceptualization, formal analysis, investigation, methodology, writing—original draft, and writing—review and editing: J. B. C. and B. C. A. M. Both authors have read and agreed to the submitted version of the manuscript.

Funding: This research received no external funding.

Conflicts of interest: The authors declare no conflicts of interest.

Data availability: This study is a critical literature review and did not generate primary data. All sources analyzed are identified in the reference list and are publicly accessible. The databases consulted and search terms used are described in the methodology section.

Ethics approval: Not applicable. The research did not involve human participants or animals.

Informed consent: Not applicable.

Use of artificial intelligence: Generative artificial intelligence tools were used to assist with the literature search and language editing of the manuscript. All content produced in this way was checked against primary sources and reviewed by the authors, who assume full responsibility for the final text.

Supplementary materials: There are no supplementary materials.

References

Amano, T., Bowker, L., & Burton-Jones, A. (2025). AI-mediated translation presents two possible futures for academic publishing in a multilingual world. *PLOS Biology, 23*(6), e3003215. https://doi.org/10.1371/journal.pbio.3003215

Amano, T., González-Varo, J. P., & Sutherland, W. J. (2016). Languages are still a major barrier to global science. *PLOS Biology, 14*(12), e2000933. https://doi.org/10.1371/journal.pbio.2000933

Amano, T., Ramírez-Castañeda, V., Berdejo-Espinola, V., Borokini, I., Chowdhury, S., Golivets, M., González-Trujillo, J. D., Montaño-Centellas, F., Paudel, K., White, R. L., & Veríssimo, D. (2023). The manifold costs of being a non-native English speaker in science. *PLOS Biology, 21*(7), e3002184. https://doi.org/10.1371/journal.pbio.3002184

Beigel, F., Packer, A. L., Gallardo, O., & Salatino, M. (2024a). OLIVA: La producción científica indexada en América Latina. Diversidad disciplinar, colaboración institucional y multilingüismo en SciELO y Redalyc (1995-2018) [OLIVA: Indexed scientific output in Latin America. Disciplinary diversity, institutional collaboration, and multilingualism in SciELO and Redalyc (1995-2018)]. *Dados, 67*(1), Article e20210174. https://doi.org/10.1590/dados.2024.67.1.307

Beigel, F., Sánchez Pereyra, A., Alonso-Gamboa, J. O., Salatino, M., Gallardo, O., Ferreira Gonçalves, A., Carrillo Romero, O., Flores Chávez, M. A., Durán Muñoz, E., & Arguello Mendoza, M. G. T. (2024b). OLIVA-2: Las revistas iberoamericanas indexadas en Biblat y Latindex: fuentes fundamentales para conocer la producción científica global [OLIVA-2: Ibero-American journals indexed in Biblat and Latindex: Fundamental sources for understanding global scientific output]. *E-Ciencias de la Información, 14*(1), 1-24. https://doi.org/10.15517/eci.v14i1.55951

Bosman, J., Frantsvåg, J. E., Kramer, B., Langlais, P.-C., & Proudman, V. (2021). *OA diamond journals study. Part 1: Findings* [Report]. Zenodo. https://doi.org/10.5281/zenodo.4558704

Crossref. (2025). *Multi-language material and translations*. https://www.crossref.org/documentation/principles-practices/best-practices/multi-language/

Di Bitetti, M. S., & Ferreras, J. A. (2017). Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. *Ambio, 46*(1), 121-127. https://doi.org/10.1007/s13280-016-0820-7

Fundação de Amparo à Pesquisa do Estado de São Paulo. (2014, January 6). *SciELO announces measures for internationalizing scientific journals*. Agência FAPESP. https://agencia.fapesp.br/scielo-announces-measures-for-internationalizing-scientific-journals/20454

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91-108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Hamel, R. E. (2007). The dominance of English in the international scientific periodical literature and the future of language use in science. *AILA Review, 20*, 53-71. https://doi.org/10.1075/aila.20.06ham

Hanauer, D. I., Sheridan, C. L., & Englander, K. (2019). Linguistic injustice in the writing of research articles in English as a second language: Data from Taiwanese and Mexican researchers. *Written Communication, 36*(1), 136-154. https://doi.org/10.1177/0741088318804821

Helsinki Initiative. (2019). *Helsinki Initiative on multilingualism in scholarly communication*. Federation of Finnish Learned Societies, Committee for Public Information, Finnish Association for Scholarly Publishing, Universities Norway & European Network for Research Evaluation in the Social Sciences and the Humanities. https://doi.org/10.6084/m9.figshare.7887059

Hyland, K. (2016). Academic publishing and the myth of linguistic injustice. *Journal of Second Language Writing, 31*, 58-69. https://doi.org/10.1016/j.jslw.2016.01.005

Lillis, T., & Curry, M. J. (2010). *Academic writing in a global context: The politics and practices of publishing in English*. Routledge.

Liu, J., He, Y., Zheng, Z., Bu, Y., & Ni, C. (2025). *AI-assisted writing is growing fastest among non-English-speaking and less established scientists* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2511.15872

Politzer-Ahles, S., Holliday, J. J., Girolamo, T., Spychalska, M., & Berkson, K. H. (2016). Is linguistic injustice a myth? A response to Hyland (2016). *Journal of Second Language Writing, 34*, 3-8. https://doi.org/10.1016/j.jslw.2016.09.003

Ramírez-Castañeda, V. (2020). Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: The case of Colombian researchers in biological sciences. *PLOS ONE, 15*(9), Article e0238372. https://doi.org/10.1371/journal.pone.0238372

Salatino, M. (2022). Los circuitos lingüísticos de la publicación científica latinoamericana [The linguistic circuits of Latin American scientific publishing]. *Tempo Social, 34*(3), 253-273. https://doi.org/10.11606/0103-2070.ts.2022.201928

Santos, B. de S. (2014). *Epistemologies of the South: Justice against epistemicide*. Routledge.

Scientific Electronic Library Online. (2020). *Criteria, policy and procedures for the admission and permanence of journals in the SciELO Brazil collection*. https://www.scielo.br/media/files/20220900-scielo-brazil-criteria-en.pdf