Scientometrics (2022) 127:7369–7395
https://doi.org/10.1007/s11192-022-04403-9

Local emergence, global expansion: understanding the structural evolution of a bi-lingual national research landscape

Matias Federico Milia[1,2] · **Ariadna Nebot Giralt**[2] · **Rigas Arvanitis**[2,3]

Received: 14 October 2021 / Accepted: 5 May 2022 / Published online: 21 June 2022

Abstract
Research institutions organize their scientific activities in an increasingly diverse landscape. In matters of global interest, research relies on an ever-more cross-disciplinary background, which reveals intriguing questions concerning the local dynamics vs. global audiences. This paper proposes new methodological tools to assess, from a strategic perspective, the evolution of a given research landscape. It relies on the Global Research Institute of Paris' recent experience, a new interdisciplinary Institute focusing on globalization topics beyond the usual economic meaning. The Institute leans on a broad and diverse set of research units of the Université de Paris and relates to the broad landscape of social sciences in France. This article charts the evolution of French authors' scientific publications on the Institute's thematic interests in French and English. It focuses on the structural features of the debate, namely the volume, the underlying historical semantic structure, and its main thematic domains. The paper offers significant evidence to understand knowledge circulation dynamics and links that non-speaking countries' scientific literature builds with the English one.

Keywords Knowledge circulation · Strategic-oriented analysis · Global research · Urban space · Technologies

Introduction

Research institutions build their research agendas in an ever-more international, dynamic, and diverse landscape. Many of these efforts are increasingly cross-disciplinary, either in collaboration or competition. Various institutions are motivated to inform their activities

✉ Matias Federico Milia
matias@milia.net

[1] FLACSO-México, Carretera al Ajusco 377, Colonia Héroes de Padierna, Tlalpan, Mexico, Mexico

[2] Global Research Institute of Paris (GRIP), Université de Paris, 45 rue des Saints-Pères, 75006 Paris, France

[3] Centre Population et Développement (Ceped), Université de Paris & Institut de Recherche Pour le Développement (IRD), 45 rue des Saints-Pères, 75006 Paris, France

through the analysis of available bibliometric data. Understanding a specific research landscape is critical to steer global and multidisciplinary research agendas (Wallace & Ràfols, 2015, 2018). Different approaches tend to emerge in research areas with an international scope, gathering varied research interests from diverse origins and disciplinary backgrounds. One way of observing these complex groupings is to assess emerging trends over time (Ràfols et al., 2010; Zitt & Bassecoulard, 1994). It is possible thus to describe the emerging tendencies and follow their evolution.

However, such an interpretative approach may reveal not so much the strategic choices of authors and affiliated institutions but rather the limitations experienced by researchers in their academic environment (Cruz-Castro & Sanz-Menéndez, 2018). Institutional structures tend to reflect the history of the scientific endeavor (Trow, 1999, p. 4). These questions are important when considering the emergence and circulation of new thematics. Local interests may differ from those appearing in the global arenas, but research in other languages tends to be overlooked in English literature (Chi, 2012). How global transformations of research affect local scientific choices is not only a challenging and interesting question; it is also central to understand the growing globalization processes (O'Brien & Arvanitis, 2019).

We present the work done at the Global Research Institute of Paris (GRIP), a new French Institute willing to map, from an institutional perspective, research areas with a global scope. In particular, GRIP wanted to investigate the possible distance between its own strategic choices and those found in the French and international literature. We have chosen an extended period, from 1980 to 2020, to better grasp the knowledge aggregation process behind these issues.

GRIP is an interdisciplinary institute focusing on globalization beyond its usual economic dimensions. GRIP builds upon the broad and diverse social science landscape inside the perimeter of the Université de Paris. Its interpretation of the concept of "Global Research" bounds to a vast array of research units scattered over Paris and relates to the broad landscape of social sciences in France. The multilingual publishing practices toward diverse audiences in social science research becomes a privileged standpoint of analysis (Kulczycki et al., 2020).

The differences in language use in science have long been raised (Garfield, 1990). However, rising technical possibilities are making it now possible to develop a discussion of this kind empirically. As international indexes limit the inclusion of many sources published in languages other than English, alternative languages can become a suitable option to understand alternative approaches to similar subjects. Studies on other European countries show that documents not indexed in citation indexes are, indeed, a valuable and often untapped source of information (Chi, 2014).

Languages that include countries with larger scientific communities can support debates somehow detached from mainstream preoccupations, a particularly notable feature, since different languages imply distinct circulation channels for published texts. Nevertheless, it has been stressed that English language is an important vehicle for international recognition (Meneghini & Packer, 2007). Research has shown that, contrary to the common assumptions about the globalization process, national research landscapes have strengthened their collaboration since the 2000' (Maisonobe et al., 2016). In matters of Global Research, such as those behind GRIP's interpretation of the subject, international acceptance goes beyond prestige; it is also necessary to establish a truly global scientific dialogue and collaborations.

GRIP chose to concentrate on three thematic areas: Global Urbanities; Circulations; and Technologies, Markets and Vulnerabilities. These topics were chosen through exchanges

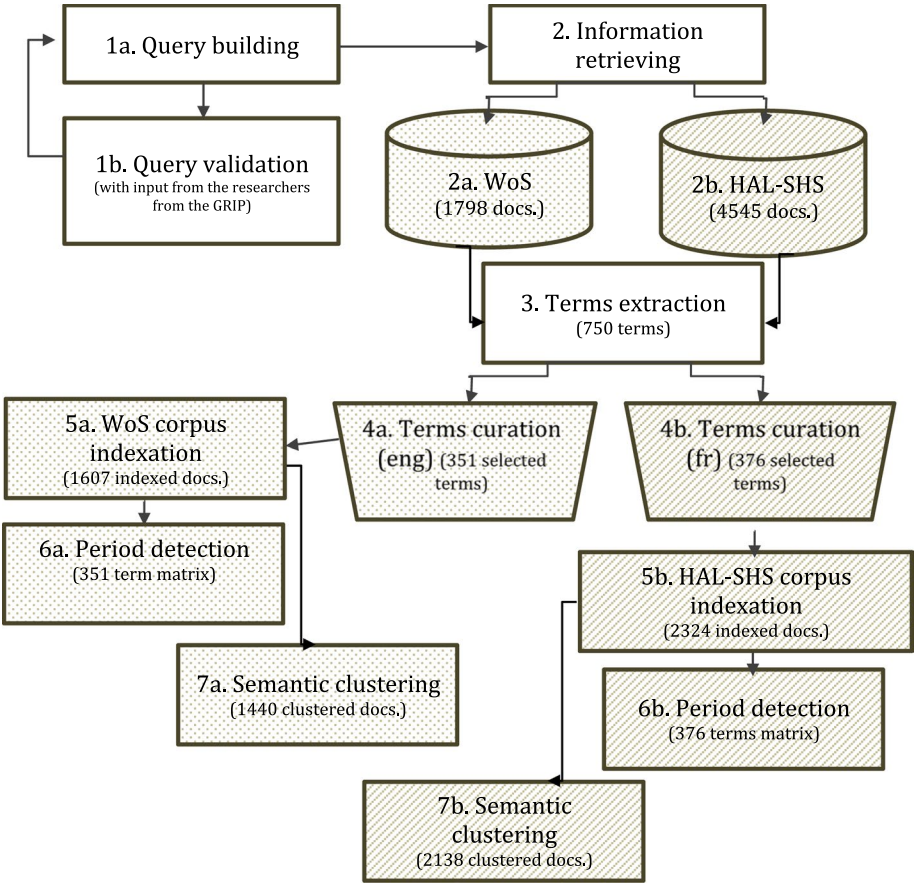

Fig. 1 Main operations conducted during the analysis

among academics located in the University, taking into account their interests, research trajectories, and previous projects, but with no previous strategic analysis. It was thus interesting to see if the Institute's topics were different or coherent with the available literature, by comparing them with topics appearing in the international 'mainstream' publications and those in French publication repositories.

Materials and methods

This article charts literature related to Global Research themes that inspired GRIP's development in two different languages (French and English). To better explain the analytical operations conducted, a detailed account has been included in Fig. 1, numbering each of the followed steps. Specific queries were built to capture documents involving each of the Institute's thematic areas. As it remains crucial to situate these thematic choices, our research focuses exclusively on French authors. To represent the GRIP's scientific objects and interests, we shared our research queries with a sample of researchers involved in the

Institute's thematic areas. In this way, we focused on refining a precise query for each thematic area. After collecting and examining our first harvest of documents, the feedback from these same researchers became an input to fine-tune our queries.

As a result, six different queries for information retrieval were written, one for each research axis in both of the chosen sources. For English documents, information was retrieved from the Web of Science database. Items written in French were accessed from the HAL-SHS collection using the API made available by the platform (CCSD, 2021). HAL-SHS is a mandatory repository for French researchers that has been poorly explored and remains highly relevant since regional databases are an essential source to study social sciences' research output (Engels et al., 2012). In both cases, the fields we retrieved were the titles, abstracts, keywords, authors, and publication years. In WoS, the search strategy was based on the documents' topics. In HAL-SHS, full texts were available and searched; we excluded doctoral thesis and other teaching-related documents. We refined and improved our search strategy by consecutive iterations after analyzing the results in a continuous dialog with the researchers, experts in the field. As a result, we established the following query lines for each language, as shown in "Appendix" (Table 1).

In this process, we gathered 1798 English documents from the Web of Science (WoS) and 4545 French documents from the HAL-SHS, both between 1980 and 2020. The query aimed at articles, books, book chapters, and research communications in both sources. We processed the two resulting collections with the help of the CorText platform using NLP techniques to obtain a fine-grain meaningful set of 750 terms describing the main semantic features of both collections. These lists of terms were manually curated[1] to eliminate noise and shallow terms in both French and English.

A *Period Detector* script was used to capture the semantic structure. We analyzed the frequency distribution of significant terms over the years and computed the optimal partitions in subsequent periods. We selected only the words with a minimum frequency of eight, 376 terms in French and 351 in English. We are able to represent distances between the semantical composition of our data for every pair of different time steps in a scale from most-similar (0) to most-dissimilar (1). Periods were computed using the gap statistic method (Tibshirani et al., 2001). As a result, we could improve the stability of differences among subsequent years. In this way, we gain access to the underlying knowledge aggregation process reflected in these documents. This method allows grasping the inner coherence in each detected period, as the distance between all pairs of years included within each period has been computed and presented using the same scale (from 0 to 1).

Textual analysis is a mapping technique used to assess research landscapes and identify thematic domains (Barbier et al., 2012). In order to describe the topical structure of analyzed documents in French and English, we produced a co-occurrence network with the same terms used for period detection. We used the Louvain algorithm (Blondel et al., 2008), a bottom-up modularity-based clustering algorithm, to detect emerging semantic clusters. Given the not deterministic nature of this process, slight variations can happen in different computations. Nevertheless, given the size of the calculated networks, additional runs were computed, still reporting a consistent community structure with the ones here reported.

[1] Operations were conducted to clean terms. Void terms were excluded. Specific terms were homologated manually when the software had not detected their synonymity. Foreign terms were kept when related to objects of interest. 257 terms were deleted or homologated of the French collection. In English, a total of 675 terms remained.

We interpreted these clusters as the thematic structure of the documents in our corpus. To reduce the complexity of the resulting network and optimize the modularity in its partition, we defined a resolution parameter (Lambiotte et al., 2014). For this optimization, we used an algorithm to define a parameter resolution value (Aynaud, 2020) of two (2) within a range from 0.1 up to 4.9.

Clusters were manually labeled according to their structure and contents, using the terms' weight and centrality within the cluster as a benchmark. French clusters were labeled in English to make them more accessible to readers. Although French detected terms were kept in their original language in Fig. 5, they were translated to English by authors and included in the "Appendix" (Table 3).

We studied the overall structure of the clusters considering terms relevance in the context of the GRIP and its connection structure within the cluster. The resulting terms were treated taking into account criticism to co-word analysis and their relation towards the context of use (Leydesdorff & Hellsten, 2006). In the term detection phase, our NLP approach allowed us to grasp a more complex set of terms—i.e., noun phrases—that better describe expressions in a language-in-use setting, detecting lexical variations in a broader context. For clarity, we included a detailed description of terms composing each cluster and each topical domain's overall density to better understand how well connected the linkages inside each cluster are.

We compared the overall structure of the terms describing French and English corpus. We stressed during our analysis two structural features to characterize discussions undertaken by French authors when comparing both languages: first, noun phrases related to theoretical approaches and objects of study; second, those pointing towards the geographical scope of the research and the involved stakeholders.

Findings

Publication growth over time

This analysis describes the themes related to the scientific interests of the GRIP in French publications in two different languages. When looking at the total number of documents complying with the criteria shown in the "Appendix" (Table 1), we discovered that within these issues, as defined above, more documents were published before 2005 in French (393 documents) than in English (115 docs), and publications grew earlier and faster in French than in English (Fig. 2). When comparing both curves, there is a significant time lag between French and English documents. The volume of publications around these issues picked up in French between 2002 and 2003, but only six years later in English, between 2008 and 2009. The French language growth seems to feed English-language papers, but with a significant time lag. Finally, we observe that the number of published English-language documents stagnates from 2018 and on.

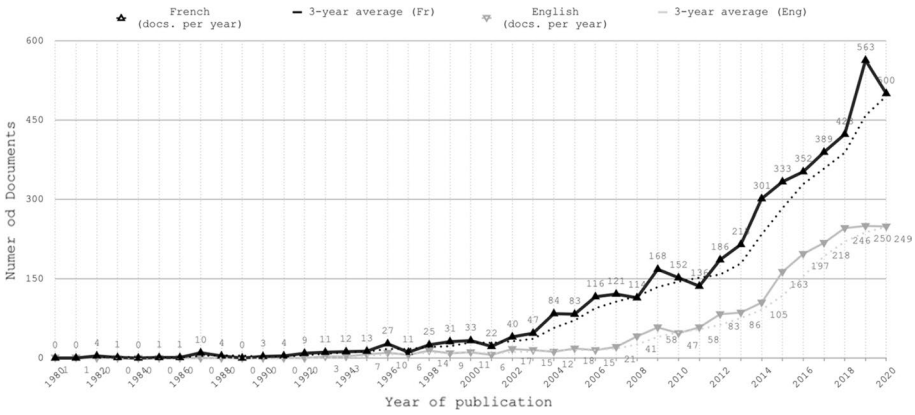

Fig. 2 Number of documents representing the Institute's thematic interests in English and French

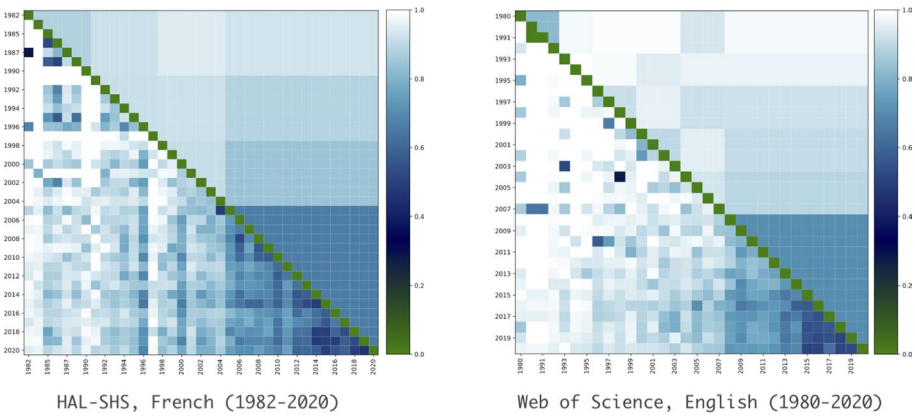

Fig. 3 Period detection and significant term correlation between years in French and English

Detected lexical period structure

The underlying semantic structure, summarized in Fig. 3,[2] allows us to understand how this growth unfolds. It charts how many terms used in one year resemble those from the other years in the series, representing this using a scale from 0 (green) to 1 (white). The higher the value, the more dissimilar those 2 years are. Dissimilarity is represented by whiter cells and similarity by darker blue-green cells. On each graphic's upper right, detected most similar periods are represented by using the same color scale. The resemblance of the pairs of years within each detected period—i.e., inner coherence—is calculated and presented, resorting to the same scale (from 0 to 1). This analysis allows us to better understand how consistent and consolidated each period is.

[2] A year-by-year detail of the numerical results for the period detection have been included in the "Appendix".

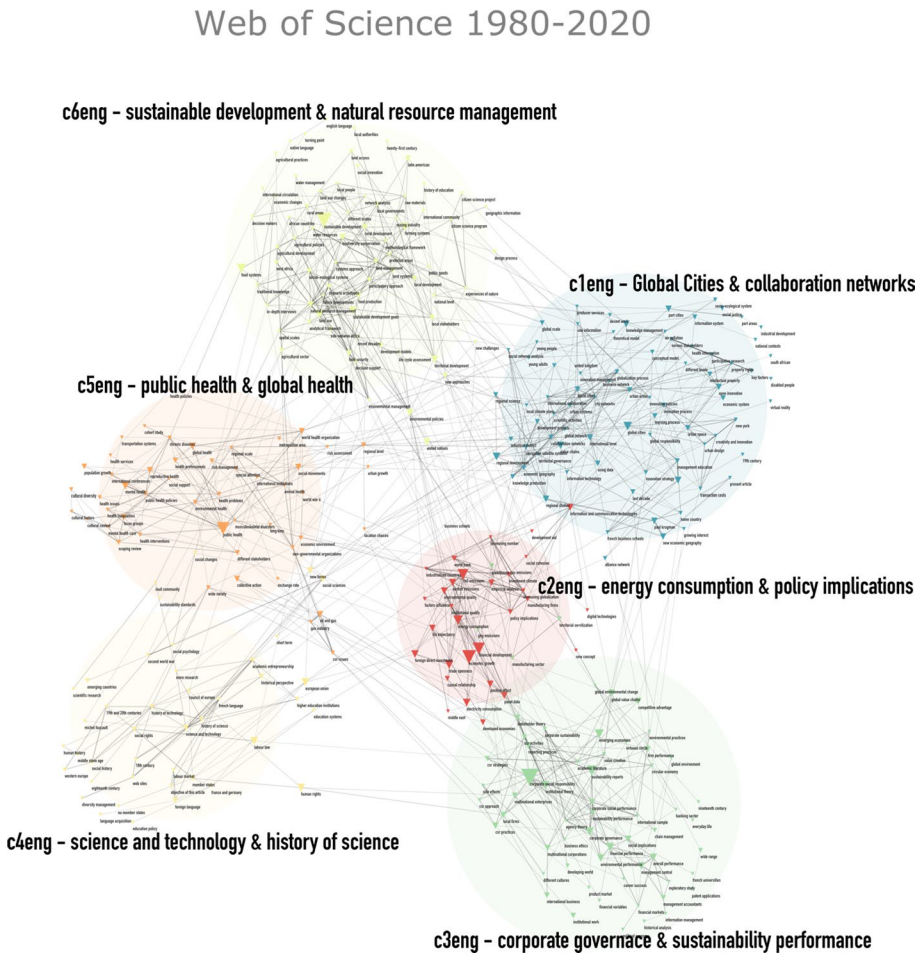

Fig. 4 English terms co-occurrence network

In French, three short periods, between four and eight years of duration, are detected. Going from 1982 to 1990, the first one has a low inner coherence (0.88). The second (1991–1997) and the third (1998–2004) periods deal with a large variety of issues; hence they also show low inner coherence ratios (0.91 and 0.92, respectively). The fifth and last is the one that covers more time, from 2005 to 2020. It is also the one with documents that resemble the most to each other (0.65). Periods relate to each other, showing a steady knowledge aggregation process when compared to this last period.

In English, the similarity between the papers over the years is more irregular and weaker. The inner coherence in the first five periods serves as an example to this: 1980 to 1992 (0.85), 1993 to 1995 (0.98), 1996 to 1999 (0.92), 2000 to 2003 (0.91), and 2004 to 2007 (0.90). Here, detected periods are more—six in total—and also shorter. Shorter periods mean that topics remain stable for a shorter time. Also, the resemblance between periods is not so strong as the one observed in the French publications and the aggregation process, showing that some of the issues addressed do not recover over time. In general, the debate in English appears looser and more erratic. Also, it is

important to stress that it is not till 2008 that the English debate starts showing stronger internal cohesion (0.71).

The feeding process observed in the publication growth over the years is consistent with the observed results. At the end of the studied period, the debate becomes better structured in French (0.65 against 0.71 in English), happening earlier in 2005 in the French debate and 2008 in English.

Overall thematic features

The most salient characteristics of documents published by French authors in English on topics related to GRIP's interests appear summarized in Fig. 4, while a detailed list of the terms that constitute each thematic domain are summarized in Table 2, available in the "Appendix". When applying our parametrization to the set of curated terms describing retrieved documents, we came up with six meaningful clusters representing main thematic domains. A first cluster (c1eng), labeled *Global Cities & collaboration networks*, focuses on global and world cities. Many perspectives rise in the subject, such as global value chains, innovation studies, or regional studies. Regarding key locations emerging from this analysis, only one city New York—and two countries—United Kingdom, South Africa—appears in our outline.

A second cluster (c2eng) relates to *energy consumption & policy implications*. It centers mostly on financial, manufacturing, and trade (policy) implications of the current energy matrix. It appears oriented chiefly towards developed economies. Notoriously, the World Bank is a relevant stakeholder connected to this discussion. A third thematic domain (c3eng) links to *corporate governance & sustainability performance*. Here, we have detected close relations between corporate governance, corporate responsibility activities, and specific performance areas, such as chain management and business practices. With a particular emphasis on sustainability, its main focus is on emerging economies.

Most notably, science and technology issues arise concerning socio-historical approaches. Then, a specific set of preoccupations rises around *science and technology & history of science* (c4eng). It seems notoriously oriented towards the European Union and higher education institutions within its boundaries. Sanitary issues that emerge in *public health & global health* (c5eng) relate primarily to chronic diseases, reproductive health, and mental health problems. The World Health Organization (WHO) appears a relevant actor for these matters, as well as Social Movements. These actors tend to be mentioned separately, either on reproductive health (social movements), or public health (WHO).

Finally, *sustainable development & natural resource management* (c6eng) shows sustainable development concerns associated principally with land-related and agricultural issues. Local actors appear as the relevant stakeholders; West Africa is linked to food security and Latin America to rural development.

French literature on the same topics shows a different configuration. Here, seven clusters result from our analysis. A detailed list of the detected clusters and their composing terms is available in the "Appendix" (Table 3), while a graphic representation of this research landscape is available in Fig. 5.

A first cluster (c1fr) stresses the interest of French authors in Africa when addressing global research issues. Under the umbrella of *Africa & contemporary societies*, Africa relates mainly to sanitary risks and waste management. Daily life descriptions in urban and rural settings characterize much of the research done in France about Africa. We did not

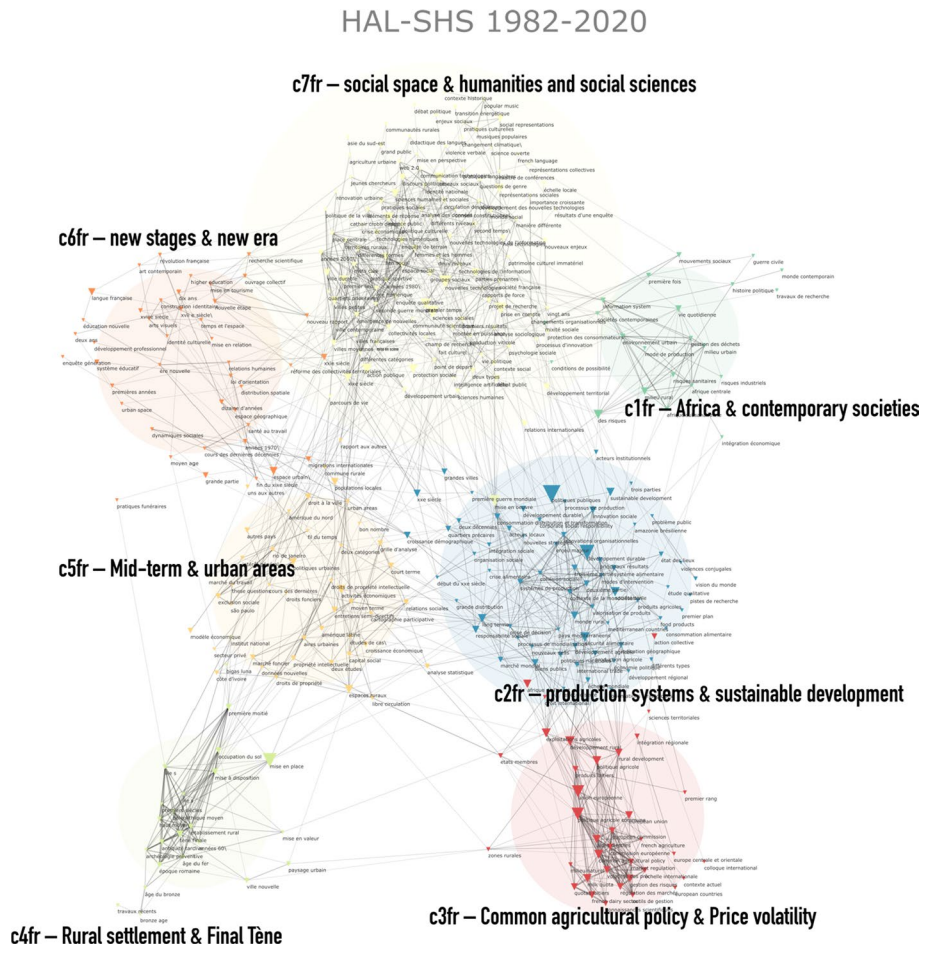

Fig. 5 French terms co-occurrence network

identify any particular actorsin this cluster, but it relates to specific places in the continent, such as Sub-Sahara or Central Africa.

In *production systems & sustainable development* (c2fr), co-occurrent terms show a similar structure to *c6eng* and *c3eng*. For French authors, this domain shows detailed accounts of processes such as food production and commercialization within a global scope. Corporate Social Responsibility is present, but not as markedly as its English counterpart. No global actors are mentioned in this cluster; only mentions of the rural world, food systems, and global markets are worth stressing. They allow grasping the kind of orientation of French-published research on the matter.

Also related to food production, *common agricultural policy & Price volatility* (c3fr) are emergent relations between different terms describing European agricultural policies. The most relevant actor here is the European Union, but there are also minor mentions to West Africa and France.

European rural settlements from prehistory to the Middle Ages, and historical approaches of land use, are the central preoccupation in *rural settlement & historical*

approaches (c4fr). A more narrow-sighted perspective appears in *mid-term & urban areas* (c5fr), stressing relations between urban policies, economic activities, and citizen rights in the city. Attention to the development of cities stresses here mid-term implications rather than deep historical roots. Local communities stand out as the most relevant actors. At the same time, various geographical settings emerge, such as Latin and North America, or specific places and countries such as Rio de Janeiro and Ivory Coast/Côte d'Ivoire.

A particular set of interests gathers here in *new stages & new era* (c6fr) in tight connection to understanding processes of change and revolution. It points toward relations between historical time—Middle Ages, XVII, XVIII, and XIX centuries—and future scenarios—XXI century. It focuses on objects such as the French language, urban space, tourism, scientific research, and the educational system. No relevant stakeholders worth mentioning are detected here.

Finally, *social space & humanities and social sciences* (c7fr) arises as perhaps the most diverse clusters in the French-published documents. It contains mostly a sociological set of interests, building strong bonds between different descriptions of social structure. It uses primarily qualitative approaches towards culture, political order, new digital technologies, and discourses. It also shows a focus mainly on a local scale and an urban setting. Urban related research emerges related to small and medium-sized towns, without any particular city being mentioned. New technologies appear as a relevant subject in relation to the social transformations. Regarding geographical entities, France and Southeast Asia are the main locations for this kind of research.

Comparison between languages

We observed comparatively the thematic domains resulting from queries that represent each thematic interest of the GRIP in English and French to identify convergence and divergence in the two bodies of literature. The most notable shared interest is *sustainable development*, which relates in French with production systems (c2fr) while in English, it relates to three different clusters focusing on energy consumption (c2eng), natural resources (c6eng), and corporate governance (c3eng). S Similarly, science and technology research is a self-standing interest in English (c4eng) while it is subsumed in a larger cluster in French (c7fr). Interest on Global Cities is also clearly identified in English (c1eng) while in French it relates to the broader concept of social space (c7fr). English publications are neatly divided into specialized clusters, when in French the same concerns are embedded into larger conceptual categories.

Our semi-automatic approach to text analysis has allowed us to observe a rather holistic approach towards Global Research in the French language. Thematics such as social space (c7fr) or sustainable development (c2fr) do not appear as singular objects but rather appear connected to a diverse collection of research objects within the cluster, indicating a tighter linguistic and semantic connections with the French corpus.

Descriptors of French documents show two other traits worth mentioning: the predominance of historical approaches, and the preeminence of Francophone countries and Latin America as research locations.

Finally, a clear difference between the two language bodies of references is the scale of the mentioned objects. For instance, when discussing Global Cities in English (c1eng), New York rises as the only relevant city, whereas the focus in French is on small and middle-sized towns. Something similar happens with actors mentioned in the clusters: the English-written literature mentions large international governance bodies (World Health

Organization, World Bank, United Nations) while French documents emphasize local communities and regional actors. Also, no specific actors arise as relevant term in French, but are referred to within a larger concept such as the rural world or food systems.

Discussion

Results have shown how similar topics can be addressed at different paces and with distinct thematic orientations in two different languages. As the volume of documents rises first in French and, later, in English, it is natural to assume that many of the ideas developed in one language get translated and adapted to the other. Our results on the semantic structure of the debate in each language confirm this "natural" pattern: the production in French has a tighter, firmer, and stronger underlying structure. Topics appear on a longer time frame, a fact we could relate to long prolonged and sustained tradition on these subjects with a steady amount of documents published in the first of the three periods. The last detected period of 15 years (2005–2020) speaks of the stabilization of the researchers' main concerns as reflected in the French documents. The delay between French and English-language publications on the same topics may be explained by a publication strategy where publishing English-language articles or books takes place after a maturation period that is visible in the previously published French-speaking literature. It has been suggested that French journals might be less selective than English-speaking journals. We doubt very much this assumption—and given the growth of number of journals worldwide in English, including the so-called predatory journals—we tend to believe that French-speaking journals, in particular in the social sciences are as selective as English-speaking.[3] But the cost of access for a French scientists in a French-speaking journal is lesser, simply because of language issues. Moreover the geographical analysis of global science collaborations through co-authorships shows that inter-institutional collaborations at the national level have strengthened (Maisonobe et al., 2016). Thus the preference of French as first publishing language, followed by English as a second choice corresponds to a strategy: matters of interest get first widely discussed in French and then published in English.

We could then use the metaphor of 'carry publishing' (as we talk of carry trading in the financial investment business) since we can consider the articles in French as a local investment maturing over time to benefit in the long term by a publication in English on the international arena. French science publications in all domains, including social sciences, have grown all over France, as result of active policies followed in the last 40 years to strengthen smaller cities, local universities, regional research centers rather than the larger metropolis (Milard & Grossetti, 2019). These new research units tend also to promote "international" publications, mainly in English, which correspond to the suggested 'carry publications', from a growing research population scattered in a wider range of cities and regions. Finally, in the promotion and research evaluation processes in France in the social sciences, the contention about English-language journals has been very intense, to the point that the national research evaluation authority (HCERES), as well as several evaluation commissions in national research organizations, have abandoned the use of pre-established lists of journals in several social sciences, or the use of bibliometric indicators.

[3] The personal experience of one of the authors (RA) as editor-in-chief of a high-impact French-speaking journal confirms this view, including the active role of local science communities and the acceptance of double-blind evaluation peer reviewing.

That entails implications in a better understanding of knowledge circulation dynamics and language interaction between English and non-English speaking scientific literatures (Kulczycki et al., 2018). This process takes place over the long term (Gordin, 2015) and relates to both scientific strategic choices and structural social constraints in the production of scientific knowledge (Hanafi & Arvanitis, 2014). The flat slope in publication rate from 2018 and on allows us to assume that these 'carry publishing' strategies have limits. Here, novelty appears as an inherent solid constraint since international interest will decrease over time.

When looking at the thematic orientation in French documents against those published in English, it becomes clear that these publishing practices have practical implications for the kind of research published in each language. The investment of knowledge and capabilities accumulated in French is not directly translated to English. Authors make changes when they present their research to an international audience reading in English and do not simply translate their work into English. Different objects are addressed at different scales in English, establishing a particular set of connections towards specialization rather than comprehensive approaches, as seen in French.

Regarding our data sources, some limitations are worth being mentioned. We have resorted to a novel data source such as HAL-SHS for the French part of our analysis. Although this is a rapid-evolving open science initiative, many normalization tasks are still lacking. These restrictions limit the possibilities of running traditional bibliometric analysis based on journals, authors, or institutions, as this information is not yet available or complete for records within HAL-SHS.

Further research should investigate these publication practices for non-English speakers; the particular case of French is of specific interest since it goes well beyond France (e.g., French-speaking Africa and North-African countries publish more in French than French scientists). How does the authorship structure develop in each language? What role authors' choices play in this process? How do research institutions, evaluation standards, and international competition (as reflected in universities' rankings) impact the publication process?

Conclusions

This research has shown that, on issues of global relevance, language matters. Specific interests can develop in a particular way within specific linguistic and geographic boundaries. The locally developed perspectives in French eventually migrate to the English language although serving different purposes, and not as simple translations from an original French-written version. Our results indicate a maturation process of thematic choices over time and different relative positions of these thematics when migrating from a local audience to an international English-speaking readership. This confirms previous observations showing that publishing in the English language in core-publishing journals confers a different value to academic knowledge than when it is published in the "local" language (Hanafi, 2011; Keim, 2016). Our case of French-speaking social sciences in a large University might convey a general pattern, but we cannot generalize to all social sciences in France, or in French-speaking countries, nor to other languages that are even more minority languages. To our knowledge, beyond observing that social sciences publish more heavily in their "national" languages than other scientific domains, little research has been done

on this fundamental issue of knowledge circulation in different languages and its impact on thematic choices (Gordin, 2015; Kulczycki et al., 2018, 2020; Ortiz, 2009).

Additionally, we suggest an empirical method that allows observing this circulation process from conceptualization in the local context in French and further dissemination through English-speaking publications. A strong interaction along the analytical process is needed to obtain fine-grained results and build query lines in the data-mining process. The interaction with researchers to build strategic-oriented query lines for information retrieval has shown a promising potential to map rising strategic research landscapes. Nevertheless, incorporating feedback can be complicated in the subsequent stages of the analysis. Since informants may not be fully aware of these language-processing techniques, they might have difficulties providing precise feedback when examining a resulting semantic map.

Lastly, the local perspective in our analysis allowed us to link the scientific debates at the national level with those appearing in English-speaking journals. It has been possible in the case of France, since local data were available, and because French social sciences also circulate beyond the national boundaries (see, e.g., 'French theory' in the USA). But as national repositories and databases are growing everywhere (see the databases such as Latindex in Latin America, or the new repositories in Asia, China, Russia, India), studying differences between locally published research in non-English speaking contexts and English-speaking international authors will be feasible and will probably reveal specific determinants that go beyond the need to diffuse more widely research results.

Appendix

See Tables 1, 2, 3, 4 and 5.

Table 1 Queries used to represent each thematic axis in WoS (English) and HAL-SHS (French)

Web of Science (English)	HAL–SHS (French)
Axis 1: global urbanities	
TS = ((((city OR cities) NEAR/2 (irregular migration)) OR ((city OR cities) NEAR/2 (migrant workers)) OR ((city OR cities) NEAR/2 (transnational labor))) OR ((((sanctuary city) OR (sanctuary cities)) OR (((municipal* OR local) NEAR/2 Government*) NEAR (*migration policy)) OR (asylum NEAR/2 (city OR cities)))) OR (((global OR world) NEAR/1 (city OR cities OR megacity OR megacities OR metropolis OR megalopolis))) OR (((((city OR cities OR urban) NEAR/2 (practice*)) OR ((city OR cities OR urban) NEAR/2 (ways of living))) AND (digital OR digitalization)))) AND CU = (FRANCE)	q = text_fulltext (((ville OR villes) "AND (cosmopolitisme OR nouvelle technologie OR globalisation OR Ségrégation OR ("intrusion exclusion" ~ 2))) OR ("droit à la ville") OR (("ville sanctuaire" ~ 2) OR ("ville accueillante")))
Axis 2: circulations	
TS = (((CSR OR (Corporate Social Responsibility)) AND (global* OR international* OR transnational*)) OR (((region* OR territor* OR local*) NEAR (development)) AND (agricultur* OR agribusiness OR industr* OR producers OR production OR productive OR productivity) AND (value chains OR export* OR liberalization OR (non NEAR tariffs) OR FDI OR (foreign investment) OR openness OR transnational OR international OR global* OR trade OR (economic cooperation))) OR (multilingualism) OR (((maker movement) OR (inventive activity) OR (history NEAR (technolog* OR innovation)) OR (thing turn))) OR (((remittances OR transnationalism) AND ((non-economic) OR political OR cultural OR scientific OR scientist* OR artist* OR (higher education) OR businessmen OR academic OR commercial))) OR (((global OR transnational OR international OR cross-national) AND (diffusion OR circulation OR flow* OR networks OR isomorphism) AND (((institutions OR institutional) NEAR model) OR (organizational practices) OR education OR teaching OR art OR culture OR music OR stantard* OR standardization OR certification)))) AND CU = (FRANCE)	q = text_fulltext (("mondialisation par le bas") OR (("nouvelle technologie" OR "nouvelles technologies") OR ("réseaux sociaux" AND "processus identitaire")) OR ("robotique humanoïde") OR ((circulation OR globalisation OR mondialisation) AND (management responsable OR RSE OR "Responsabilité Sociale des Entreprises")) OR (métropolisation OR "Aménagement du territoire" OR "dévelopement régional") OR (plurilinguisme) OR (transfert OR musique OR norme OR idée* OR savoirs OR argent OR politiques))

Table 1 (continued)

Web of Science (English)	HAL–SHS (French)
Axis 3: technologies, market logics and vulnerabilities	
TS = ((((STS OR "science studies" OR (science NEAR/2 technology NEAR/2 society)) AND ("global south" OR periphery OR "latin america" OR "latin-america" OR postcolonial OR Asia OR Africa OR India))) OR ((medicine OR health) AND (globalization OR transnational)) OR ((((“digital technologies” OR “information and communication technologies”) AND health) OR mHealth OR “m-Health” OR “eHealth” OR “digital health”) AND (gender OR women)) OR ((anthropocene OR globalization OR globalizing OR transnational) AND (pollution OR environment OR environmental)) OR (((“science” OR “scientific” OR “research”) AND ("crowdsourcing")) OR (“crowd science”) OR (“participative research”) OR (“participatory research”) OR (“citizen science”))) AND PY = (2009–2020) AND CU = (FRANCE)	q = text_fulltext (((STS OR “science studies” OR “science and technology studies” OR technoscience) AND (Sud OR Suds)) OR ((globalisation OR mondialisation) AND (medicines OR santé)) OR (((santé AND (numérique OR mobile)) OR (mhealth OR mSanté)) AND (maternelle OR sex* OR genre*)) OR ((globalisation OR mondialisation) AND (anthropocène OR environnement OR pollution)) OR (“Sciences participatives” OR “science participative” OR (“sciences citoyennes”) OR (“recherche participative”)))

Table 2 English detected terms, grouped by cluster

Cluster	Terms	Density
c1eng—global cities and collaboration networks	Urban areas, collaboration networks, city networks, international level, world cities, open innovation, scientific activities, international collaboration, urban space, globalization process, management education, property rights, participative research, economic geography, navigation satellite systems, various stakeholders, Regional Studies, innovation process, knowledge production, creativity and innovation, urban design, New York, intellectual property, global cities, Using data, French business schools, business schools, urban systems, Paul Krugman, transaction costs, regional development, regional science, New Economic Geography, key factors, social network analysis, territorial governance, business network, development projects, innovation policies, Global Responsibility, knowledge management, global network, innovation strategy, home country, learning process, innovation management, producer services, value chains, port cities, United Kingdom, air pollution, young adults, young people, conceptual model, health information, information system, last decade, use information, port areas, decent work, present article, socio-ecological system, theoretical model, local climate plans, information technology, global scale, different levels, alliance network, industrial district, economic system, growing interest, nineteenth century, disabled people, virtual reality, social justice, national contexts, industrial development, South African	0.89
c2eng—energy consumption and policy implications	Energy consumption, industrialized countries, carbon emissions, policy implications, causal relationship, social cohesion, trade openness, increasing globalization, life expectancy, environmental quality, manufacturing firms, Middle East, foreign direct investment, panel data, institutional quality, increasing number, digital technologies, developed economies, investment climate, factors influence, development aid, information and communication technologies, financial development, World Bank, empirical analysis, economic growth, new concept, positive effect, electricity consumption, CO2 emissions	0.9

Table 2 (continued)

Cluster	Terms	Density
c3eng—corporate governance and sustainability performance	Corporate social responsibility, financial performance, corporate social performance, sustainability performance, corporate governance, overall performance, sustainability reports, reporting practices, Social implications, agency theory, management control, environmental performance, corporate sustainability, stakeholder theory, institutional theory, academic literature, multinational corporations, value creation, international sample, financial markets, management accountants, CSR strategies, multinational enterprises, exploratory study, circular economy, CSR activities, CSR practices, global value chains, business ethics, emerging economies, firm performance, side effects, environmental practices, virtuous circle, CSR approach, local firms, GHG emissions, manufacturing sector, territorial servitization, chain management, international business, competitive advantage, information management, global environmental change, banking sector, global environment, greenhouse gas emissions, different cultures, financial variables, career success, developing world, wide range, political economy, product market, institutional work, nineteenth century, everyday life, French universities, patent applications, historical analysis	0.9
c4eng—science and technology and history of science	Science and technology, history of science, history of technology, academic entrepreneurship, web sites, French language, social sciences, social history, member states, social rights, nineteenth and twentieth centuries, eighteenth century, foreign language, labour market, more research, historical perspective, social psychology, higher education institutions, objective of this article, France and Germany, European Union, Second World War, Western Europe, new forms, social changes, labour law, Council of Europe, Inuit community, human rights, Middle Stone Age, diversity management, human history, education systems, Michel Foucault, scientific research, emerging countries, eighteenth century, EU Member States, short term, education policy, sustainability standards, language acquisition	0.89

Table 2 (continued)

Cluster	Terms	Density
c5eng—public health and global health	Public health, public health policies, non-governmental organizations, environmental health, chronic diseases, animal health, scoping review, reproductive health, World Health Organization, oil and gas, mental health care, gas industry, mental health, CSR issues, regional scale, special attention, focus groups, international conferences, health problems, global health, health interventions, regional level, health inequalities, social support, cultural context, risk management, health professionals, different stakeholders, cultural factors, wide variety, social movements, musculoskeletal disorders, collective action, risk assessment, World War II, cohort study, health issues, cultural diversity, population growth, long time, urban growth, economic environment, transportation systems, health services, health policies, metropolitan area, location choices, exchange rate	0.86
c6eng—sustainable development and natural resource management	Land use, natural resource management, land management, food production, participatory approach, social-ecological systems, agricultural policies, land systems, protected areas, rural development, spatial scales, farming systems, biodiversity conservation, future developments, water resources, agricultural development, recent decades, local development, food security, in-depth interviews, local governments, West Africa, environmental policies, systems approach, experiences of nature, citizen science program, traditional knowledge, Sustainable Development Goals, environmental management, different scales, land use changes, local stakeholders, territorial development, local people, raw materials, methodological framework, mining industry, sub-Saharan Africa, African countries, rural areas, economic changes, agricultural sector, decision makers, analytical framework, decision support, sustainable development, Latin American, network analysis, land access, public goods, citizen science project, new approaches, international community, water management, food systems, national level, new challenges, development models, local authorities, design process, geographic information, history of education, native language, United Nations, social innovation, life cycle assessment, twenty-first century, English language, international circulation, agricultural practices, turning point	0.9

Table 3 French detected terms, grouped by cluster

Cluster	Terms	Density
c1fr—Africa and contemporary societies	Central Africa, waste management, urban environment, urban environment, sub-Saharan Africa, rural environment, information system, RISKS, daily life, contemporary societies, political history, health risks, social movements, social mix, first time, Civil War, industrial risks, contemporary world, economic integration, conditions of possibility, mode of production, research work	0.87
c2fr—production systems and sustainable development	Production systems, product valorization, social cohesion, rural world, context of globalization, modes of intervention, organizational innovations, decision making, geographical indication, social innovation, food system, food security, mediterranean countries, agricultural production, globalization process, second part, main results, long term, third part, civil society, international trade, world market, agricultural products, consumption, distribution and transformation, new strategies, major stake, sustainable development, regional development, political economy, agricultural development, food products, social organization, Social Responsibility production process, precarious neighborhoods, global scale, two decades, beginning of the twentieth century, new challenges, Corporate Social Responsibility, social integration, sustainable development, large-scale distribution, Mediterranean countries, local actors, state of the art, twentieth century, national policies, sustainable development, public problem, domestic violence, demographic growth, food crisis, world view, greenhouse effect, large cities, public policies, food consumption, strategic orientations, Brazilian Amazon, public goods, World War I, three parts, first plan, qualitative study, institutional actors, different types, avenues of research	0.8
c3fr—common agricultural policy and price volatility	Common agricultural policy, price volatility, Common agricultural policy, market regulation, European Commission, direct aid, risk management, market regulation, milk quota, agricultural policy, management tools, international scale, milk quotas, French dairy sector, rural development, French agriculture, European Union, Rural Development, scientific knowledge, dairy products, natural environment, European Union, farms, rural areas, Member States, Central and Eastern Europe, regional integration, territorial sciences, international law, European countries, first rank, West Africa, collective action, current context, international conference	0.91
c4fr—rural settlement and historical approaches	Final Stone Age, rural settlement, third century, disposal, High Middle Ages, Late Antiquity, Middle Paleolithic, Iron Age, early centuries, first half, Roman Period, preventive archaeology, second century, 1960s, land use, settlement, Bronze Age, development, new town, National Institute, Bronze Age, urban landscape, recent work, Bigas Luna	0.95

Table 3 (continued)

Cluster	Terms	Density
c5fr—mid-term and urban areas	Urban areas, medium term, rural areas, Rio de Janeiro, these issues, urban policies, social exclusion, intellectual property rights, case studies, São Paulo, social capital, property rights, land market, urban areas, economic activities, free movement, two studies, North America, general interest, relationship to others, short term, new data, two categories, local populations, Ivory Coast, labor market, statistical analysis, time line, Latin America, other countries, semi-structured interviews, land rights, rural commune, participatory mapping, private sector, economic model, economic growth, large number, some to others	0.8
c6fr—new stages and new era	Funeral practices, professional development, Generation Survey, contemporary art, reform of the territorial communities, spatial distribution, Middle Ages, educational system, setting in relation, eighteenth century, two years, urban space, higher education, new education, French language, ten years, large part, geographical space, orientation law, international migration, French Revolution, early years, social dynamics, scientific research, ten years, urban space, cultural identity, last decades, seventeenth century, XXI century, time and space, human relations, collective work, visual arts, identity building, health at work, end of the nineteenth century, 1970s, new relationship, tourism, new stage	0.78

Table 3 (continued)

Cluster	Terms	Density
c7fr—social space and humanities and social sciences	Different forms, digital era, small towns, medium-sized towns, emergence of new, social link, Keywords, qualitative survey, 1980s, men and women, rural territories, 2000s, first place, new technologies, cultural policy, local authorities, World War II, French society, public space, information technologies, central place, power relations, nineteenth century, circulation of discourses, social networks, sociological analysis, different categories, rise in power, political discourse, different way, development of new technologies, social groups, social psychology, communication technologies, second time, data analysis, social world, first time, social representations, web 2.0, starting point, first results, digital technologies, gender issues, Constitutional Council, Popular Music, social context, popular music, social sciences, social practices, taking into account, elements of response, staging, public action, sustainable city, different levels, social representations, social sciences cultural practices, new information technologies, production, organizational changes, research project, contemporary city, priority districts, urban development, city policy, language practices, two types, twenty years, field survey, national identity, artificial intelligence, intangible cultural heritage, innovation process, consumer protection, new stakes, political life, sports practice, new technologies, economic crisis, implementation, urban renewal, Cathair Crobh Dearg, language didactics, growing importance, young researchers, public debate, perspective, verbal violence, stakeholders, climate change, research field, general public, cultural fact, life course, scientific community, urban agriculture, energy transition, survey results, French language, open science, rural communities, two levels, international relations, historical context, French cities, social protection, social issues, territorial development, collective representations, local scale, lecturer, Southeast Asia, political debate	0.9

Table 4 Full results of the period detection and significant term correlation between years in French

	1982	1984	1985	1986	1987	1988	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
1982	0	0.88	0.88	0.88	0.88	0.88	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95	0.95	0.95
1984	1.00	0	0.88	0.88	0.88	0.88	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95	0.95	0.95
1985	1.00	1.00	0	0.88	0.88	0.88	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95	0.95	0.95
1986	1.00	1.00	0.51	0	0.88	0.88	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95	0.95	0.95
1987	0.29	1.00	0.90	0.80	0	0.88	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95	0.95	0.95
1988	1.00	1.00	0.56	0.49	0.90	0	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95	0.95	0.95
1990	1.00	1.00	1.00	1.00	1.00	1.00	0	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.95	0.95	0.95	0.95	0.95
1991	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
1992	1.00	1.00	0.84	0.67	0.93	0.83	1.00	1.00	0	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
1993	1.00	1.00	0.87	0.73	0.95	0.86	1.00	1.00	0.86	0	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
1994	1.00	1.00	0.91	0.82	0.89	0.91	1.00	1.00	0.92	0.84	0	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
1995	1.00	1.00	0.71	0.79	0.91	0.62	1.00	1.00	0.73	0.84	0.92	0	0.91	0.91	0.91	0.91	0.91	0.91	0.91
1996	0.68	1.00	0.83	0.82	0.77	0.79	1.00	0.93	0.80	0.86	0.68	0.82	0	0.91	0.91	0.91	0.91	0.91	0.91
1997	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.91	1.00	1.00	0	0.91	0.91	0.91	0.91	0.91
1998	1.00	1.00	0.96	0.92	0.98	0.96	1.00	1.00	0.97	0.98	0.96	0.93	0.91	1.00	0	0.91	0.91	0.91	0.91
1999	0.99	1.00	0.97	0.93	0.98	0.96	1.00	1.00	0.90	0.88	0.90	0.97	0.79	0.93	1.00	0	0.91	0.91	0.91
2000	0.85	1.00	0.85	0.91	0.88	0.85	1.00	0.92	0.89	0.80	0.89	0.80	0.89	1.00	0.90	0.95	0	0.91	0.91
2001	1.00	0.79	1.00	1.00	1.00	1.00	1.00	1.00	0.79	0.95	0.96	0.95	0.92	0.92	0.98	0.98	0.86	0	0.91
2002	0.96	1.00	0.77	0.66	0.93	0.83	1.00	1.00	0.85	0.90	0.90	0.79	0.74	1.00	0.94	0.83	0.88	0.90	0
2003	0.99	0.96	0.88	0.94	0.99	0.95	0.95	0.86	0.91	0.89	0.86	0.95	0.81	1.00	0.97	0.95	0.85	0.88	0.84
2004	1.00	0.98	0.94	0.97	0.99	0.98	0.92	0.95	0.95	0.93	0.93	0.94	0.90	1.00	1.00	0.92	0.91	0.82	0.85
2005	0.89	0.96	0.82	0.91	0.80	0.89	0.91	0.98	0.90	0.92	0.87	0.88	0.74	1.00	0.91	0.95	0.84	0.82	0.81
2006	0.99	1.00	0.88	0.90	0.98	0.91	0.82	1.00	0.88	0.88	0.82	0.91	0.77	1.00	0.97	0.84	0.73	0.89	0.89
2007	0.96	0.98	0.90	0.96	0.89	0.92	1.00	0.86	0.92	0.87	0.82	0.91	0.75	0.96	0.93	0.89	0.79	0.90	0.79
2008	0.99	0.98	0.85	0.91	0.93	0.91	0.91	1.00	0.91	0.91	0.78	0.89	0.70	0.93	0.94	0.89	0.79	0.90	0.84

Table 4 (continued)

	1982	1984	1985	1986	1987	1988	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
2009	1.00	0.98	0.96	0.98	0.97	0.96	0.84	0.95	1.00	0.93	0.78	0.90	0.79	0.98	0.95	0.77	0.79	0.93	0.90
2010	0.97	0.96	0.93	0.92	0.83	0.89	0.93	1.00	0.89	0.86	0.83	0.90	0.68	0.96	0.83	0.84	0.68	0.89	0.88
2011	0.95	0.96	0.88	0.89	0.93	0.89	0.89	1.00	0.93	0.91	0.84	0.93	0.81	1.00	0.82	0.89	0.75	0.94	0.89
2012	0.94	0.98	0.86	0.86	0.93	0.93	0.98	0.93	0.90	0.89	0.77	0.84	0.68	1.00	0.91	0.84	0.80	0.91	0.81
2013	1.00	0.92	0.87	0.90	0.94	0.87	0.97	0.84	0.90	0.93	0.88	0.85	0.84	0.91	0.92	0.91	0.74	0.87	0.86
2014	0.98	0.95	0.79	0.83	0.91	0.82	0.92	1.00	0.84	0.84	0.86	0.80	0.67	0.91	0.87	0.77	0.71	0.85	0.73
2015	0.97	0.93	0.86	0.90	0.95	0.89	0.91	0.95	0.88	0.84	0.87	0.83	0.64	0.95	0.87	0.88	0.65	0.84	0.80
2016	0.97	0.95	0.88	0.88	0.91	0.93	1.00	0.90	0.90	0.88	0.83	0.97	0.83	0.97	0.89	0.85	0.70	0.86	0.86
2017	0.93	0.97	0.89	0.91	0.93	0.94	0.95	0.91	0.93	0.88	0.90	0.89	0.77	1.00	0.89	0.93	0.76	0.86	0.82
2018	0.97	0.96	0.75	0.88	0.93	0.91	0.99	0.96	0.86	0.82	0.85	0.83	0.70	0.91	0.78	0.88	0.77	0.92	0.78
2019	0.98	0.98	0.80	0.80	0.95	0.84	0.96	0.98	0.86	0.84	0.86	0.82	0.71	1.00	0.85	0.81	0.66	0.90	0.73
2020	0.95	0.97	0.89	0.92	0.95	0.89	0.97	0.88	0.90	0.85	0.90	0.83	0.67	1.00	0.82	0.88	0.77	0.89	0.78

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
1982	0.95	0.95	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1984	0.95	0.95	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1985	0.95	0.95	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1986	0.95	0.95	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1987	0.95	0.95	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1988	0.95	0.95	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1990	0.95	0.95	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1991	0.91	0.91	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
1992	0.91	0.91	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
1993	0.91	0.91	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
1994	0.91	0.91	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
1995	0.91	0.91	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88

Table 4 (continued)

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
1996	0.91	0.91	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
1997	0.91	0.91	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
1998	0.91	0.91	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
1999	0.91	0.91	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
2000	0.91	0.91	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
2001	0.91	0.91	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
2002	0.91	0.91	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
2003	0	0.91	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
2004	0.83	0	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
2005	0.84	0.47	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2006	0.78	0.85	0.72	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2007	0.79	0.83	0.70	0.65	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2008	0.83	0.80	0.67	0.50	0.68	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2009	0.85	0.83	0.72	0.58	0.76	0.54	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2010	0.81	0.87	0.71	0.61	0.70	0.57	0.56	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2011	0.89	0.90	0.74	0.71	0.77	0.71	0.67	0.57	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2012	0.86	0.88	0.77	0.67	0.72	0.77	0.68	0.61	0.63	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2013	0.81	0.85	0.79	0.71	0.72	0.72	0.66	0.59	0.73	0.62	0	0.65	0.65	0.65	0.65	0.65	0.65	0.65
2014	0.79	0.82	0.72	0.62	0.70	0.67	0.68	0.59	0.68	0.55	0.57	0	0.65	0.65	0.65	0.65	0.65	0.65
2015	0.77	0.83	0.69	0.64	0.59	0.59	0.66	0.58	0.69	0.55	0.52	0.48	0	0.65	0.65	0.65	0.65	0.65
2016	0.88	0.88	0.81	0.76	0.77	0.72	0.70	0.67	0.74	0.60	0.67	0.61	0.57	0	0.65	0.65	0.65	0.65
2017	0.86	0.86	0.76	0.73	0.79	0.71	0.74	0.63	0.74	0.68	0.66	0.62	0.55	0.64	0	0.65	0.65	0.65
2018	0.81	0.86	0.67	0.73	0.69	0.70	0.61	0.62	0.72	0.62	0.61	0.48	0.48	0.59	0.52	0	0.65	0.65
2019	0.80	0.88	0.78	0.64	0.68	0.67	0.68	0.61	0.70	0.59	0.60	0.47	0.48	0.48	0.57	0.51	0	0.65
2020	0.79	0.86	0.74	0.70	0.70	0.74	0.69	0.66	0.75	0.58	0.59	0.57	0.48	0.54	0.57	0.48	0.45	0

Table 5 Full results of the period detection and significant term correlation between years in English

	1980	1981	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
1980	0	0.95	0.85	0.85	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99
1981	0.95	0	0.85	0.85	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99
1991	1.00	1.00	0	0.85	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99
1992	0.94	1.00	1.00	0	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99
1993	1.00	1.00	1.00	1.00	0	0.98	0.98	0.98	0.98	0.98	0.98	0.95	0.95	0.95
1994	1.00	1.00	1.00	1.00	1.00	0	0.98	0.98	0.98	0.98	0.98	0.95	0.95	0.95
1995	0.62	1.00	1.00	0.98	1.00	1.00	0	0.98	0.98	0.98	0.98	0.95	0.95	0.95
1996	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0	0.92	0.92	0.92	0.96	0.96	0.96
1997	1.00	1.00	1.00	0.98	0.87	1.00	1.00	0.88	0	0.92	0.92	0.96	0.96	0.96
1998	1.00	1.00	1.00	0.98	1.00	1.00	0.97	0.97	1.00	0	0.92	0.96	0.96	0.96
1999	0.86	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.67	1.00	0	0.96	0.96	0.96
2000	1.00	1.00	1.00	1.00	0.92	1.00	1.00	1.00	1.00	0.91	0.83	0	0.91	0.91
2001	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.81	0	0.91
2002	0.89	1.00	1.00	0.98	1.00	1.00	0.97	1.00	1.00	0.93	0.96	0.75	0.90	0
2003	1.00	1.00	1.00	1.00	0.53	1.00	1.00	0.95	0.95	1.00	0.94	1.00	1.00	0.97
2004	0.87	1.00	1.00	0.99	1.00	1.00	0.97	1.00	0.97	0.28	0.96	1.00	1.00	0.91
2005	1.00	1.00	1.00	0.97	0.86	1.00	1.00	1.00	0.92	0.98	0.97	1.00	0.88	0.90
2006	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.99	0.98	0.98	0.98	0.95	1.00	1.00
2007	0.88	0.65	0.65	0.98	0.87	1.00	0.93	0.98	0.89	0.83	1.00	0.96	0.92	0.98
2008	1.00	1.00	1.00	1.00	1.00	1.00	0.96	0.97	0.97	0.97	0.96	0.97	0.93	0.95
2009	0.96	1.00	1.00	0.98	1.00	0.93	0.99	0.98	0.98	0.95	0.97	0.91	0.93	0.86
2010	1.00	1.00	1.00	0.98	1.00	0.93	1.00	0.58	0.75	0.93	0.86	0.92	0.92	0.95
2011	1.00	1.00	1.00	0.98	0.97	1.00	1.00	0.95	0.98	0.95	0.81	0.89	0.96	0.91
2012	1.00	1.00	1.00	0.99	0.98	0.96	0.94	0.94	0.96	0.96	0.98	0.94	0.96	0.87
2013	0.95	1.00	1.00	0.98	1.00	0.94	0.87	0.91	0.95	0.91	0.96	0.91	0.99	0.90
2014	0.97	1.00	1.00	0.95	0.95	1.00	1.00	0.92	0.97	0.95	0.95	0.91	0.86	0.91
2015	1.00	0.97	0.97	0.98	0.91	0.97	0.96	0.91	0.95	0.92	0.95	0.94	0.96	0.92
2016	0.97	0.98	0.98	0.98	0.92	1.00	0.98	0.89	0.91	0.92	0.89	0.89	0.88	0.90
2017	0.96	0.96	0.96	0.97	0.91	0.98	0.98	0.97	0.88	0.91	0.96	0.89	0.92	0.84
2018	0.86	1.00	1.00	0.98	0.92	0.98	0.97	0.92	0.95	0.93	0.97	0.89	0.96	0.83
2019	0.93	1.00	1.00	0.98	0.96	0.97	0.95	0.93	0.97	0.92	0.97	0.88	0.91	0.88
2020	0.97	0.98	0.98	0.96	1.00	0.95	0.94	0.91	0.95	0.92	0.94	0.93	0.93	0.92

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
1980	0.99	0.94	0.94	0.94	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
1981	0.99	0.94	0.94	0.94	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
1991	0.99	0.94	0.94	0.94	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
1992	0.99	0.94	0.94	0.94	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
1993	0.95	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
1994	0.95	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
1995	0.95	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
1996	0.96	0.92	0.92	0.92	0.92	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1997	0.96	0.92	0.92	0.92	0.92	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1998	0.96	0.92	0.92	0.92	0.92	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
1999	0.96	0.92	0.92	0.92	0.92	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93

Table 5 (continued)

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
2000	0.91	0.96	0.96	0.96	0.96	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
2001	0.91	0.96	0.96	0.96	0.96	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
2002	0.91	0.96	0.96	0.96	0.96	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
2003	0	0.96	0.96	0.96	0.96	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
2004	0.93	0	0.93	0.93	0.93	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
2005	0.99	0.93	0	0.93	0.93	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
2006	0.85	0.99	0.95	0	0.93	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
2007	1.00	0.85	0.92	0.97	0	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
2008	0.96	0.93	0.92	0.93	0.89	0	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71
2009	0.98	0.90	0.89	0.97	0.86	0.85	0	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71
2010	0.93	0.97	0.98	0.94	0.97	0.84	0.81	0	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71
2011	0.97	0.96	0.85	0.92	0.92	0.88	0.75	0.80	0	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71
2012	0.92	0.93	0.91	0.96	0.93	0.89	0.80	0.84	0.82	0	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71
2013	0.92	0.93	0.88	0.91	0.86	0.77	0.79	0.84	0.80	0.77	0	0.71	0.71	0.71	0.71	0.71	0.71	0.71
2014	0.91	0.95	0.91	0.87	0.94	0.85	0.71	0.78	0.80	0.81	0.74	0	0.71	0.71	0.71	0.71	0.71	0.71
2015	0.88	0.91	0.87	0.83	0.89	0.75	0.72	0.78	0.77	0.71	0.66	0.70	0	0.71	0.71	0.71	0.71	0.71
2016	0.86	0.89	0.72	0.89	0.81	0.79	0.67	0.70	0.73	0.74	0.71	0.59	0.58	0	0.71	0.71	0.71	0.71
2017	0.80	0.92	0.83	0.87	0.85	0.75	0.73	0.73	0.77	0.74	0.64	0.67	0.58	0.55	0	0.71	0.71	0.71
2018	0.80	0.91	0.82	0.88	0.86	0.77	0.73	0.80	0.80	0.73	0.67	0.67	0.67	0.56	0.57	0	0.71	0.71
2019	0.95	0.87	0.83	0.92	0.88	0.77	0.71	0.79	0.79	0.77	0.66	0.66	0.56	0.53	0.55	0.52	0	0.71
2020	0.88	0.94	0.89	0.87	0.86	0.74	0.75	0.75	0.75	0.73	0.68	0.68	0.52	0.52	0.64	0.52	0.54	0

Acknowledgements Research funded by the IdEx platform, from the Université de Paris, and the GRIP. A previous version of this paper was presented at the 18th International Society for Scientometrics and Informetrics Conference, that took place between 12 and 15 July 2021. We thank two acute anonymous reviewers for comments in a previous version of this article.

References

Aynaud, T. (2020). *Python-louvain, Louvain algorithm for community detection.* Retrieved from https://github.com/taynaud/python-louvain. Accessed 9 Oct 2021.

Barbier, M., Bompart, M., Garandel-Batifol, V., & Mogoutov, A. (2012). Textual analysis and scientometric mapping of the dynamic knowledge in and around the IFSA community. In I. Darnhofer, D. Gibbon, & B. Dedieu (Eds.), *Farming systems research into the 21st century: The new dynamic* (pp. 73–94). Springer.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008.

CCSD. (2021). *API HAL. API Archive Ouverte HAL.* Aarchive ouverte HAL-SHS (Sciences de l'Homme et de la Société).

Chi, P.-S. (2012). Bibliometric characteristics of political science research in Germany. *Proceedings of the American Society for Information Science and Technology, 49*(1), 1–6.
Chi, P.-S. (2014). Which role do non-source items play in the social sciences? A case study in political science in Germany. *Scientometrics, 101*(2), 1195–1213.
Cruz-Castro, L., & Sanz-Menéndez, L. (2018). Autonomy and authority in public research organizations: Structure and funding factors. *Minerva, 56*, 1–26.
Engels, T. C. E., Ossenblok, T. L. B., & Spruyt, E. H. J. (2012). Changing publication patterns in the social sciences and humanities, 2000–2009. *Scientometrics, 93*(2), 373–390.
Garfield, E. (1990). The languages of science revisited: English (only) spoken here? *Current Contents, 511*(1), 3–17.
Gordin, M. D. (2015). *Scientific Babel: How science was done before and after global English*. The University of Chicago Press.
Hanafi, S. (2011). University systems in the Arab East: Publish globally and perish locally vs publish locally and perish globally. *Current Sociology, 59*(3), 291–309.
Hanafi, S., & Arvanitis, R. (2014). The marginalization of the Arab language in social science: Structural constraints and dependency by choice. *Current Sociology, 62*(5), 723–742. https://doi.org/10.1177/0011392114531504
Keim, W. (2016). The international circulation of social science knowledge. *Revue D'anthropologie Des Connaissances*. https://doi.org/10.3917/rac.030.0042
Kulczycki, E., Engels, T. C. E., Pölönen, J., Bruun, K., Dušková, M., Guns, R., Nowotniak, R., et al. (2018). Publication patterns in the social sciences and humanities: Evidence from eight European countries. *Scientometrics, 116*(1), 463–486.
Kulczycki, E., Guns, R., Pölönen, J., Engels, T. C. E., Rozkosz, E. A., Zuccala, A. A., Bruun, K., et al. (2020). Multilingual publishing in the social sciences and humanities: A seven-country European study. *Journal of the Association for Information Science and Technology, 71*(11), 1371–1385.
Lambiotte, R., Delvenne, J.-C., & Barahona, M. (2014). Laplacian dynamics and multiscale modular structure in networks. *IEEE Transactions on Network Science and Engineering, 1*(2), 76–90.
Leydesdorff, L., & Hellsten, I. (2006). Measuring the meaning of words in contexts: An automated analysis of controversies about "Monarch butterflies", "Frankenfoods", and "stem cells." *Scientometrics, 67*(2), 231–258.
Maisonobe, M., Eckert, D., Grossetti, M., Jégou, L., & Milard, B. (2016). The world network of scientific collaborations between cities: Domestic or international dynamics? *Journal of Informetrics, 10*(4), 1025–1036.
Meneghini, R., & Packer, A. L. (2007). 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, 8*(2), 112–116.
Milard, B., & Grossetti, M. (2019). Excellence everywhere? Regional development of French scientific output and visibility. *Regional Studies, 53*(10), 1459–1469. https://doi.org/10.1080/00343404.2019.1581362
O'Brien, D., & Arvanitis, R. (2019). The Transformation of Research in the South: An introduction. In R. Arvanitis & D. O'Brien (Eds.), *The transformation of research in the south* (pp. 1–6). IRD Éditions.
Ortiz, H. (2009). Embarking on financial industry research: A response to anthropologists' common methodological concerns. *Anthropology News, 50*(7), 25–25.
Ràfols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology, 61*(9), 1871–1887.
Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (statistical Methodology), 63*(2), 411–423.
Trow, M. A. (1999). *Biology at Berkeley*. Center for Studies in Higher Education.
Wallace, M., & Ràfols, I. (2015). Research portfolio analysis in science policy: Moving from financial returns to societal benefits. *Minerva, 53*(2), 89–115.
Wallace, M., & Ràfols, I. (2018). *Institutional shaping of research priorities: A case study on avian influenza*. Social Science Research Network.
Zitt, M., & Bassecoulard, E. (1994). Development of a method for detection and trend analysis of research fronts built by lexical or cocitation analysis. *Scientometrics, 30*(1), 333–351.