Non-English Journals and Papers in Physics: Bias in Citations?

Liming Liang, Ronald Rousseau and Zhen Zhong

This article has been published in:

Proceedings of the ISSI 2011 Conference

Edited by Ed Noyons, Patrick Ngulube and Jacqueline Leta

Published by: ISSI, Leiden University and University of Zululand

2011

pp. 463-473

Non-English Journals and Papers in Physics: Bias in Citations?

Liming Liang[1], Ronald Rousseau[2] and Zhen Zhong[3]

[1] *liangliming1949@sina.com*

[1]Henan Normal University, Institute for Science, Technology and Society, Xinxiang, 453007 (China)

Dalian University of Technology, WISE Lab, Dalian, 116024 (China)

[2] ronald.rousseau@khbo.be

KHBO (Association K.U.Leuven), Industrial Sciences and Technology, Zeedijk 101, 8400 Oostende (Belgium)

University of Antwerp (UA), IBW, Venusstraat 35, 2000 Antwerp, Belgium

[3] *clock_town@126.com*

Dalian University of Technology, WISE Lab, Dalian, 116024 (China)

Abstract

This study aims at investigating, at the journal as well as the article level, if there is a huge difference in citations between English-language publications and non-English publications. The investigation uses the Web of Knowledge as data source and focuses on the field of physics. Using a precise definition of a "non-English journal" we filter out nine non-English journals, scattered in six physics subfields. Average received citations per paper (CPP) of the non-English journal(s) are compared with the CPP of pure English journals, and this in the same subfield. We clearly observe that non-English journals are inferior - in number of citations received - to pure English journals and this in all physics subfields studied. Further, twelve physics journals were chosen as sample journals to compare the CPP of non-English papers with that of pure English papers in the same journal. The result of this comparison is that for the majority of these journals and for most of the publication years the CPP of non-English papers is lower than that of the pure English papers. Finally, analyzing linguistic characteristics of the literature citing nine non-English physics journals confirm the own-language preference in physics.

Introduction

Already in 1978 Garfield (1978) wrote, see also (Garfield, 1979) that for articles with at least one French author the average number of received citations over the period he studied was 7.45 for those written in English, while for those written in French it was only 2.9. There was however nothing particular about French. Garfield noted that similar differences also hold for German vs. English (German scientists), Russian vs. English (Russian scientists) and Spanish vs. English (Spanish scientists). For the Spanish case he mentions that an English-language article receives on average 2.4 more citations than a Spanish language one (of course, all his data come from the Science Citation Index). Garfield's opinion is that French scientists publish their best work in English (and outside France). As the term bias refers to prejudice based on unreasoned judgment, the low citation values for French language articles are not the result of bias, according to Garfield's opinion. Yet, it is clear that if the same finding is independently published once in English and once in another language, the English article will garner the most citations. The introduction of the notion of co-citation by Small (1973),

cited 530 times (2011-1), and Marshakova (1973), cited 87 times (2011-1), may be a case in point in our field.

Yet, real bias may occur when comparing countries based on SCI data. This has clearly been shown by van Leeuwen et al. (2000, 2001). Depending on the percentage of English-language articles covered by the SCI (nowadays WoS) a country may look better or worse, without any real difference in publication quality (even as measured through citations). They recently reiterated this point (van Raan et al., 2011) stating that the effect of non-English journals is particularly evident in application-oriented fields and in the context of university rankings. Yitzhaki (1988) observed that English-speaking scientists make very little use of foreign language material and studied the reverse problem, namely the use of the own language by non-English speaking scholars. In doing so, he introduced the language self-citation index, and initiated a search for measures of "own-group preference" (Bookstein & Yitzhaki, 1999; Egghe & Rousseau, 2004).

In order to counter the significant negative language bias in country comparisons, as they call it, Grupp et al. (2001) propose a two-dimensional representation. This representation uses two indicators: scientific regard and international alignment. These indicators are variations of the better known MECR (mean expected citation rate) and NMCR (normalized mean citation rate). Lin & Zhang (2007) observed that there exists a group of Chinese scientists in nanoscience and technology who publish in Chinese and have few connections with the rest of the world (read: the English-language world). Surprisingly, this group seems to increase.

We note that the problem regarding languages has also been observed in business-related studies too. Depending on the mother tongue and cultural accommodation respondents may answer the same question differently, so that results of surveys may measure cultural or linguistic differences, and not the differences that were intended (Harzing et al., 2009).

In this contribution we investigate if there still is such a huge difference in citations between English-language publications and other ones, and this in the field of physics.

Data and Methods

The data about physics journals, publications and citations for this study were obtained from Thomson Reuter's Web of Knowledge. The time span considered is 1978-2007. We took the following steps.

Step 1: Construction of the sample set of physics journals.

We collected the journals which occurred in at least one of the following eight JCR categories during the period 2000-2007.

1) Physics, Applied: with 103 journals.
2) Physics, Atomic, Molecular & Chemical: with 40 journals.
3) Physics, Condensed Matter: with 70 journals.
4) Physics, Fluids & Plasmas; with 25 journals.
5) Physics, Mathematical: with 44 journals.
6) Physics, Multidisciplinary: with 87 journals.
7) Physics, Nuclear: with 25 journals.
8) Physics, Particles & Fields: with 25 journals.

Although the sum of the number of journals is 419 there are actually only 352 different journals in this set as many journals are categorized in more than one JCR category. The *International Journal of Modern Physics B*, for example, belongs to three categories. This set of 352 journals will be referred to as the physics journal set.

Step 2: Definition of “non-English” journals.

A journal in the physics journal set is called a non-English journal over a certain period if in each year of this period the journal published at least one non-English paper and during the whole period the percentage of non-English papers is at least 50. Here, only ‘normal’ articles, review articles and proceeding papers are considered as papers.

In the “Journal Information” section offered by the JCR many journals are labelled as “Multi-language”. However, when retrieving papers from all journals that might be considered non-English, including the “Multi-language” ones, we found only nine journals that satisfied our definition of a non-English journal. These nine journals are shown in Table 1, numbers 1 to 9. This set of nine journals consists of one German journal and three Chinese, three Russian, and two French journals. These nine journals are used in our study on the citation difference between non-English and English-language journals.

For the study on the citation difference between non-English and English-language papers we selected some more journals that have published at least 200 non-English papers. In this way Table 1 is augmented with six more physics journals (numbers 10 to 15).

Step 3: Comparison of journals 1 to 9 with pure English-language journals.

The purpose of this comparison is to find the position of non-English journals in their own disciplinary citation environment. For each of the nine non-English journals the number of citations received over the period between the publication year and 2007 were collected and this for each publication year separately. The oldest publication year depends on the first year for which the journal satisfied the definition of a non-English journal. The average number of citations received per article (denoted as CPP) was determined for each publication year. All pure English journals (journals in which all published articles are in English), in each of the studied subfields were considered as one mega journal (one per subfield) and their CPPs, for each publication year, used for comparisons.

In step 3 we performed an analysis on journal level. In the next step we will perform an analysis on article level.

Step 4: Selection of articles in journals 4 to 15 for a comparison on article level.

Again the indicator CPP was used. The CPP time sequence of non-English papers was compared with that of English papers published *in the same journal*. Journals 1 to 3 were not used as the number of English papers they published was considered to be too low for valid comparisons. As explained above, journals 10 to 15 were selected, because each of them published enough non-English papers to form a comparable non-English paper set.

Table 1. Nine non-English physics journals and six sample journals with more than 200 non-English papers each

No.	JCR abbrev. title	Publication period with non-English papers	Subfield	Total papers	English papers	Non-English papers	Share of non-English papers
1	*USP FIZ NAUK*	1978-2001	Multidisciplinary	1621	2	Russian 1619	99.9%
2	*ACTA PHYS SIN-CH ED*	1999-2007	Multidisciplinary	6642	80	Chinese 6562 Spanish 1	98.8%
3	*IZV AKAD NAUK FIZ*	1993-2002	Multidisciplinary	3663	50	Russian 3613	98.6%
4	*CHINESE J CHEM PHYS*	2002-2005	Atomic, Molecular & Chemical	532	34	Chinese 498	93.6%
5	*PTB-MITT*	1978-2003	Multidisciplinary	815	109	German 705 Georgian 1	86.6%
6	*METALLOFIZ NOV TEKH*	1995-2007	Condensed Matter	1871	286	Russian 1218 Ukrainian 359 Czech 7 German 1	84.7%
7	*HIGH ENERG PHYS NUC*	1998-2007	1) Nuclear 2) Particles & Fields	2483	436	Chinese 2047	82.4%
8	*VIDE*	1978-2003	Applied	1332	391	French 941	70.6%
9	*ANN PHYS-PARIS*	1978-2007	Multidisciplinary	1485	469	French 1016	68.4%
10	*REV MEX FIS*	1992-2007	Multidisciplinary	2313	1644	Spanish 660 French 9	28.9%
11	*ANN PHYS-BERLIN*	1978-1994	Multidisciplinary	874	653	German 222	25.3%
12	*J PHYS IV*	1991-2006	Multidisciplinary	10767	8933	French 1834	17.0%
13	*HELV PHYS ACTA*	1978-1999	Multidisciplinary	1754	1522	German 120 French 112	13.2%
14	*Z NATURFORSCH A*	1978-1997	Multidisciplinary	3828	3532	German 295 French 1	7.7%
15	*CAN J PHYS*	1978-2007	Multidisciplinary	5099	4884	French 215	4.2%

Findings

Non-English journals receive less citations than English-language journals, and this in the same subfield

The indicator CPP was taken as the criterion to assess the citation position of non-English journals. Journals 1 to 9 are used for these comparisons, which are made on the JCR subfield level. These nine journals belong to six subfields: journals 1,2,3,5 and 9 all belong to the same subfield, namely “Physics, Multidisciplinary”. These five journals taken together were compared with the collective (mega journal) of 52 pure English journals in the same subfield. Journal 7 belongs to two subfields, “Physics, Nuclear” and “Physics, Particles and Fields”, and hence compared with the pure English journals in both subfields. Journals 4, 6 and 8 belong to “Physics, Atomic, Molecular & Chemical”, “Physics, Condensed Matter”, “Physics,

Applied", respectively. The three journals were compared with the pure English journals in these three subfields. Figure 1 presents the results of the comparisons based on the CPP of the non-English journals with those of the English journals, and this in the six physics subfields mentioned above.

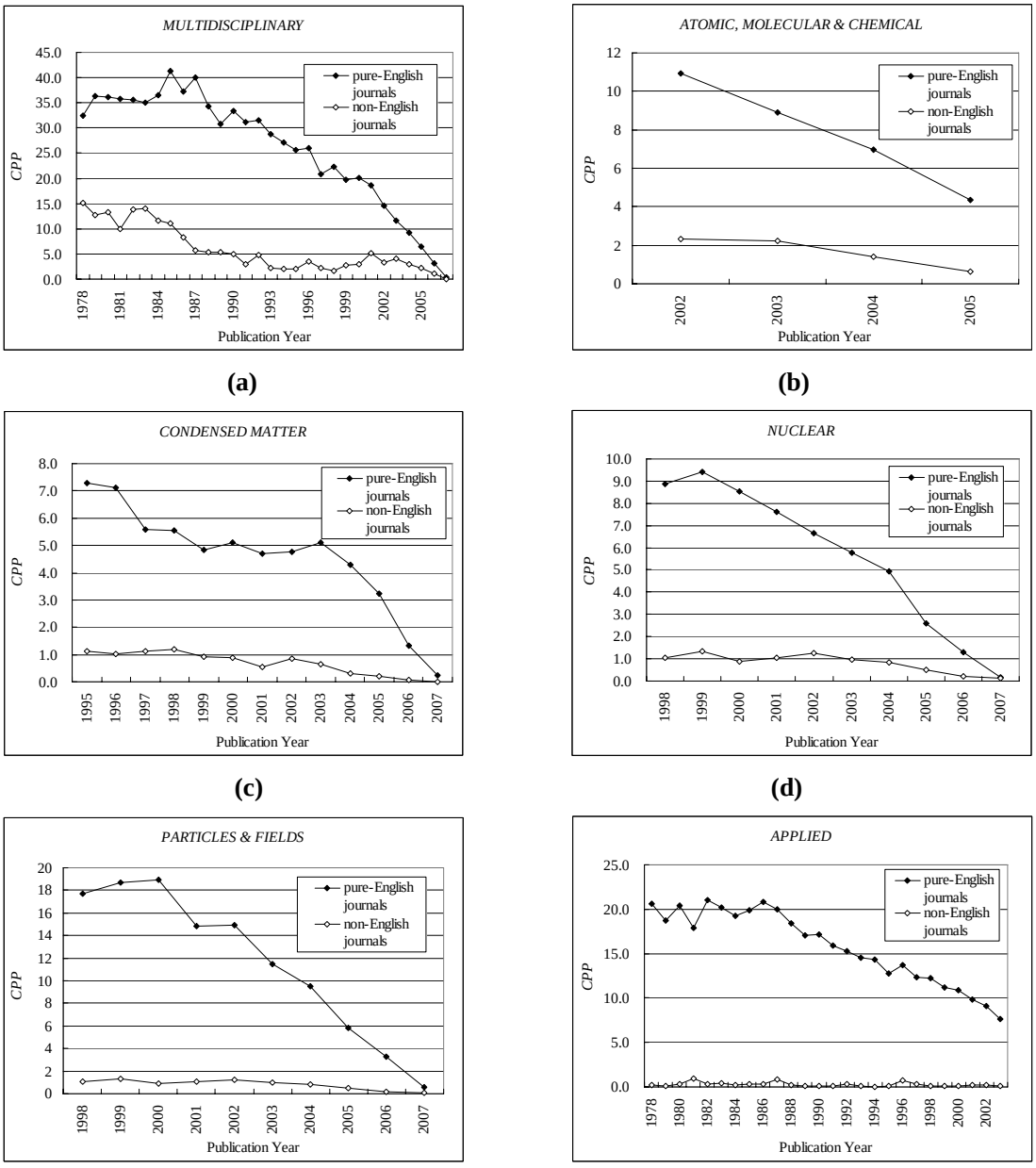


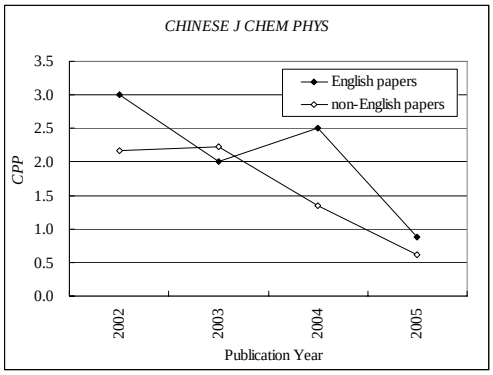
Figure 1 CPP comparison between non-English journals and pure English journals

The six pairs of curves shown in Figure 1 clearly show that no matter what subfield, and no matter which publication year, or even how many journals are compared, all the CPP curves of the pure English journals are situated highly above the CPP curves of the non-English journal(s). There is no doubt that the CPP of the non-English journal(s) is much lower than the average level of the pure English journals' CPP. This phenomenon reveals that

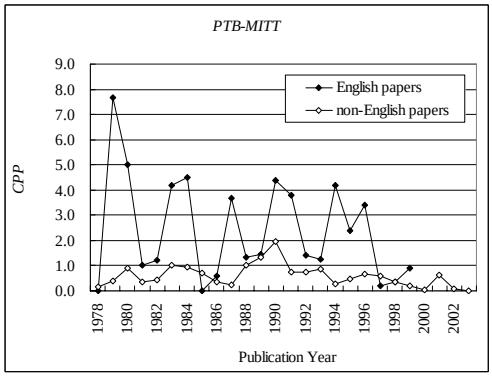
non-English journals are inferior to English-languages journals in the average number of citations received.

Non-English papers receive generally less citations than English-language ones, and this in the same journal

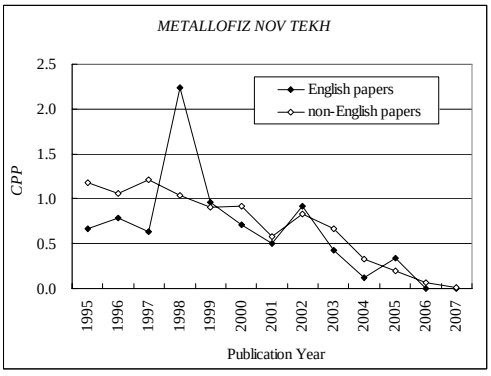
Again CPP is used as the criterion for assessing the citation position of non-English papers. Comparisons of non-English papers with English papers are made for publications in the same journal. The CPP of non-English papers and that of English papers were matched by publication year. Journal 4 to 15 from Table 1 were selected as sample journals. These 12 journals are scattered over six physics subfields. Figure 2 consists of 12 panels. Each panel represents one journal.



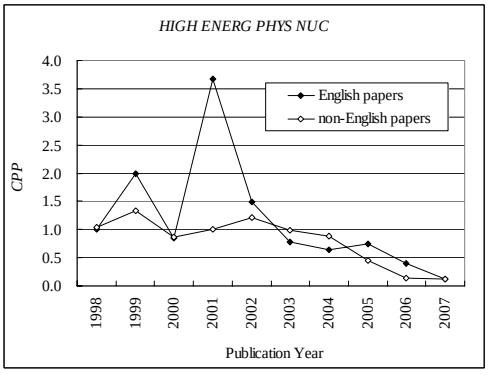
(a)



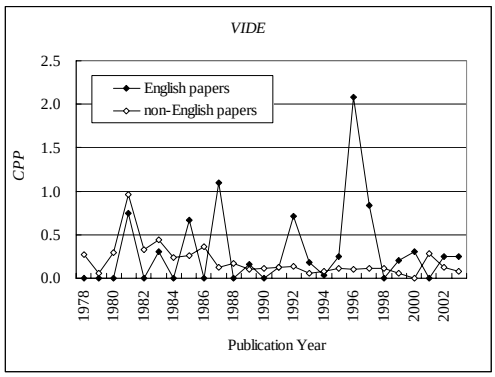
(b)



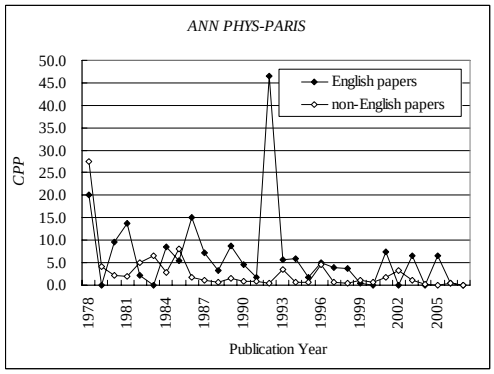
(c)



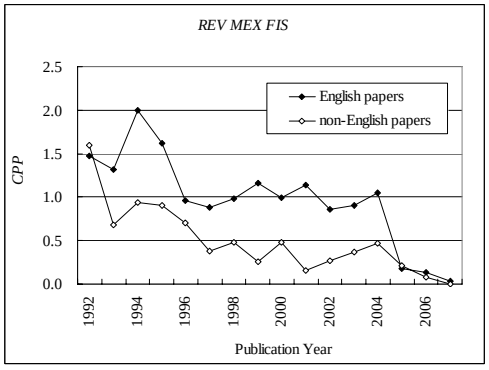
(d)


VIDE
English papers
non-English papers
CPP
Publication Year

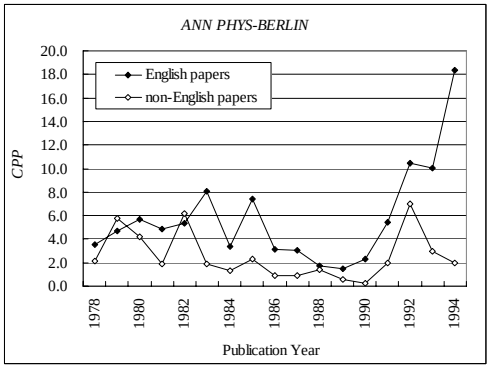
(e)


ANN PHYS-PARIS
English papers
non-English papers
CPP
Publication Year

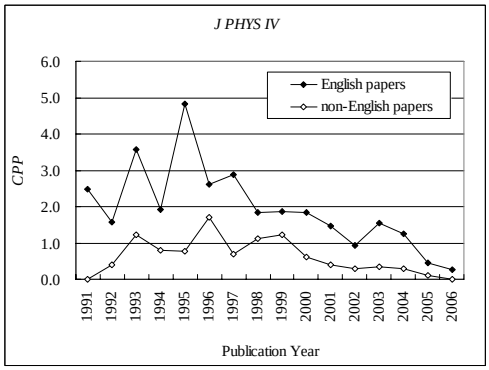
(f)


REV MEX FIS
English papers
non-English papers
CPP
Publication Year

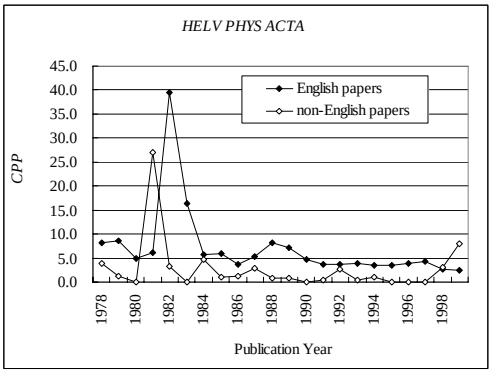
(g)


ANN PHYS-BERLIN
English papers
non-English papers
CPP
Publication Year

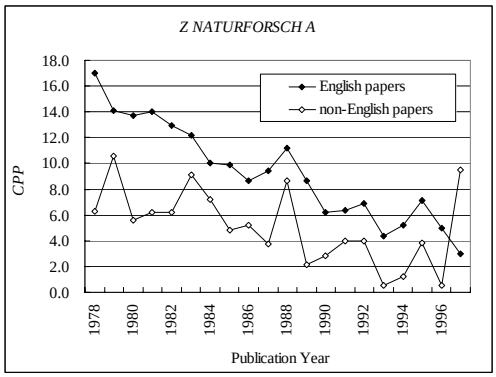
(h)


J PHYS IV
English papers
non-English papers
CPP
Publication Year

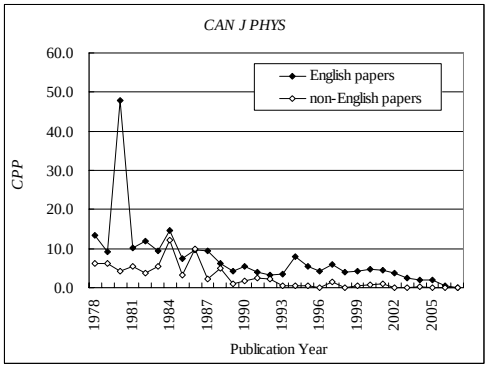
(i)


HELV PHYS ACTA
English papers
non-English papers
CPP
Publication Year

(j)



(k)



(l)

Figure 2 CPP comparison between non-English papers and English papers published in the same journal

We refer to the CPP curve of English papers as Curve-E and to the CPP curve of non-English papers as Curve-N. According to the relative position of Curve-E and Curve-N in the different panels the twelve pairs of curves, representing results for twelve journals, can be classified into the following four types. In the bracket after the journal's name we provide the following information: the letter symbol of the panel in Figure 2, the main non-English language, and the percentage of years in which Curve-N is located above Curve-E.

Type Ⅰ

Curve-N is situated above Curve-E in at most 5% of the years studied. Type Ⅰ consists of 3 journals：*J PHYS IV* (i, French, 0%), *CAN J PHYS* (l, French, 3%) and *Z NATURFORSCH A* (k, German, 5%).

Type Ⅱ

Curve-N is situated above Curve-E in 10 to 15 percentage of the years studied. Type Ⅱ consists of 5 journals: *ANN PHYS-PARIS* (f, French, 10%), *PTB-MITT* (b, 11.5% , German), *ANN PHYS-BERLIN* (h, German, 11.8%), *REV MEX FIS* (g, Spanish, 12.5%), and *HELV PHYS ACTA* (j, German and French, 13.6%).

Type Ⅲ

Curve-N is situated above Curve-E in 20 to 25 percentage of the years studied. Type Ⅲ consists of 2 journals, *HIGH ENERG PHYS NUC* (d, Chinese, 20%) and *CHINESE J CHEM PHYS* (a, Chinese, 25%).

Type Ⅳ

Curve-E and Curve-N intersect many times. Type Ⅳ consists of 2 journals, *VIDE* (e, French, 50%) and *METALLOFIZ NOV TEKH* (c , Russian and Ukrainian, 61.5%). For VIDE, in half of its 26 publication years Curve-N is above Curve-E; while the journal *METALLOFIZ NOV TEKH* covers 13 publication years, and in eight years Curve-N is situated above Curve-E.

Ten of the twelve physics journals belong to Types Ⅰ, Ⅱ, Ⅲ, that is to say that for most journals

and for most publication years the CPP of non-English papers is smaller than the CPP of English papers published in the same journal. Obviously, non-English physics papers receive less citations than English-language ones, and this in the same journal.
We note, in particular, the case of *J PHYS IV* (number 12 in our list). It is not a non-English journal as its French papers account for only 17%. However, the absolute number of French papers is big, reaching 1834. It is the only journal whose curve-E is totally located above its curve-N with no exceptional year (panel i). This is a very clear case that English-language papers are favoured.

Other statistics

Apart from the above detailed citation analysis we collected other statistics from Thomson Reuter's Web of Knowledge. Concretely, we investigated some characteristics of documents (no restriction on document type) citing the papers published in the nine non-English journals. Here, the journal's "publication period with non-English papers" is the cited period. The citing window begins from the first publication year and ends in 2007. Citing literatures' main language and authors' country affiliation of each journal are listed in Table 2, which reveals, on journal level, that language self-citation is the primary factor leading to an unbalanced number of citations in these journals, see e.g. (Ren & Rousseau, 2002).

Table 2. Linguistic characteristics of the literatures citing the nine non-English journals

No.	JCR abbrev. title	Journal language	Percentage of citing literature in "Journal language"	Journal self-citation rate	Percentage of citing literature with at least one author from the country of "Journal language"
1	*USP FIZ NAUK*	Russian	42.68%	4.29%	71.39%, from Russia or USSR
2	*ACTA PHYS SIN-CH ED*	Chinese	54.13%	50.11%	97.58%, from China
3	*IZV AKAD NAUK FIZ*	Russian	26.87%	22.08%	82.28%, from Russia
4	*CHINESE J CHEM PHYS*	Chinese	46.07%	42.06%	90.91%, from China
5	*PTB-MITT*	German	13.65%	7.21%	39.57%, from German
6	*METALLOFIZ NOV TEKH*	Russian	35.72%	51.94%	15.42%, from Russia
		Ukrainian	13.83%		72.74%, from Ukraine
7	*HIGH ENERG PHYS NUC*	Chinese	48.70%	54.34%	90.03%, from China
8	*VIDE*	French	5.86%	2.47%	31.48%, from France
9	*ANN PHYS-PARIS*	French	3.88%	2.46%	35.45%, from France

Concerning the *J PHYS IV* we compared citations received for the journal's 8933 English papers and 1834 French papers, respectively.

- For documents citing English papers: journal self-citation rate 4.29%; percentage of citing literatures in English is 97.99%, in French 0.75%; percentage of papers with at least one author from France is 20.47%.
- For documents citing French papers: journal self-citation rate 9.08%; percentage of citing literatures in English is 88.01%, in French 11.08%; percentage of papers with at least one author from France is 62.94%.

These data confirm the own-language preference factor.

Conclusion and Discussion

Conclusion

We defined so-called “non-English” journals and filtered nine non-English journals, scattered in six physics subfields, from 352 physics journals. The average citation per paper (CPP) of the non-English journal(s) was compared with the CPP of pure English journals in the same subfield. Based on these studies we conclude that non-English journals are inferior in number of citations to pure English journals and this in all physics subfields studied.

Twelve physics journals were chosen as sample journals to compare the CPP of non-English papers with that of the pure English papers in the same journal. The result of this comparison is that for the majority of these journals and for most of the publication years the CPP of non-English papers is lower than that of the pure English papers. Based on this result we conclude that non-English papers receive less citations than English-language ones.

Discussion

Why do non-English journals and non-English papers attract less citations? The evident reason is the language barrier. English is the primary language of science. Most scientists can read English, therefore, they may cite English journals and English papers. There are, however, not so many scientists knowing Chinese, Russian, and other non-English languages if these language are not their native language. Accordingly, few scientists are able to read Chinese or Russian and other non-English papers. Consequently, these non-English papers have less chance to be cited. Our investigation on the documents, which cite the papers published in the nine non-English journals, reveals that language self-citation is the primary factor leading to biased citations in these journals.

Another reason causing the biased citation can be attributed to the database we used. Our statistics deal with the Chinese, French, German, Russian and Spanish journals and papers within the Thomson Reuters database. Though Thomson Reuters indexes some non-English journals, and some English journals also publish non-English papers, the proportion of these non-English journals and papers is rather low. Hence, the potential citing pool of non-English papers and journals covered by the Web of Science is not big enough to form a citing source of large scale. It is in this context that local databases (e.g. CSTPC, CSCD and CNKI in China; Scielo in Latin America) may play an important role as a counterbalance to observed preference to English articles in international databases (besides of course as retrieval tools).

We discussed unbalanced citing to non-English journals in an English-language environment. Naturally, we may ask ourselves: is there a reciprocal phenomenon that English-language journals are cited less in a non-English language environment, for example, Chinese-language environment? We made an investigation on physics journals published in China. Based on the CSTPC (Chinese Scientific and Technical Papers and Citations) database, created by the Institute of Scientific and Technical Information of China, we compared the impact factors of English-language journals and Chinese-language journals in physics. The results are shown in Table 3. It seems that the results only partially support the reciprocal phenomenon. Often English-language journals are not very popular, but still, some rank among the top five of more than thirty physics journals.

Table 3 Impact factors of English-language journals and Chinese journals published in China in physics

Journal title	2007		2008		2009	
	Impact factor	Rank in 33 journals	Impact factor	Rank in 34 journals	Impact factor	Rank in 36 journals
ADVANCES IN ATMOSPHERIC SCIENCES	0.674	10				
CHINESE JOURNAL OF CHEMICAL PHYSICS	0.450	21	0.272	28	0.222	30
CHINESE PHYSICS	1.960	1				
CHINESE PHYSICS B			0.714	11	1.108	5
CHINESE PHYSICS C			0.354	24	0.221	31
CHINESE PHYSICS LETTERS	0.555	17	0.956	7	0.544	12
COMMUNICATIONS IN THEORETICAL PHYSICS	0.484	19	0.405	22	0.400	19
CHINESE OPTICS LETTERS			0.890	8	1.415	2

Finally, returning to Garfield's observations, it may simply be that non-English papers are just of less quality. We offer the following approach for discussion and investigation. Consider a non-English paper, say a German one, and determine a comparable English-language one. By *comparable* we mean: dealing with the same, or at least a very similar topic, published in the same publication year, and published in a journal in the same subfield. Then a panel of scientists that are not from an English speaking country and not from a German speaking country, but can fluently read these two languages, performs a peer review of each of the two comparable articles. When this is done for a large number of articles (say 100) then one would know if there really is a quality difference. Our guess is that on average there is a difference and this for two reasons. If a scientist is able to he/she will, in many cases, try to publish in a highly visible English-language journal; second if a scientist is poor at English he/she may not know the English literature very well, or not at all, and hence lag behind. Of course it would be very interesting to detect, using the approach we outlines, non-English top quality articles, unknown to most Anglo-Saxon scientists (remember the Sputnik era!).

Acknowledgements

This work is sponsored by the National Natural Science Foundation of China. NSFC project 70973118.

References

Bookstein, A. & Yitzhaki, M. (1999). Own-language preference: a new measure of "relative language self-citation". *Scientometrics*, 46(2), 337-348.

Egghe, L. & Rousseau, R. (2004). How to measure own-group preference? A novel approach to a

sociometric problem. *Scientometrics*, 59(2), 233-252.

Garfield, E. (1978). Do French scientists who publish outside of France and/or in English do better research? *Current Contents*, May 29. Reprinted in: *Essays of an Information Scientist*, 3 (1980), pp. 498-503. Philadelphia: ISI Press.

Garfield, E. (1979). Journal citation studies 32. Canadian journals, Part 2: an analysis of Canadian research published at home and abroad. *Current Contents*, August 20. Reprinted in: *Essays of an Information Scientist*, 4 (1981), pp. 249-253. Philadelphia: ISI Press.

Grupp, H., Schmoch, U. & Hinze, S. (2001). International alignment and scientific regard as macro-indicators for international comparisons of publications. *Scientometrics*, 51(2), 359-380.

Institute of Scientific and Technical Information of China (2007, 2008, 2009). Chinese S&T Journal Citation Reports. Beijing: Scientific and Technical Publishing House.

Lin, MW. & Zhang, JJ. (2007). Language trends in nanoscience and technology: the case of Chinese-language publications. *Scientometrics*, 70(3), 555-564.

Marshakova, I. V. (1973). System of document connections based on references (in Russian). *Nauchno-Tekhnicheskaya Informatsiya, series 2*, 6, 3-8.

Ren, SL. & Rousseau, R. (2002). International visibility of Chinese scientific journals. *Scientometrics*, 53(3), 389-405.

Small, H. (1973). Co-citation in the scientific literature: a new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265-269.

Van Leeuwen, T. N., Moed, H.F., Tijssen, R.J.W., Visser, M.S. & van Raan, A.F.J. (2000). First evidence of serious language-bias in the use of citation analysis for the evaluation of national science systems. *Research Evaluation*, 8(2), 155-156.

Van Leeuwen, T. N., Moed, H.F., Tijssen, R.J.W., Visser, M.S. & van Raan, A.F.J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51(1), 335-346.

Van Raan, T., van Leeuwen, T. & Visser, M. (2011). Non-English papers decrease rankings. *Nature*, 469, p. 34.

Yitzhaki, M. (1988).The language barrier in the humanities: measures of language self-citation and self-derivation – the case of biblical studies. In: *Informetrics 87/88* (Egghe & Rousseau, eds), pp. 301-314. Amsterdam: Elsevier.