

Document de travail N° 2018-05

Axe : Macroéconomie Appliquée Finance et Mondialisation

La dynamique des SHS françaises dans le Web of Science : un manque de représentativité ou de visibilité internationale ?

Abdelghani Maddi,
maddiabdelghani@gmail.com
CEPN, UMR-CNRS 7234, Université Paris 13, Sorbonne Paris Cité

Aouatif De La Laurencie,
aouatif.de-la-laurencie@hceres.fr
OST, HCERES

Novembre 2018

Mots clés : publications scientifiques, Web Of Science, SHS, bibliométrie, spécialisation scientifique

Codes JEL : D83

La dynamique des SHS françaises dans le Web of Science : un manque de représentativité ou de visibilité internationale ?

Aouatif De-La-Laurencie, OST, Hcéres
Abdelghani Maddi, CEPN, Paris 13 : OST, Hcéres

Résumé

Notre étude analyse la dynamique des sciences humaines et sociales (SHS) françaises dans le Web of Science (WOS), comparée à celle de quatre pays européens non anglophones, à savoir, l'Allemagne, l'Espagne, l'Italie et les Pays-Bas. Le point de départ de notre étude est le fait que les SHS françaises semblent doublement pénalisées dans les bases de données internationales à l'instar du WOS ou de Scopus. Premièrement, ces bases de données recensent majoritairement des articles publiés dans des revues à comité de lecture (une condition pour qu'une revue soit intégrée dans ces bases de données). Or, les pratiques de publications en SHS induisent que les articles ne soient pas majoritaires (avec une certaine disparité parmi les disciplines). Secondement, les bases de données internationales contiennent un biais de surreprésentation de l'anglais au détriment des autres langues.

Pour cette étude, nous avons utilisé les données issues de la base de données WOS entre 2000 et 2015. Nos premiers résultats indiquent que la dynamique de publication en SHS est très contrastée selon les pays. Certains pays présentent une dynamique très forte à l'instar de l'Espagne qui croît de façon exponentielle. La France est le pays dont le nombre d'articles en SHS progresse le moins.

Pour tenter de fournir une explication à cette observation, nous avons émis deux hypothèses. La première concerne l'indicateur de spécialisation. Ce dernier permet de mesurer la part d'une discipline dans l'ensemble des publications d'un pays ; plus la part est importante, plus le pays est dit spécialisé dans la discipline en question. Notre hypothèse consiste à dire qu'étant spécialisée dans des disciplines SHS qui publient peu d'articles, la France a peu de publications recensées et est peu visible dans les bases internationales. Symétriquement, les disciplines de non spécialisation de la France publient beaucoup d'articles, et comme la France n'est pas spécialisée, elle ne sera pas bien représentée également. La seconde hypothèse est liée à la langue de publication. Selon cette hypothèse, les chercheurs français ont tendance à publier plus en leur langue nationale et peu en anglais, contrairement aux chercheurs dans les autres pays.

L'analyse des types de document des disciplines de spécialisation et de non spécialisation de la France indique que la part des articles de la France est beaucoup plus élevée que celle du monde pour toutes les disciplines. Autrement dit, la France publie plus d'articles que la moyenne mondiale dans toutes les disciplines SHS qu'elle y soit spécialisée ou non. Ce résultat nous mène à poser la question suivante : Est-ce que seuls les articles français passent la barrière à l'entrée du WOS ?

L'étude de la langue de publication montre que la part de l'anglais diffère selon les pays. La France présente la part de publications en anglais la plus faible. Les Pays-Bas ont la part la plus importante avec une part qui s'approche de 100%, suivis de l'Italie ayant une part qui avoisine 90%. Pour ce qui est de la dynamique, plusieurs comportements sont observés. Les Pays-Bas et l'Italie présentent une part assez stable entre 2000 et 2015. L'Allemagne et la France augmentent la part de leurs publications en anglais avec une évolution plus prononcée pour l'Allemagne. L'Espagne a un comportement assez particulier. En plus de

l'augmentation des publications en anglais, L'Espagne a très fortement accru ses publications en sa langue nationale dans le WOS contrairement à la France. Les résultats sur les langues vérifient notre seconde hypothèse. La France étant le dernier pays en termes de part de publications en anglais peine à affirmer sa présence dans le WOS.

Abstract

Our study analyzes the dynamics of France's scientific production in the social sciences and humanities (SSH), compared to that of four non-English European countries, namely, Germany, Spain, Italy and the Netherlands. Our analysis is based on the observation that the SSH of these countries are doubly penalized in international databases like the WOS or Scopus. Firstly, because the nature of SSH means that in these disciplines articles are not in the majority (with disparities depending on the discipline). However, bibliometric databases mainly include articles published in peer-reviewed journals (a condition for a journal to be included in these databases). Secondly, international databases contain a bias of overrepresentation of English to the detriment of other languages.

To do this, we used the data from the WOS database. Our first results indicate that the dynamics of SSH are very contrasting according to the countries. Some countries have a very strong dynamic, like Spain, which is exponentially growing. France is the country with the least progress in SSH.

In an attempt to provide an explanation for this observation, we made two assumptions. First, we assume that France, as it is specialized in disciplines which are likely to publish few articles, has few recorded publications in International databases. Symmetrically, France's non-specialist fields publish a lot of articles, and since France is not specialized, it will not be well represented as well. Second, we assume that French authors tend to publish more in their national language and few in English, unlike the other countries.

The analysis of the structure of the specialization and non-specialization disciplines of France indicates that the share of French articles is much higher than that of the world for all disciplines. In other words, the French publish more articles than the world average in all SSH disciplines whether specialized or not. This result leads us to ask the following question. Do only French articles pass the barrier at the entrance of the WOS? Or is it easier to write an article in English than a book for French researchers?

The study of the language of publication shows that the part of English differs according to the country. France has the lowest proportion of publications in English. The Netherlands has the largest share with a share approaching 100%, followed by Italy with a share of around 90%. In terms of dynamics, several behaviors are observed. Apart from the two countries, with the largest share with a fairly stable share between 2000 and 2015, Germany and France increase the share of their publication in English with a more pronounced evolution for Germany. Spain has a rather particular behavior. In addition to publications in English, Spain has greatly increased its publications in its national language in the WOS unlike France. The results of the second hypothesis seem to answer, but in part, to our questioning. France being the last country in terms of share of publications in English struggles to assert its presence in the WOS.

Introduction

Les indicateurs bibliométriques utilisés dans l'évaluation d'acteurs de recherche en sciences humaines et sociales (SHS) soient constamment contestés en France, jugés trop réducteurs de l'évaluation de leur production. L'utilisation des indicateurs est un enjeu majeur pour ces disciplines dans la mesure où l'évaluation quantitative de la recherche et des chercheurs s'impose de plus en plus. Cette situation est due au fait que les calculs sont effectués à partir des bases de données qui présentent des biais, telles que le WOS ou Scopus qui recensent majoritairement des articles publiés dans des revues scientifiques et des actes de colloques, au détriment d'autres types de produits (ouvrages, monographies, chapitres d'ouvrage, traductions, etc.). Ces bases de données favorisent également les publications écrites en anglais au détriment des autres langues. Les indicateurs calculés à partir de ces bases ne permettent pas de mesurer la production réelle en SHS car elles ignorent la spécificité des pratiques de publication dans ces disciplines.

Cependant, malgré ces problèmes de représentativité des bases de données internationales, certains pays, non anglophones, enregistrent une forte progression de leurs publications en SHS dans ces bases. L'Espagne, l'Italie ou encore l'Allemagne enregistrent une forte croissance de leurs publications en SHS dans le WOS. À l'opposé, la France enregistre une croissance moins importante dans le WOS alors qu'elle compte une grande communauté de chercheurs dans ces disciplines.

L'objectif de cette étude est d'analyser les raisons de cette progression relativement faible des SHS françaises comparées à celles des principaux pays européens non anglophones. Pour ce faire, nous examinons deux hypothèses en lien avec les biais du WOS en faveur des articles de revues d'une part et de l'anglais d'autre part.

La première hypothèse consiste à dire que la France est spécialisée dans des disciplines qui publient peu d'articles et, inversement, peu spécialisée dans les disciplines qui publient beaucoup d'articles. C'est la raison pour laquelle la production française est moins présente dans les bases de données bibliométriques recensant majoritairement des articles.

La seconde hypothèse tient compte du biais linguistique. Elle peut être formulée comme suit : les chercheurs français en SHS publient en langue nationale plutôt qu'en anglais, contrairement aux chercheurs des autres pays européens non anglophones qui ont tendance à publier en anglais. Les publications françaises sont donc moins visibles.

Pour vérifier ces deux hypothèses, nous avons identifié les disciplines de spécialisation et de non spécialisation de la France en SHS. Nous avons analysé la distribution des types de production par disciplines dans le WOS afin de comparer la part des articles au niveau mondial, et pour la France spécifiquement. Nous avons analysé l'évolution du volume et de la part des publications écrites en anglais de la France, tout en la comparant avec les principaux pays non-anglophones ayant une langue nationale forte ; le référentiel pays compte la France, l'Allemagne, l'Espagne et l'Italie, pays européens non anglophones de taille comparable. Les Pays-Bas sont ajoutés pour les indicateurs indépendamment du volume de publications.

Notre objectif est d'analyser la question d'un point de vue descriptif afin de caractériser les spécificités, disciplinaires et linguistiques, de la France par rapport à ses principaux « concurrents » scientifiques. Il ne s'agit ici pas dans ce travail d'analyser l'impact des

politiques publiques sur la recherche ni l'efficacité des choix scientifiques stratégiques en matière de recherche.

La première présente une revue de littérature sur les spécificités des SHS et leur présence dans les bases de données internationales. La seconde section cadre l'analyse en présentant l'évolution des volumes et de spécialisation des pays sélectionnés. Les sections trois et quatre vérifient respectivement les deux hypothèses présentées ci-dessus. Enfin, la discussion des résultats pose de nouvelles questions comme pistes pour de futurs approfondissements.

1. Revue de littérature : spécificités des SHS et couverture de bases de données

1.1. Spécificités des sciences humaines et sociales

En sciences humaines et sociales, Hicks (1999, 2004) souligne que les livres jouent un rôle central dans la diffusion des connaissances, contrairement aux sciences de la matière et de la vie. Hicks (2004) distingue quatre principaux types de produits. Les articles publiés dans des revues internationales. Les ouvrages qui constituent l'unité de savoir la plus courante en sciences sociales. Ils viennent en deuxième position après les articles en termes d'utilisation dans l'analyse bibliométrique car leur recensement dans les bases de données est très incomplet. Les articles ayant une orientation nationale constituent le troisième type de produits en sciences sociales. Ils sont, en général, publiés dans des revues nationales dont le lectorat est plus restreint. Leur impact académique est relativement faible, mais leur impact social peut être élevé¹ (Gruzd *et al.* 2013 ; Mohammadi *et al.* 2013, 2014, 2015a, 2015b ; Mohammadi, 2014 ; Chen *et al.* 2015). Il en va de même pour la littérature non-académique (ou les produits de vulgarisation) qui constitue le quatrième et dernier type des produits scientifiques en sciences sociales définis par Hicks. La littérature non-académique telle que les documentaires, les articles de presse, les conférences et séminaires de vulgarisation est destinée à un public large.

Nederhof (2006) définit trois autres caractéristiques spécifiques aux SHS relatives à la forme des écrits scientifiques, au public auquel sont destinés les travaux de recherche et aux pratiques de collaboration entre les chercheurs. Les développements théoriques et la revue de littérature sont souvent importants en SHS. Les travaux de recherche en SHS sont destinés à un public plus large. Enfin, la recherche en SHS est moins collaborative qu'en sciences de la matière et de la vie car le nombre d'auteurs par publication est relativement faible. Nederhof conclut sur le fait que les méthodes bibliométriques utilisées en sciences de la matière et de la vie peuvent être utilisées pour évaluer les publications en SHS, à une double condition : développer la couverture des revues et des publications non recensées dans le WOS ; et élargir la part des livres dans l'analyse. Cette recommandation ne s'applique pas de la même façon à l'ensemble des disciplines, mais dépend des pratiques de publications et donc du taux de couverture des différents types de produits existant dans chaque discipline.

Il existe des disciplines en SHS ayant des comportements de publication et de citation similaires à ceux des sciences de la matière et de la vie. Ainsi, en économie, gestion, psychologie et sciences de l'information, les articles jouent un rôle important dans la diffusion des connaissances (Hicks, 2004 ; Huang *et al.*, 2008). Il en va de même selon McDonald

¹ Pour une revue détaillée sur les travaux étudiant l'impact social des publications scientifiques, voir : Sugimoto *et al.*, 2017.

(2003) en philosophie. En effet, il a étudié le comportement de citation dans trois disciplines : philosophie, histoire et littérature. Sur un échantillon de 173 publications totalisant près de 12 000 citations, les livres (21 sur 173) rassemblent plus de 59% de l'ensemble des citations. Cette moyenne masque des comportements différents : en philosophie, 44% des citations renvoient à des livres et 54% à des articles, alors qu'en histoire et en littérature, ces parts sont respectivement de 57% / 39% et 78% / 16%.

Knievel et Kellsey (2005) fournissent une analyse comparative sur huit disciplines : philosophie, musique, religion, littérature, linguistique, études de l'antiquité, philosophie et histoire. Leur analyse porte sur plus de 9 000 références d'articles de recherche publiés en 2002. Elles montrent qu'au sein des sciences humaines il existe des divergences dans les pratiques de citation. En moyenne, les citations de livres représentent 76%. Ce taux varie entre 51% (philosophie) et 89% (religion). Il existe en effet des disciplines dans lesquelles les articles de recherche publiés dans des revues jouent un rôle central dans la diffusion des connaissances, notamment la philosophie (51%) et la linguistique (61%). Larivière et al (2006) montrent qu'au sein des SHS, la part des d'articles est croissante : elle est passée de 40 à 48% entre 1981 et 2000, mais plus précisément de 60 à 70% en psychologie et de 45 à 57% en économie. Pour les autres disciplines le taux est relativement stable. L'histoire, la littérature et la catégorie « autres sciences humaines » citent beaucoup moins les articles, avec des taux, respectivement, de 34%, 26% et 21%.

Heinzkill (2007) analyse 20 000 références de 555 articles publiés dans des revues littéraires. Ses résultats indiquent que les citations des livres représentent plus de 78% dont 45% de références citées ont plus de 20 ans. Dans ce cas, il faut tenir compte du fait qu'une partie de ces citations renvoient aux œuvres originales étudiées, ce qui est une particularité des études littéraires.

En somme, les produits de recherche en SHS peuvent être scindés en quatre principaux groupes : les ouvrages (incluant les ouvrages individuels, les chapitres d'ouvrages, directions d'ouvrages, ouvrages collectifs, etc.) ; les articles scientifiques (publiés dans des revues à comité de lecture ou sans comité de lecture) ; les actes de conférences, communications et séminaires ; enfin, les produits de vulgarisation (tout produit destiné à des non spécialistes). Cette conclusion sur les produits de recherche peut également s'appliquer en sciences de la matière et de la vie, bien que la proportion des autres produits (différents des articles) soit relativement faible.

1.2. Couverture des bases de données : quelle place pour les SHS ?

Les différents types de publication en SHS ne sont pas bien représentés dans les bases de données internationales comme le WoS et Scopus, dont les processus de sélection privilégient l'article dans des revues à comités de lecture et les publications en anglais ou disposant au moins un résumé en anglais.

Selon Leydesdorff (2003) 55% des références dans la base de données répertoriant les publications en sciences sociales à savoir le Social Sciences Citations Index (SSCI) ne sont pas indexées dans le WOS, contrairement aux sciences naturelles par exemple qui comptent un taux de couverture des références (dans le SCI) de 79%. C'est également le cas des sciences humaines, mais de façon plus accentuée. Comme montré par Knievel et al. (2005), les articles de recherche viennent loin derrière les livres avec un taux de citations de 25% contrairement aux livres ayant un taux de 74%. Les données de Butler (1998) sur les sciences sociales australiennes indiquent une corrélation forte et négative (-0.83) entre la part des

articles indexés dans le SSCI et la part des livres dans une discipline. Autrement dit, plus le nombre de livres augmente dans une discipline, moins la couverture de cette discipline dans le SSCI est forte.

Mongeon et Paul-Hus (2016) comparent le WoS et Scopus en prenant la base Ulrich comme référence. Ils proposent une analyse comparative sur trois volets, disciplinaire (en quatre grandes disciplines), géographique et linguistique. Sur un ensemble de 63 013 revues académiques recensées dans Ulrich, le WoS recense 13 605 revues, tandis que Scopus en recense 20346. Le WoS et Scopus contiennent tous deux des biais géographiques et linguistiques avec une surreprésentation des revues et pays anglophones. Quant au type de publication, les deux bases recensent majoritairement des articles, ce qui peut poser problème notamment en sciences humaines et sociales où les pratiques de publication varient selon les disciplines. Les deux bases de données présentent des divergences notables, Scopus est beaucoup plus large que le WoS, ce qui conduirait à des résultats différents lors de l'analyse. En recherche biomédicale tout comme en sciences humaines et sociales, le WoS peut être considéré comme un sous-ensemble de Scopus car la quasi-totalité des revues recensées par le WoS sont également présentes dans Scopus. Les revues recensées uniquement par Scopus sont de l'ordre de 42 % en science de la matière et de la vie et en ingénierie, 58 % en recherche biomédicale, 61 % en sciences sociales et 50 % en sciences humaines. Les deux bases de données partagent en moyenne 42 % de revues communes, toutes disciplines et tous pays confondus.

Concernant les revues françaises, 80 % des revues de la base Ulrich en sciences sociales ne sont pas recensées dans le WoS. Ce taux est d'environ 58 % en sciences humaines, 43 % en science de la matière et de la vie et en ingénierie, 60 % en recherche biomédicale. Le même constat s'applique pour les revues francophones : Scopus a une plus large couverture des publications écrites en français que le WoS.

Harzing et al. (2017a) et Harzing et al. (2017b) ont comparé la couverture de quatre bases de données bibliométriques (WoS, Scopus, Google Scholar et Microsoft Academic). Leur échantillon est formé des publications de 145 chercheurs représentant les cinq principaux domaines de la science (sciences, sciences de la vie, ingénierie, sciences sociales et sciences humaines). Ils concluent que Microsoft Academic constitue une « excellente alternative » pour l'évaluation bibliométrique, à condition d'améliorer la couverture des livres notamment en SHS, ainsi que la qualité des métadonnées.

Hug et Brändle (2017) fournissent une étude comparative beaucoup plus large en analysant la couverture des trois bases de données, Microsoft Academic, WoS et Scopus. Ils analysent la couverture des publications des chercheurs de l'Université de Zurich présentes dans l'archive ouverte « Zurich Open Archive and Repository (ZORA) » créée en 2006. Ils analysent un sous ensemble de ZORA constitué de 62 791 publications publiées entre 2008 et 2015. Leurs résultats montrent que les trois bases de données ont des taux de couverture similaires. Pour les types de document, ils montrent que les trois bases de données recensent majoritairement des articles publiés dans des revues. Par contre, Microsoft Academic a une meilleure couverture des autres types de documents que le WoS ou Scopus, notamment les actes de conférences. Scopus a la part la plus élevée des articles que les deux autres bases de données. Concernant la couverture par discipline, les taux de couverture sont assez similaires entre les trois bases. Cependant, Microsoft Academic a une meilleure couverture des sciences humaines et sociales que les deux autres bases de données.

En résumé, les sciences humaines et sociales sont les moins représentées avec des taux de couverture relativement faibles comparés à ceux des sciences de la matière et de la vie. Cela est dû notamment aux spécificités des sciences humaines et sociales disposant d'un éventail de produits scientifiques qui ne se limite pas aux articles publiés dans des revues à comité de lecture. Par ailleurs, les SHS ont une orientation nationale et régionale.

2. Panorama de la production en SHS dans des pays européens non anglophones

L'objectif de cette section est de fournir des éléments de cadrage liés à l'évolution des volumes de publications en SHS ainsi qu'à l'évolution de la spécialisation dans ce domaine.

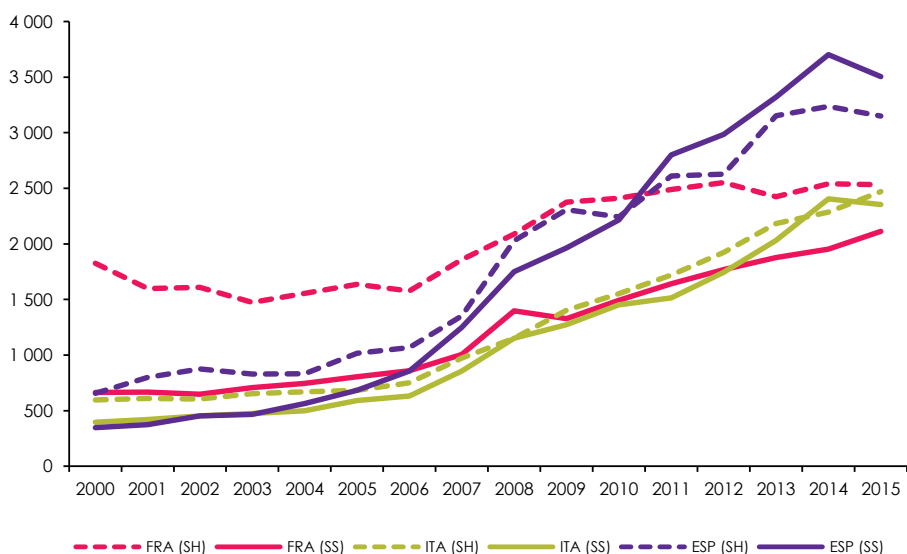
Au niveau mondial, le nombre de publications a quasiment doublé en SH (sciences humaines) dans le WOS passant de 180994 publications en 2000 à 320536 en 2015. Les publications en SS (sciences sociales) ont eu une évolution plus forte ; le nombre de publications a augmenté de 203% sur la période (378640 en 2015 contre 124903 publications en 2000).

a. Indicateurs de production pour le référentiel

Le graphique 1 montre une évolution spectaculaire du nombre de publications de l'Espagne en sciences sociales (SS) et en sciences humaines (SH)². Elle approche 4 000 publications par an en SS et 3000 en SH, alors qu'en 2000, l'Espagne se situait loin derrière la France en SH (600 publications contre près de 2 000 pour la France). L'Espagne a multiplié son nombre de publications par 10 sur la période 2000-2015. À l'opposé, la France enregistre une faible croissance et se fait dépasser rapidement, en plus de l'Espagne pour les deux disciplines, par l'Italie en SS à partir de 2012. L'Italie enregistre une multiplication du nombre de publications par près de 7 au cours de la période.

² Une publication est comptabilisée pour un pays si au moins un des auteurs de la publication a une adresse dans ce pays.

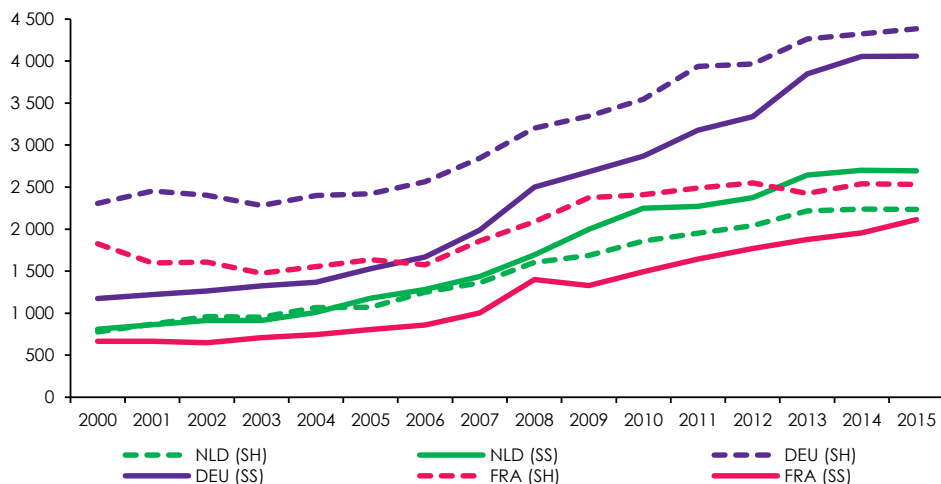
Graphique 1 : Nombre de publications en SHS, France, Italie et Espagne, 2000-15



Source : Rapport de l'OST « La position scientifique de la France dans le monde 2000-2015 »

Le graphique 2 montre que l'écart entre l'Allemagne et la France en SS s'est creusé entre 2000 et 2015. C'est également le cas pour les SH. L'Allemagne, avec un volume de publications relativement important, progresse sur la période. Il en va de même pour les Pays-Bas qui, dépassés par la France en volume en SH en 2000, réduisent leur écart entre 2000 et 2015. À l'opposé, les Pays-Bas ont un volume de publications en SS plus élevé que celui de la France depuis 2000. L'écart entre les deux pays s'est creusé également sur la période suite à une progression plus forte des hollandais comparés aux français.

Graphique 2 : Nombre de publications en SHS, France, Pays-Bas et Allemagne, 2000-15



Source : Rapport de l'OST « La position scientifique de la France dans le monde 2000-2015 »

La France a ainsi moins progressé en nombre de publications que les pays du référentiel, au cours des 15 premières années du 21^e siècle dans les disciplines des SHS.

b. Indices de spécialisation pour le référentiel

Le tableau 1 fournit les indices de spécialisation des pays du référentiel en 2000 et 2015, ainsi que le pourcentage de l'évolution, pour les deux disciplines SS et SH.

L'indice de spécialisation scientifique d'un pays dans une discipline est défini par la part de la discipline dans les publications du pays, normalisée par cette même part de la discipline dans les publications mondiales. Plus l'indice de spécialisation est supérieur à 1, plus le pays est dit « spécialisé » dans la discipline considérée.

Tableau 1 : indices de spécialisation en SH et SS du référentiel, 2000 et 2015

Pays	Spécialisation en sciences humaines			Spécialisation en sciences sociales		
	2000	2015	Évolution SH	2000	2015	Évolution SS
NLD	1,09	1,80	+66%	1,18	1,55	+32%
ESP	0,72	1,46	+104%	0,40	1,17	+194%
DEU	0,87	1,16	+32%	0,46	0,77	+66%
FRA	0,95	0,97	+2%	0,36	0,58	+61%
ITA	0,47	0,96	+105%	0,32	0,65	+103%

Source : Base OST, Web of Science, calculs OST

Le tableau 1 montre que les niveaux et les évolutions des indices de spécialisation varient sensiblement au sein du référentiel. Une très forte progression est observée pour l'Espagne dans les deux grandes disciplines ; SS et SH. L'indice de spécialisation de ce pays passe ainsi de 0,72 à 1,46 en SH, soit une progression de 104%. La France était plus spécialisée que l'Espagne en 2000, avec un indice égale à 0,95 et son indice n'a quasiment pas bougé sur la période car il est de 0,97 en 2015 (soit une évolution de 2%). La situation se renverse en 2015 entre les deux pays en termes de spécialisation en SH. Les Pays-Bas renforcent également leur spécialisation ainsi que l'Allemagne. L'Italie, bien qu'elle ait un indice de spécialisation très faible en 2000, a rattrapé la France en 2015, avec un indice de spécialisation de 0,96. Soit une progression de 105%, équivalente à celle de l'Espagne (104%).

En SS, la tendance se confirme pour l'Espagne dans une plus grande mesure, passant d'un indice de spécialisation de 0,40 à 1,17 entre 2000 et 2015, soit une multiplication par 4. La France augmente de 66% son indice de spécialisation en SS mais reste en-dessous de la moyenne mondiale. C'est également le cas pour l'Allemagne. L'Italie enregistre une plus forte progression et se place juste derrière l'Espagne en termes d'évolution. Les Pays-Bas, déjà spécialisés en 2000, renforcent leur spécialisation en 2015, passant de 1,18 à 1,55 (soit une évolution de 32%).

Dans les deux sections qui suivent, nous verrons sur le cas de la France qui présente une croissance limitée avec des indices de spécialisation faibles pour les deux grandes disciplines des sciences sociales et sciences humaines. Autrement dit, l'hypothèse de « sous-représentation » dans le WOS s'applique moins aux autres pays de notre référentiel car ils ont une dynamique forte dans ses disciplines dans le WOS. La France semble davantage impactée par le biais de la base de données. Raison pour laquelle nous n'avons pas intégré les pays du référentiel dans la section 3 qui concerne les types de production.

3. Quel type de production pour les SHS françaises le WOS?

Dans cette section, nous examinons la première hypothèse selon laquelle les disciplines de spécialisation de la France publieraient peu d'articles. Comme la base de données internationale du WOS privilégie les articles publiés dans des revues, la production française

y sera mal représentée. Il est également possible de formuler cette hypothèse à partir des disciplines de non spécialisation de la France : les articles de revues y constitueraient le type le plus courant. De ce fait, la France est sous-spécialisée dans ces disciplines, elle est donc sous-représentée dans le WOS.

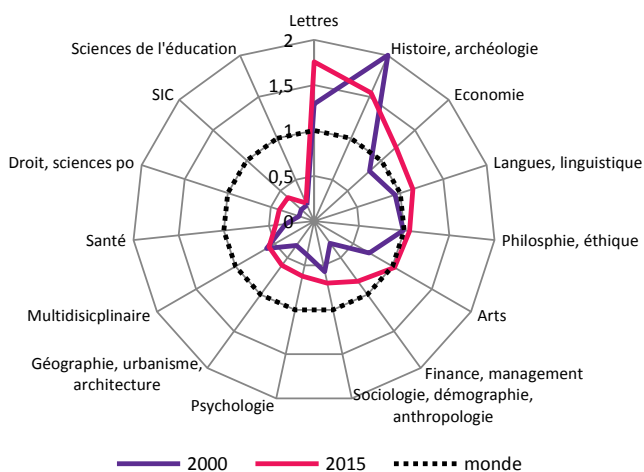
Nous avons regroupé, pour une lecture simplifiée, les 81 disciplines de recherche SHS du WOS en 15 grands disciplines qui se répartissent en 6 disciplines SH et 8 disciplines SS, plus un domaine « multidisciplinaire » (tableau 2 et annexe 1).

Tableau 2 : Les 15 disciplines SHS issues du regroupement des disciplines de recherche du WOS

Sciences sociales	Sciences humaines
Santé	Arts
SIC (sciences de l'information et de la communication)	Histoire, archéologie
Sociologie, démographie, anthropologie	Langues, linguistique
Droit, sciences politiques	Lettres
Economie	Philosophie, éthique
Education	Psychologie
Finance, management	-
Géographie, urbanisme, architecture	-
Multidisciplinaire	

Le graphique 3 fournit les indices de spécialisation pour les deux dates 2000 et 2015. Au sein des sciences humaines, la France a un indice de spécialisation qui passe de 2,6 à plus de 2,1 en histoire/archéologie. Par ailleurs, la France renforce sa spécialisation dans quasiment toutes les disciplines, notamment en lettres (1,3 à 1,8), langue/linguistique (0,9 à 1,1) et dans le domaine des arts (0,7 à 1,0). En sciences sociales, la France est devenue spécialisée en économie au cours de la période, avec un indice de 1,2 en 2015. L'indice de spécialisation de la France a fortement progressé en finance, management (de 0,30 à 0,82) tout en restant au-dessous de 1.

Graphique 3 : spécialisation de la France dans les disciplines des SHS ; 2000-15

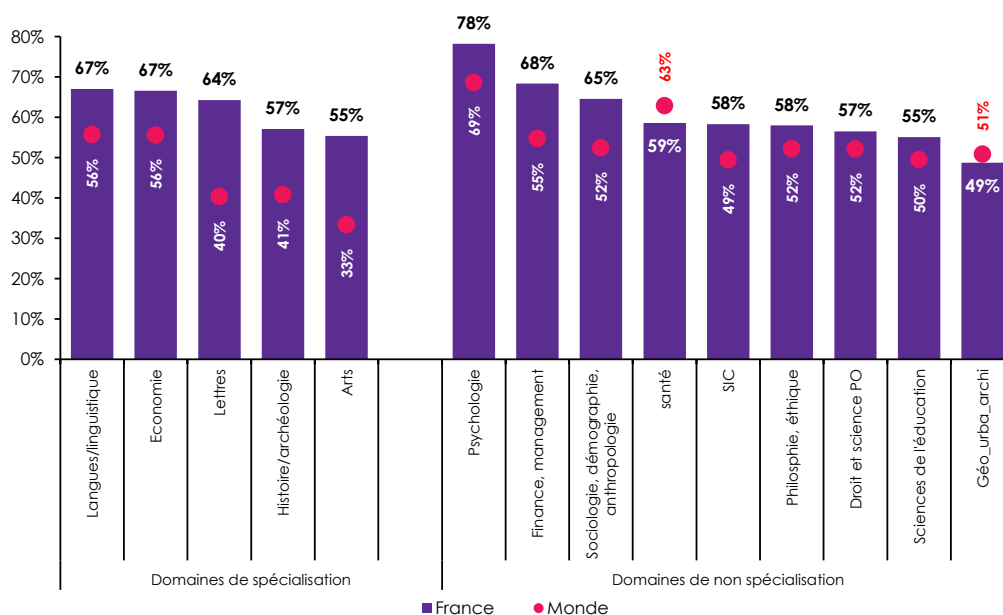


Source : Base OST, Web of Science, calculs OST

Pour résumer, les disciplines de spécialisation (indice de spécialisation supérieur à 1) de la France en SHS en 2015 en SH l'histoire/archéologie, lettres, langues/linguistique et arts et en SS, l'économie. Dans les 10 disciplines restant la France n'est pas spécialisée.

Une fois que nous avons déterminé les disciplines de spécialisation et de non spécialisation de la France en SHS, il nous est possible maintenant de tester la première hypothèse portant sur la nature des publications. Le graphique 4 montre la part des articles dans les disciplines de spécialisation et de non spécialisation (SHS) de la France en 2015 dans le WOS.

Graphique 4 : Part des articles dans les disciplines de spécialisation et de non spécialisation (SHS) de la France, 2015 (WOS).



Source : Base OST, Web of Science, calculs OST

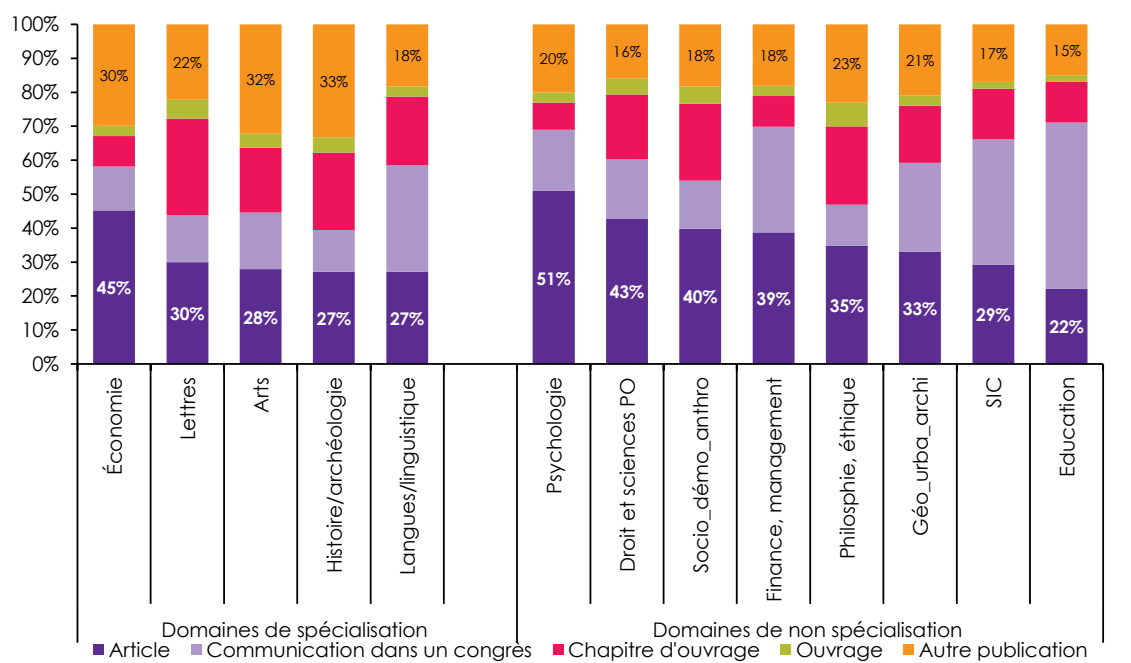
Nous constatons que, au niveau mondial, la part des articles dans les cinq disciplines de spécialisation de la France varie selon les disciplines. Les articles sont majoritaires en langues/linguistique et en économie (soit une part de 56%). À l'inverse, la part est relativement faible pour les arts (33%), les lettres (40%) et l'histoire/archéologie (41%). Dans le cas de la France, la part des articles est beaucoup plus élevée pour l'ensemble des disciplines où elle est spécialisée. Sa part est de 67% en langues/linguistique et en économie, 64% en lettres, suivies de l'Histoire/archéologie et Arts respectivement avec des parts de 57% et 55 %.

En ce qui concerne les disciplines où la France n'est pas spécialisée, on observe une tendance similaire. C'est-à-dire, globalement, que la part des articles dans les publications françaises dépasse celle du monde, hormis pour « santé » et « géographie, urbanisme et architecture ».

Afin d'approfondir l'analyse de la distribution des types de publications par discipline, nous avons examiné l'archive ouverte française HAL³. L'objectif est d'étudier dans un premier temps, la place qu'occupent des articles par rapport aux bases de données internationales. Dans un second temps, nous souhaitons vérifier si les disciplines de spécialisation et de non spécialisation de la France se comportent de la même façon.

Le graphique 5 montre la répartition des publications selon le type de dépôt dans HAL-SHS pour les disciplines de spécialisation et de non spécialisation de la France en 2015. Nous constatons la même tendance que dans le WOS mais avec une part moins importante de l'article dans HAL-SHS. Les disciplines de non spécialisation de la France publient relativement plus d'articles que les disciplines de spécialisation. Dans les disciplines de spécialisation de la France la part des articles n'excède pas 30% (en dehors l'économie). La psychologie est la discipline ayant la part des articles la plus élevée dans HAL-SHS (51%), suivie de l'économie (45%). En sciences de l'éducation, la part des communications dans des congrès est très importante (elle avoisine le 30%).

Graphique 5 : Part des articles dans les disciplines de spécialisation et de non spécialisation (SHS) de la France, 2015 (HAL-SHS).



Source : HAL, calculs OST

L'analyse des données WOS et HAL permet de vérifier notre première hypothèse : Les disciplines où la France est spécialisée ne sont pas dominées par les articles. Paradoxalement, la France publie plus d'articles que la moyenne mondiale dans toutes ces disciplines; mais aussi dans les disciplines SHS où elle n'est spécialisée. Ces résultats pourraient être expliqués de deux façons. Premièrement, les autres publications françaises (ouvrages, chapitres d'ouvrages...) seraient majoritairement publiées en langue française sans résumé en anglais ou sans traduction pour les ouvrages, donc non recensées dans le WOS. Deuxièmement la production scientifique de la France en SHS serait relativement faible

³ HAL est une archive ouverte française lancée en 2001 sur la base du volontariat, les chercheurs déposent leurs productions. <https://hal.archives-ouvertes.fr/>

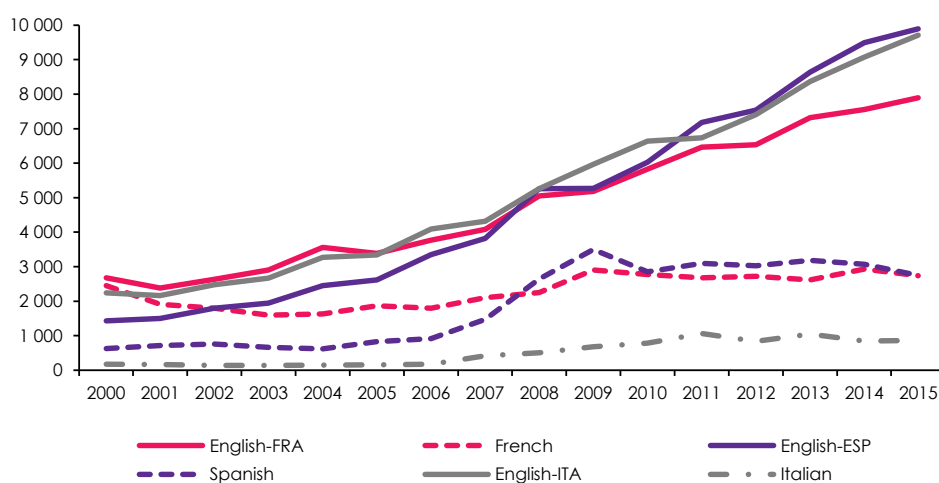
comparée à celle d'autres pays, et la production scientifique française ne serait effectivement pas spécialisée dans un certain nombre de disciplines des SHS.

4. La part élevée de publication en langue nationale est-elle une spécificité française ?

Dans cette section, nous explorons, dans un premier temps, l'évolution des volumes de publication en langues nationales et en anglais des 5 pays de notre référentiel, à savoir la France, l'Espagne, l'Italie, l'Allemagne et les Pays-Bas. Dans un second temps, nous analysons l'évolution de la part de l'anglais dans les publications en SHS de chacun de ces pays.

Il s'agit de vérifier si la sous-représentation des SHS de la France peut s'expliquer par le fait que les français publient plus en langue nationale, alors que les bases bibliométriques internationales recensent plutôt les publications en anglais.

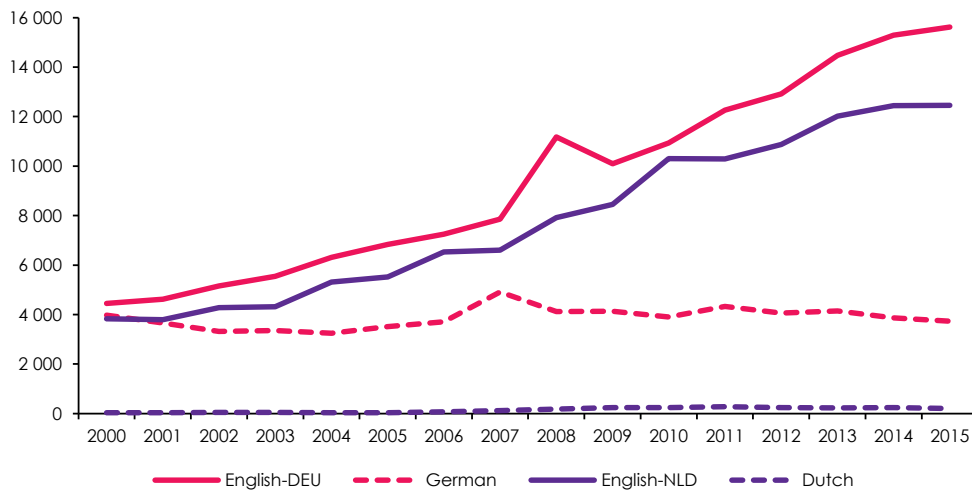
Graphique 6 : nombre de publications en SHS selon la langue, France, Espagne et Italie



Le graphique 6 montre l'évolution du volume de publications (tous types de documents confondus) en anglais et en langue nationale. Le volume de publications en anglais croît pour les trois pays ; France, Espagne et Italie. Mais la tendance de la progression n'est pas la même. En 2000, le volume de publications en anglais de la France excède celui des deux autres pays. L'Italie comme l'Espagne ont enregistré une progression plus forte que celle de la France et la dépassent, respectivement, à partir de 2005 (pour l'Italie) et 2008 (pour l'Espagne). La progression de l'Espagne est tellement forte qu'elle surpasse également l'Italie à partir de 2010.

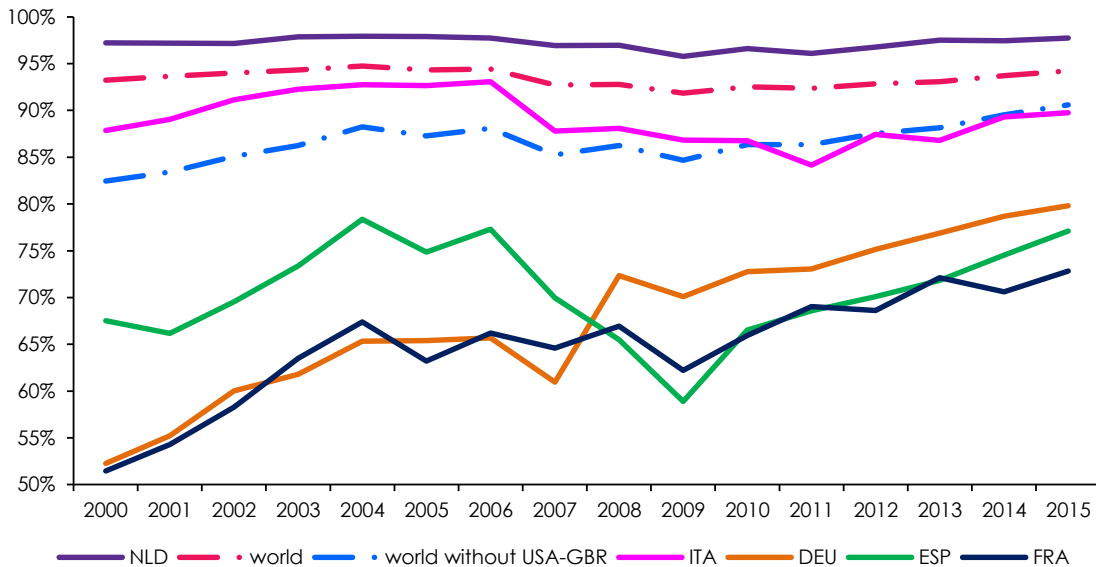
Quant aux langues nationales, le graphique 6 montre une tendance intéressante. L'Espagne croît fortement son volume de publications dans sa langue nationale notamment entre 2006 et 2009, période pendant laquelle elle dépasse le volume des publications françaises écrites en français. Cette tendance se tasse à partir de 2010 mais le volume de publications de l'Espagne en langue nationale reste globalement supérieur à celui de la France. À l'opposé, l'Italie publie peu dans sa langue nationale, comparée à la France et à l'Espagne.

Graphique 7 : nombre de publications en SHS selon la langue, Allemagne et Pays-Bas en SHS



Tout comme le graphique 6, le graphique 7 montre une progression du nombre de publication en anglais, à la fois pour l'Allemagne et pour les Pays-Bas. Le volume des publications allemandes écrites en anglais excède (avec un pic en 2008) celui des publications hollandaises sur la période. À l'opposé, les publications en langue nationale poursuivent leur baisse pour l'Allemagne, et sont marginales pour les Pays-Bas.

Graphique 8 : part de l'anglais dans les publications nationales en SHS des pays du référentiel



Le graphique 8 montre l'évolution de la part de l'anglais sur l'ensemble des publications en SHS des pays du référentiel. La part de l'anglais dans les publications néerlandaises est élevée ; elle s'approche des 100% sur la période. De même, l'Italie a une part de publications en anglais qui avoisine les 90%. Une tendance exceptionnelle est également observée pour le cas de l'Allemagne, qui, en 2000, comptait un peu plus de 50% de publication en anglais ; la part a atteint les 80% en 2015 avec une tendance à la hausse soutenue, notamment depuis 2009. Le cas de l'Espagne est particulier ; sa part des publications écrites en anglais

en SHS est passée de 65 à 75 % entre 2000 et 2006. Puis la part chute jusqu'en 2009 où elle passe en dessous des 60%. Cette période correspond à la forte croissance du volume de publications écrites en espagnol (voir graphique 6). Après 2009, la part de l'anglais dans les publications espagnoles repart à la hausse et excède désormais celle de la France. La France quant à elle, augmente également sa part de publications en anglais. Cependant, étant donné sa tendance, elle représente le pays ayant la part la plus faible de publications écrites en anglais au sein de notre référentiel (72.8%). L'écart entre la France et les quatre autres pays s'élargit à partir de 2014.

Au regard des deux précédents graphiques, notre hypothèse construite autour de la langue de publication se vérifie. Hormis le cas de l'Espagne, la France représente le pays ayant la part de publication en anglais la plus faible. Et réciproquement, une part de publication en langue nationale la plus forte. Cependant, la forte progression des publications de l'Espagne en langue nationale (en plus de celles en anglais) laisse des questions en suspens. Notamment, pourquoi les espagnols publient plus que les français dans leur langue nationale ? Ou, formulé autrement, pourquoi les publications en espagnol sont plus visibles dans le WOS que les publications en Français ?

5. Conclusions et discussion

L'objectif de ce travail était d'analyser la dynamique des SHS françaises dans le WOS en les comparant avec les principaux pays européens non anglophones. Nous sommes partis de l'hypothèse selon laquelle la France serait doublement pénalisée dans les bases internationales qui ignorent à la fois les spécificités des pratiques des SHS dans lesquelles les articles ne sont pas majoritaires, et la prédominance de la langue nationale dans ce domaine.

L'évolution du nombre de publications en SHS dans le WOS est contrastée, entre l'Espagne qui publie beaucoup plus depuis 2000 et la France qui peine à suivre la dynamique des autres pays. Autrement dit, au sein des pays du référentiel confrontés théoriquement aux mêmes biais de la base de données, la France enregistre une moindre progression.

Deux hypothèses sont susceptibles d'expliquer ce phénomène. La première consiste à dire que la France, en SHS, se spécialise dans des disciplines où la part des articles est relativement faible. Et donc ses publications majoritaires, étant donné la spécialisation, sont des produits autres que les articles qui ne seraient pas recensés dans le WOS. Inversement, si nous considérons plutôt les disciplines de non spécialisation de la France, la proportion des articles dans ces disciplines serait très élevée. Et comme la France n'est pas spécialisée, elle a une plus faible part de publications dans ces disciplines, raison pour laquelle sa présence dans le WOS est faible par rapport aux autres pays.

Les résultats issus des tests de la première hypothèse indiquent que dans les quatre disciplines de spécialisation de la France, la part des articles pour le monde n'est pas dominante (50% en histoire/archéologie, 46% en lettres, 40% en arts) sauf en langues/linguistique (64%) en 2015. Paradoxalement, la France publie plus d'articles que la moyenne mondiale dans les SHS où elle est spécialisée. La part des articles pour la France dans ces disciplines se situe entre 68 et 74% en 2015. Par conséquent, notre première hypothèse ne se vérifie que partiellement. La France publie plus d'articles que la moyenne mondiale dans pratiquement toutes les disciplines des SHS, qu'elle y soit spécialisée ou non. Ce résultat laisse émerger beaucoup de questions : est-ce que seuls les articles français passent la barrière à l'entrée du WOS ? Ou encore, est-ce qu'il est plus facile de publier un article en anglais qu'un ouvrage

pour les chercheurs français sachant que le temps d'élaboration, de maturation et de rédaction d'un ouvrage de recherche en SHS est plus long que pour un article ?

C'est pour cette raison que nous avons émis une deuxième hypothèse qui pourrait constituer une partie de l'explication. Selon cette seconde hypothèse, les chercheurs français en SHS publient peu en anglais, moins que les autres pays non anglophones. C'est pour cela qu'une partie non négligeable de leur production n'apparaît pas dans les bases de données internationales.

Le volume des publications françaises en anglais progresse plus que celui des publications en français certes. Néanmoins, les autres pays enregistrent une augmentation beaucoup plus forte des publications en anglais. L'Espagne part de loin puis dépasse la France en termes de volume même pour les publications écrites en espagnol. La part de l'anglais pour les Pays-Bas avoisine les 100%, l'Italie 90% et on constate également une progression pour l'Allemagne des publications écrites en anglais. Par ailleurs, la part de l'anglais de l'Espagne n'est pas très loin de celle de la France (bien que l'écart s'accroisse à partir de 2014). L'Espagne a accru ses publications en sa langue nationale dans le WOS contrairement à la France. Les résultats de la seconde hypothèse nous semblent répondre, mais en partie, à nos questionnements. La France étant le dernier pays en termes de part de publications en anglais, peine à affirmer sa présence dans le WOS.

Ainsi, la langue peut constituer une partie de l'explication de la faible visibilité des SHS françaises dans le WOS, mais des questions restent en suspens, notamment pour le cas de l'Espagne. Pourquoi les espagnols ont eu une aussi forte dynamique des publications écrites en espagnol dans le WOS contrairement aux français? S'agit-il d'un artéfact de la base de données qui favorise la langue espagnole (utilisée largement en Amérique latine et en Amérique du Sud) ou adopte une stratégie de rattrapage de l'Espagne ? Ou bien, s'agit-il plutôt d'une vraie progression de l'Espagne, corolaire d'une stratégie de recherche efficiente rendant les chercheurs plus productifs ? Le rapport de l'OST (2018) a montré que l'Espagne est le premier pays ayant bénéficié de l'intégration (ou de la création) de nouvelles revues dans le WOS en sciences sociales. Il nous semble intéressant d'analyser en profondeur cette question afin de savoir ce qui a permis à l'Espagne un tel sursaut.

En résumé, il est partiellement possible de répondre à notre question quant à la faible présence des SHS françaises dans le WOS. Cela ne semble pas uniquement dû à une mauvaise couverture disciplinaire de la base de données dans la mesure où il y a des écarts importants entre pays non anglophones. Par contre, le manque d'adaptation aux standards internationaux de publication, notamment publier des articles en anglais ou au moins avec un résumé en anglais, peut constituer une partie de l'explication.

Pour avoir des explications plus pertinentes, il sera intéressant d'analyser en profondeur les politiques publiques liées à l'édition scientifique, à l'open science et leur impact sur la visibilité des chercheurs français. Une analyse des systèmes de recherche des pays de notre référentiel peut également fournir des éléments d'explication.

Références bibliographiques

- Butler, L. (1998). Personal communication of unpublished data: Cf. Hicks (2004), 479.
- Chen, K., Tang, M., Wang, C., & Hsiang, J. (2015). Exploring alternative metrics of scholarly performance in the social sciences and humanities in Taiwan. *Scientometrics*, 102(1), 97–112. <http://doi.org/10.1007/s11192-014-1420-6>.
- Glänzel W., Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing and Management*, 35, 31-44.
- Gruzd, A., & Goertzen, M. (2013). Wired Academia: Why Social Science Scholars Are Using Social Media (pp. 3332–3341). Presented at the 46th Hawaii International Conference on System Sciences, Maui, Hawaii: IEEE. <http://doi.org/10.1109/HICSS.2013.614>.
- Harzing, A. W., & Alakangas, S. (2017a). Microsoft Academic: Is the phoenix getting wings? *Scientometrics*, 110(1), 371–383.
- Harzing, A. W., & Alakangas, S. (2017b). Microsoft Academic is 1 year old: The Phoenix is ready to leave the nest. *Scientometrics*, 112(3), 1887–1894.
- Heinzkill, R. (2003). References in scholarly English and American literary journals thirty years later: A citation study. *COLLEGE & RESEARCH LIBRARIES*. Volume: 68. Issue: 2. Pages: 141-153.
- Hicks, D. (2004). The four literatures of social science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research: The use of publication and patent statistics in studies of S&T systems* (pp. 476–496). Dordrecht: Kluwer Academic Publishers.
- Hug, S. E., & Brändle, M. P. (2017). The coverage of Microsoft Academic: Analyzing the publication output of a university. *Scientometrics*. <https://doi.org/10.1007/s11192-017-2535-3>.
- Katina Strauch (Westport, Conn.: Libraries Unlimited, 2003), 144.
- Knievel, J.E., & Kelley, C. (2005). Citation analysis for collection development: A comparative study of eight humanities fields. *The Library Quarterly* 75(2): 142-168.
- Larivière et al (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*. Volume: 57. Issue: 8. Pages: 997-1004.
- Leydesdorff, L. (2003). Can networks of journal–journal citation be used as indicators of change in the social sciences? *Journal of Documentation*, 59 (1), 84–104.
- Leydesdorff, L. (2003). Can networks of journal–journal citation be used as indicators of change in the social sciences? *Journal of Documentation*, 59 (1), 84–104.
- McDonald, J. (2003). "The Book is Dead," in *Charleston Conference Proceedings, 2001*, ed.
- Mohammadi, E. (2014). *Identifying the Invisible Impact of Scholarly Publications: A Multi-disciplinary Analysis Using Altmetrics* (Ph.D. thesis). University of Wolverhampton, Wolverhampton, UK.

Mohammadi, E., & Thelwall, M. (2013). Assessing non-standard article impact using F1000 labels. *Scientometrics*, 97(2), 383–395. <http://doi.org/10.1007/s11192-013-0993-9>.

Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology*, 65(8), 1627–1638. <http://doi.org/10.1002/asi.23071>.

Mohammadi, E., Thelwall, M., & Kousha, K. (2015b). Can Mendeley Bookmarks Reflect Readership? A Survey of User Motivations. *Journal of the Association for Information Science and Technology*, <http://doi.org/10.1002/asi.23477>.

Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (2015a). Who reads research articles? An altmetrics analysis of Mendeley user categories. *Journal of the Association for Information Science and Technology*. <http://doi.org/10.1002/asi.23286>.

Mongeon, P., & Paul-Hus, A. (2016), The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213-228.

OST (2018), *La position scientifique de la France dans le monde, 2000-2015*, Hcéres, Paris.

Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, 68(9), 2037–2062.

Waltman L. (2016), A review of the literature on citation impact indicators. *Journal of infometrics*. 10, 365–391 (2016). doi: 10.1016/j.joi.2016.02.007.