



Identifying and mitigating linguistic inequalities in the management of patent information in Europe

Gazzola, M. (2015). Identifying and mitigating linguistic inequalities in the management of patent information in Europe. *World Patent Information*, 40. <https://doi.org/10.1016/j.wpi.2014.12.001>

[Link to publication record in Ulster University Research Portal](#)

Published in:
World Patent Information

Publication Status:
Published (in print/issue): 15/02/2015

DOI:
[10.1016/j.wpi.2014.12.001](https://doi.org/10.1016/j.wpi.2014.12.001)

Document Version
Publisher's PDF, also known as Version of record

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.



Identifying and mitigating linguistic inequalities in the management of patent information in Europe



Michele Gazzola*

Institut für Finanzwissenschaft, Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany

ARTICLE INFO

Article history:
Available online 22 January 2015

Keywords:
European patent office
Language policy
Patent information
Translation costs
Language
Multilingualism

ABSTRACT

The current language regime of the European Patent Office, based on three official languages (namely, English, French and German) entails different types of inequalities among European users of the European patent system. Such inequalities concern the distribution of translation costs borne by European applicants when they file a European patent application, and the costs to access patent information published by the Office. This article identifies and characterises inequalities occurring at the level of patent information management, and it discusses some possible measures to mitigate them. Three measures are proposed, that is, rationalising and harmonising the European patent information system, introducing tools to facilitate the retrieval of multilingual information, and enhancing the infrastructure for the dissemination of knowledge on the use of patent information. This article also presents some unpublished figures revealing that the European patent system is more multilingual than commonly believed. There is a mismatch between the current language regime of the Office and the actual needs of European innovators for multilingual patent information. These results confirm the relevance of new initiatives to manage patent information more effectively by reducing existing inequalities in this area.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction¹

Policies related to the use of languages in Intellectual Property (IP) organisations (or “language regimes”) can have differing effects on users interacting with such organisations. First, the choice made by a patent organisation about its official languages can asymmetrically affect the costs of access to patenting procedures borne by applicants whose primary language is one of the organisations’ official languages, and the costs borne by users whose primary language is not official. The latter, for instance, must bear the translation costs at the moment of filing, during the patent application procedures, substantive examination of the patent application, various communications with the office and the negotiations dealing with possible amendments to the scope of the patent application. Second, a language regime determines the extent to which patent information must be published in (and possibly translated into) a given language. This, in turn, has an effect on the

relative information costs borne by users who need such information. As a result, inequalities in the distribution of costs develop within the patent system and more widely in the technology market.

The European Patent Office (EPO), for example, has three official (or procedural) languages, that is, English, French and German. European patent applications are accepted, examined and published only in these three languages. The trilingual language policy of the EPO entails different types of cost inequalities between applicants whose first language is English, French and German, and all other users of the European patent system. Such inequalities are not properly identified and discussed in the literature. The debate on the language-related costs of the European patent system, in fact, has often revolved around the excessiveness of translation validation costs in the Contracting States of the EPC at the post-grant stage (e.g. [1,2]), but little attention has been paid to the effects of a trilingual language regime on the costs of access to patenting procedures before a patent application is filed, during the patent application procedures and the substantive examination of the application. Moreover, inequalities at the level of patent information costs are usually ignored. The existing literature on information costs in patents and IP (e.g. [3,4,5]) mentions the role of language in patents as an aspect affecting information costs for

* Tel.: +49 (0)30 2093 99411.

E-mail address: gazzola@hu-berlin.de.

¹ List of non-standard abbreviations: LA (London Agreement), CLIR (Cross-lingual information retrieval), EPC (European Patent Convention), MT (Machine Translation), PATLIB (Patent information centres), PCT (Patent Cooperation Treaty).

users of patent information such as individual inventors, industry, consultants, academics, but only from the point of view of those who may find patent jargon an obstacle. In other words, these contributions address the issue of the complexity of *language* used in patents, but not the issue of *languages* or linguistic diversity in IP policies.

The purpose of this article is to contribute to filling this gap by clarifying how the language regime of the EPO influences the distribution of costs of access to patenting procedures and in particular the distribution of information costs in the European patent system. It shows that the choice of policy-makers of looking at translations and interpreting as pure cost — or sometimes a deadweight loss that should be reduced as much as possible [6] — obfuscates the importance and the relevance of the implicit costs of language regimes, that is, the language-related costs borne by users of the European patent system that arise from their interaction with the patenting authority. As Pool and McFann correctly note: “It is wrong to claim that having many official languages is necessarily inefficient. As more native languages are made official, translation costs rise but adoption costs fall. The tendency to regard multiple official languages as inefficient may reflect a state-centred neglect of costs incurred by individuals in adapting to language policies” [7].

The article is structured as follows. Section 2 characterises the distributive effects related to the EPO’s language regime, focussing on inequalities emerging at the level of patent information costs. Section 3 proposes some measures at the level of information management that could be undertaken in order to mitigate existing linguistic inequalities, even without the need for a major reform of the current EPO language regime. Section 4 reveals that the European patent system is more multilingual than commonly believed, and that there is a mismatch between the current language regime of the EPO and the actual needs of European innovators for multilingual patent information. These results confirm the relevance of new initiatives to manage patent information more effectively and to reduce existing inequalities in this area. Section 5 summarises and concludes the article.

This article focuses on the European patent and on the 38 Contracting States of the European Patent Convention (EPC). It does not examine the agreement on a European patent with unitary effect (“unitary patent”) approved by 25 Member states of the European Union through an enhanced cooperation in 2012. At the time of writing of this article, in fact, the agreement on a Unified Patent Court—the third and last component of the “patent package” setting up a single and specialised patent jurisdiction—is still not in force, and therefore the unitary patent is not a reality yet. However, some implications of the translation arrangements of the unitary patent adopted in 2012 are briefly discussed in Section 2.2.

2. The distributive effects of the EPO’s language regime

2.1. Background: the language regime of the EPO

The official languages of the EPO are English, French and German. Applicants who choose to file in a non official language are subsequently required to produce a translation of their application into one of the EPO’s official languages within two months of filing (Rule 6 of the EPC). The elected official language is used for all the proceedings before the Office, and for the publication of the European application filed and the European patent granted. During the patenting process, all written and oral communication between the applicant and the Office takes place in the designated EPO official language. Applicants can use any non-official language in oral proceedings (opposition and appeal), but they must provide and

thus pay for interpretation services.² When the patent application meets the patenting criteria, the Office issues an intention to grant. At this stage the applicant has to file a translation of the claims in the two official languages of the EPO other than the language of the proceedings (Rule 71 EPC). Since all applicants must provide such translations, no distributive consequence occurs at this stage. After the patent grant, European patents need to be validated at the national level. Such validation requires the payment of validation fees, and then of renewal fees every year until protection is sought up to a period of 20 years. In the majority of the EPC Contracting States it is required to provide the translation of the patent document (or parts of it, usually the claims) into one of the official languages of the countries in which patent protection is sought (see Section 2.3 for more details).

2.2. Implicit costs of language regimes

There are four categories of language-related implicit costs associated with the EPO language regime [8], that is:

1. *Admission costs*: costs of translating a European patent application and other relevant documents which had been originally filed in an admissible non-EPO language. Recall that “admissible non-EPO languages” are the Contracting States’ official languages which are not English, French or German. The average admission cost for a standard patent application is at least € 1700 ([6], adapting [9]).
2. *Interaction costs*: translation costs related to communication exchange (oral and written) between the user and the Office on formalities as well as substantial issues (e.g. amendments to the claims, opposition and appeal proceedings). No figures on these costs are available. However, they are likely to be far from negligible, as the process of negotiation with the EPO (especially the examination of patent applications) can last several months or even some years.
3. *Granting costs*: costs of translating the claims into the two official languages of the EPO other than the language used in proceedings. The average cost of translating the claims into two official languages is € 680 ([6]).
4. *Information costs*: costs of access, retrieval and comprehension of patent information published by the Office in a language that users do not understand or do not master fluently. Information costs arise from searching for an existing technical solution in prior art for the purposes of patent landscaping and freedom to operate analysis, drafting a new application, and opposition purposes. As for interaction costs in opposition and appeal proceedings, information costs are incurred not only by applicants involved in patenting procedures but also by third parties carrying out information monitoring and analysis. Information costs in patents are not, of course, exclusively language-related, and their distribution is influenced by several factors such as, for instance, the extent of digitalisation of literature, the level of technical know-how of users willing to access the literature and the development of information management skills. This article, nevertheless, focuses on the role of language policies in explaining the inequalities in the distribution of information costs.

Note that the concept of implicit cost is relevant also for international or multinational companies. Even if the staff is multilingual and therefore capable of effectively working in a foreign

² Rule 4 provides that the party has to provide for interpretation. This means that the party has to pay for interpretation services and find the interpreters.

language, this does not mean that admission, interaction or information costs are zero. The cost of multilingual staff is, on average, higher than the cost of monolingual staff [15], and the use of external language services for certain purposes (e.g. editing and proofreading) may be required.

The distribution of implicit costs among users is affected by three aspects characterising the EPO's language regime, that is, (i) the number of official and working languages; (ii) the extent of the translation of documents supplied; and (iii) the system of financial compensation for language-related costs with rebates for some filers. The current language regime of the EPO generates two types of inequality among European users, which, at least to a certain extent, can bias the competition among European inventors through an uneven access to the protection of innovation.

First, at the level of admission costs to patenting procedures and interaction with the EPO, the current language regime of the Office generates an unequal distribution of costs between users whose primary language is one of the EPO official languages and users whose primary language is an admissible non-EPO language. The languages of the EPO are English, French and German, but European applicants are resident in 38 EPC Contracting States with 28 different official languages. The compensation scheme in force before April 2014 provided for a fee reduction of 20% in some but not all fees (e.g. € 40 on the filing fee and € 311 on the examination fee). Such compensation was not sufficient to cover the admission and interaction costs. Some authors ([10,8]) have assessed the impact of admission and interaction translation costs on the global average cost that an applicant must bear in order to get a patent granted (this includes, among other things, general fixed costs that cover application, search, examination, grant and renewal fees up to the fifth year of the patent). Results reveal that the average cost of accessing patenting procedures for European applicants whose first language is not English, French or German before April 2014 was at least 30% higher than the average cost borne by English-, French- or German-speaking applicants (€ 8012 vs € 6180 per patent). An English-only language regime would increase such inequality to at least 34%. Besides, the current language regime of the EPO creates an awkward paradox: patenting innovations in Europe is *de facto* cheaper for non-European applicants who are resident in a country sharing an official language with the EPO (e.g. Canada or the United States) than for the majority of the residents in the EPC Contracting States.

From April 2014, the compensation rebate was increased to 30%, but only for small and medium-sized enterprises, natural persons, or non-profit organisations, universities or public research organisations. This reduced the existing language-related cost inequalities to 27%. Large companies, however, are no longer entitled to any compensation rebate. As a result, the reform of the compensation scheme of the EPO does not solve the problem of costs inequalities among European applicants. Some still existing disparities could be further reduced (albeit not cancelled) in the future when the European unitary patent will enter into force, but only for applicants who are residents in the 28 Member States of the EU.³

³ The Council Regulation (EU) No 1260/2012 of 17 December 2012 implementing enhanced cooperation in the area of the creation of unitary patent protection with regard to the applicable translation arrangements provides that translation costs from one of the official language of the EU into one of the official language of the EPO will be compensated, but only up to a certain ceiling. It is still not clear what the maximum level of the compensation will be, but it is unlikely that admission translation costs will be entirely reimbursed. Moreover, the compensation scheme applies only to admission translation costs: other translation costs (e.g. interaction translation costs) are not taken into account. Finally, Council Regulation applies only to SMEs, natural persons, non-profit organisations, universities and public research organisations (thus, large companies are not included). Note that no compensation is possible for applicants who are residents in EPC Contracting States that are not member of the EU (e.g. Turkey or Norway).

The second type of distributive effect of the trilingual language policy of the EPO regards information costs. The trilingual language regime of the EPO, in fact, affects the stock and the nature of multilingual patent information available to European inventors, and it biases the distribution of the costs to have access to such information. As patents published are of the most important sources for the creation of new patents, lack of multilingual patent information can have a negative impact on the innovation process [10]. In the next section, the main existing inequalities at the level of distribution of information costs will be identified and characterised.

2.3. Inequalities in the distribution of information costs

The Patent literature is recognised as one of the main sources of specialised knowledge for the production of innovation, and indirectly of new patents ([11,12]). Burrone and Singh Jaiya [12], for example, estimate that patent documents contain 70% of the world's accumulated technical knowledge. This knowledge is necessarily published in at least one language and language policies in IP organisations affect precisely the extent to which patent information is available in different languages. The value of patent information is not limited to its importance as a source of prior art knowledge, but it is intrinsically related to the correct functioning and balance of the patent system. Language of publication is a vital aspect of the patent bargain between innovators and society since the granting of a temporary monopoly to an individual is given in exchange for a gain in information for society as a whole, via disclosure.

There are two channels through which the availability of multilingual patent information in the European patent system is affected, that is:

1. *A limited access to multilingual patent information.* In the European patent system patents are published either in English, French or German (only the claims are available in the three languages at grant). Hence, innovators whose primary language is not one of the Office's official languages have at their disposal, when the patent is granted, a reduced collection of information on the basis of which they can make decisions on their R&D or technology management strategies. European patent applications and granted patents are only published in one of the Office's official languages even when they were first filed in an admissible non-EPO language. In spite of being non-authentic, the applications originally filed in admissible non-EPO languages are a valuable source of high quality multilingual patent information which could be exploited for a number of activities (with the possible exclusion of legal analysis). These multilingual applications are not always available on the EPO's free-of-charge databases such as *Espacenet* and the *European Patent Register*.

Obviously, users can wait and rely on translations provided by the patent holder when the patent granted is validated at the national level, but in several cases such translations are not available either because the patent granted is not validated in a given country (and therefore not translated into the local language) or because of the application of the London Agreement (cf. below). These users will bear relatively higher costs and undergo time-consuming efforts to process information, which is likely to slow down their responsiveness and decision-making with respect to changes in the technology market. This is particularly relevant for SMEs, where lack of skills in foreign languages is more frequent than in large companies ([13,14]), and outsourcing information monitoring services is not always a viable option.

2. *The London Agreement.* The London Agreement (LA) was signed in 2000 and came into force in 2008 modifying the application of Article 65 of the EPC according to which national patent offices may require a translation of granted European patents for validation. The states who are parties to the LA choose to dispense partially or totally with translation rights. The LA creates a situation in which a patent can take force in a given national territory in spite of being written in a foreign language. Up to the present, 18 out of 38 countries have ratified the LA. Albania has *de facto* adopted it. Of those 19 states, only eight have completely waived their translation rights namely, Liechtenstein, Luxembourg, France, Germany, Monaco, Switzerland, the UK and Albania. Ireland has recently adapted its national law in accordance with the LA provisions and ratification is expected in the near future. In practice, the only countries which chose to dispense totally with translation requirements (with the exception of Albania) are those which share at least one official language with the Office, and hence have access to a translation of the claims in their (or one of their) official language(s) anyway. The remaining nine states that have ratified the LA still require either a translation of the claims into their official language or/and a translation of the whole document into English. Generally speaking the LA affects transparency and the principle according to which law must be clearly understandable by those who are subject to it [19], and is likely to increase the information costs borne by applicants. In fact, the LA makes it possible to decrease the post-grant translation costs ([2,1]), but it also reduces the extent to which patent documents are translated, indirectly affecting the stock of legal and technical information available in a given language (in particular as regards the description of the invention), and the relative information costs borne by users.

In sum, English-, French-, or German-speaking users, including some non-European applicants, have a competitive advantage with respect to other European innovators who incur higher information costs and time-consuming efforts to process information, with a potential effect on innovation output patterns and patent behaviour. As a result, the latter may decide to dispense with the use of patent information or carry out partial research exposing their companies to a number of risks: patent infringement, duplication of research efforts, failure to evaluate signals of change in the technology market. In addition, the inequalities in the distribution of information costs can influence patent behaviour and innovation output: a lack of or only partial information could affect the quality of future applications, lowering the probability of patent grant; failure to identify rights' holders and licensing partners can hinder follow-on improvements; patent information helps in matching inventors with producers and in strengthening the link between research and industrial application. Finally, inequalities in the distribution of information costs reduce transparency in the technology market; knowledge of the available technologies and their alternatives provides technology purchasers with a greater bargaining power, placing them in a better position to negotiate a license, and technology transfer contracts.

Access to patent information can be improved through machine translation (MT), on which the EPO is increasingly relying (for example through the EPO-Google Patent Translate software). However, inequalities in the distribution of information costs would remain since the readability of machine-translated texts is notably worse than standard translation. Jones et al. [16] have carried out three experiments to compare the readability of MT with human translation. They conclude that "reaction times, [that is, the time for a user to understand and react to the information she is processing,] are slower when reading MT output in

comparison with human translation" [16]. The quality of machine translation, especially that provided by Google Translate, has increased and it can provide "a fairly good performance at translating individual words and phrases [...], but not translating complex sentences" [17]. Nevertheless, "in no case, either in or outside the patent world, can MT be considered a substitute for human translation even when its quality almost approximates the quality delivered by a human translator" [18].

3. Mitigating linguistic inequalities in patent information management

To our knowledge, there is no quantitative estimate of the inequalities related to information costs presented in the previous section. Nevertheless they exist, and they can have an impact on the innovation process in Europe: as patent information is the most valuable input for further innovation, hindrances to patent information in a given language could have a negative impact of the innovation activities of those who work mainly in such a language [10]. Hence, it is no surprise that the awareness of the importance of appropriate strategies to manage multilingual patent information has increased in the specialised literature (e.g. [20,21]).

Some authors have looked at the possibility of introducing full compensatory monetary transfer to reduce cost inequalities at the level of access to patenting procedures [8], but little research has been carried out on inequalities in the distribution of information costs. Hence, this section discusses, with no claim of exhaustiveness, some measures at the level of the management of patent information that can mitigate inequalities in the distribution of information costs within the European patent system without the need to reform radically the current language regime of the EPO. These measures aim at expanding the corpus of information available in admissible non-EPO languages, improving the conditions of retrievability, and developing patent information awareness and management skills at the local level.

1. *Harmonisation and rationalisation.* Expanding the corpus of available information in admissible non-EPO languages does not necessarily mean translating more; it may simply imply making a more effective use of the multilingual information already available. A great amount of existing multilingual information is dispersed in the system among the EPO and national patent offices with different levels of accessibility and digitalisation of patent information. Harmonising levels of digitalisation and rationalising the information system by centralising available data on one single platform is a major priority in order to improve the dissemination of patent information in non-official languages. This, in turn, would help reduce the cost of understanding patent literature and, on the whole, lower information costs.

Rationalisation involves strong cooperation and the exchange of information, especially with the offices in EPC countries where the official languages are not English, French or German. National offices possess the Euro-direct applications originally filed in the various admissible non-EPO languages, and also the applications first filed nationally (claiming national priority) which are subsequently also filed at the EPO. In addition, national offices store translations of the European granted patents validated at national level. Human translations for validation purposes have represented an essential input source to feed the EPO-Google Patent Translate software. These data are usually made available on *Espacenet* and the *European Patent Register*, but they can only be retrieved through a manual search (that is, one has first to find a given application in its language of procedure through search engines; then the original version of the

application can be found by manually seeking the scanned PDF version in the interface of the search engine). A related problem regards the conditions of accessibility, which may substantially vary depending on the national office providing them. Finally, translations of the claims of published European applications are often available in several admissible non-EPO languages since they can be requested by member states to enforce provisional protection on their territory according to the Article 67 EPC. These translations are also a useful source of patent information and they could be made available on a centralised search engine.

2. *Better information retrieval tools.* In order to lower costs related to information retrieval, cross-lingual retrieval tools could be implemented on patent databases which allow users to perform prior art searches in their own language. The intangibility of inventions on one side and the ambiguous language of patents on the other produce a jargon which is difficult to comprehend and which requires experience of patent specific designation for keyword searches. Since 2010 cross-lingual information retrieval tools (CLIR) have been available on *Patentscope*, the database of the World Intellectual Property Organisation (WIPO). WIPO's CLIR provide synonyms and an automatic translation of search queries by using statistical MT⁴ developed internally with the software *Moses* and trained with abstracts and titles. As discussed before, MT today can provide an acceptable performance at translating individual words and phrases [17]. Search results on CLIR can be refined according to the relevant technical area so as to reduce noise and errors. The application of MT technologies to cross-lingual tools allows users to search for patents using inputs from their primary language, a function that MT on patent databases alone does not offer. The use of classification codes, at least in part, has helped to better manage multilingualism language in data retrieval; however, a keyword search remains a complementary strategy to classification definition, often with the purpose of tailoring the queries to the users' needs.
3. *Improving awareness.* Finally, increasing awareness of the possibilities offered by the patent information system, notably by strengthening patent information management skills, knowledge of patent jargon patterns and the use of classification codes could lessen the costs of retrieving and understanding patent literature for all innovators. Resulting from the cooperation between the EPO and the EPC Contracting States national patent offices, a network of patent information centres (PATLIB) was established on the EPC territory. PATLIBs play a pivotal role in disseminating patent information and providing assistance at the local level. At present there are more than 300 such centres in Europe. Some function as simple patent libraries while others offer more sophisticated products and services. The EPO is currently attempting to harmonise these centres' activities across the EPC region in order to transform them into cutting-edge information providers for non-experts also. The advantage of PATLIBs is their connection with, and proximity to, the local industry as well as their ability to work in the local language. Despite these recent efforts PATLIBs are not as well-known as they should be, especially among SMEs and local authorities which often underestimate the potential of patent information or lack the appropriate training to exploit these data successfully.

Before concluding, it is necessary to put in perspective our proposals to reduce the inequalities in the distribution of

information costs in the European patent system by addressing and discussing the question of their practical relevance. The claim is sometimes made that multilingualism as such is a barrier to the spread of patent information and that an English-only language regime would correspond to real-life practices as English is already the most frequently used language in patenting procedures at European level (e.g. [22]). The next section shows that the reality is more complex and that the optimality of an English-only solution is open to question.

4. The linguistic profile of European patent holders

From 1978 to 2011, English was the procedural language of 71.5% of all European patent applications filed with the EPO, followed by German (21.6%) and French (6.9%).⁵ The frequency of use of the three official languages changes if one focuses on patent granted: between 1978 and 2011 English was the procedural language of 64.7% of all European patents granted, followed by German (26.6%) and French (8.7%).

These figures, however, refer to the overall number of European patents granted or the number of European patent applications filed, including those granted to, or filed by, intensive users of the European patent system such as the United States and Japan. Results change considerably if one focuses on European applicants, that is, natural or legal persons having their residence or principal place of business within an EPC Contracting State. This choice can be justified by the fact that the EPO was established with the purpose of fostering and promoting innovation in Europe. In other words, the EPO is supposed to serve primarily the interests of its constituent countries. This does not mean, of course, that users from non-EPC countries must be discriminated. A European patent is in principle equally accessible for European and non-European users. Nevertheless, the fact that the current compensation scheme for translation costs applies only to European applicants and not, say, to Asian applicants,⁶ reveals that the former already benefit from a special treatment justified precisely in terms of the promotion of the interests of the EPC Contracting States.

In what follows, the number of European patents granted is used as a proxy for innovative activities instead of the number of European patent applications. Patent granted are, in fact, a better indicator of innovation than patent applications, as the number of patents granted denotes inventions that are industrially applicable, novel and non-obvious. Recall that, on average, only 42% of European patent applications filed become actual patents [23].

Table 1 presents the proportion of European patents granted to European applicants for the period 1978–2011, by groups of countries defined according to their official language(s). Figures for the period 2007–2011 are reported to track recent trends. For European countries with more than one official language (e.g. Belgium, Finland, Luxembourg, Switzerland), using census data, the number of patents granted to the residents of such countries has been split according to the percentage of the population speaking a given language as a first language or a primary language of education. EPC Contracting States have been grouped according to their

⁵ If not specified differently, all tables and results presented in this section have been elaborated by the author on the basis of a comprehensive database kindly provided by the EPO's statistics unit. The author wishes to thank the statistics unit of the EPO for its support.

⁶ That is, natural or legal persons having their residence or principal place of business within a Contracting State having a language other than English, French or German as an official language, and to nationals of that State who are resident abroad (Rule 6 of the EPC). No compensation applies to applicants who are residents in non-European countries that do not share any official language with the EPO (say, Japan or the Republic of Korea).

⁴ In statistical machine translation translations are generated by a computer that learned how to translate from vast amounts of texts translated by humans.

Table 1
Proportion of European patents granted to European applicants according to their country of residence (historical and recent trends).

Group of countries by official language	Percentage of European patent granted to European applicants (1978–2011)	Group of countries by official language	Percentage of European patent granted to European applicants (2007–2011)
German-speaking	50.5%	German-speaking	49.9%
French-speaking	18.1%	French-speaking	17.2%
English-speaking	8.7%	Italian-speaking	7.5%
Dutch-speaking	6.8%	Dutch-speaking	7.1%
Italian-speaking	6.6%	English-speaking	6.7%
Swedish-speaking	4.2%	Swedish-speaking	4.9%
Finnish-speaking	1.7%	Finnish-speaking	2.2%
Denmark	1.3%	Denmark	1.6%
Spain	0.8%	Spain	1.2%
Others	1.3%	Others	1.7%
Total (in absolute terms)	597,341		151,723

Source: EPO Statistics Database.

official language(s), because a language regime has an impact on the relative position of applicants defined by language rather than nationality.

Innovators in German-speaking countries own half of all the European patents granted to European inventors, followed by residents in the French-speaking countries. European English-speaking countries are those with the third largest group of owners of European patents, followed by the Italian- and Dutch-speaking groups. European English-speaking countries, however, were overtaken by Italian- and Dutch-speaking countries in 2006 and in 2007, respectively.

Differences in the intensity of use of the European patent system by European inventors are also reflected in the languages used in filing. Table 2 shows the proportion of European patents granted to European inventors by languages in which the application was originally filed (for the periods 1978–2011 and 2007–2011).

The data show that on average German and not English is the language most frequently used by European users in filing patent applications that eventually turn into European patents. French still plays a significant role. The share of other languages used in filing such as Italian, Dutch and Swedish is relatively low, but this is partially due to the trilingual policy of the EPO.⁷ Several applicants, especially those who are resident in the Nordic countries, prefer to file European patent applications directly in one procedural language as they know that a translation into one of these languages must be provided in any case.

The share of European patent that were originally filed in English has been increasing over time, but the majority of European patent holders still do not use it at the moment of filing. In addition, English is not the most common language of procedure used by European patentees. From 1978 to 2011, for example, the language of procedure chosen by the holder of European patents was German in 48.3% of cases, followed by English at 35.8% and French at 15.9%. From 2007 to 2011, these percentages were 43.8% for German, 42.9% for English and 13.3% for French respectively.

These percentages, however, are just average values which do not reflect the complexity of the linguistic profile of European inventors. In fact, considerable differences exist between technological sectors, as shown in Table 3 in the Appendix. Table 3 reveals

Table 2
Proportion of European patents granted to European inventors by languages in which the application was originally filed (historical and recent trends).

Language	1978–2011	2007–2011
German	48.2%	43.8%
English	29.5%	36.7%
French	15.8%	13.2%
Italian	2.6%	2.9%
Dutch	1.3%	0.8%
Swedish	1.1%	0.8%
Spanish	0.6%	0.8%
Others	0.9%	1.0%

Source: EPO Statistics Database.

that the “importance” of a language (measured in terms of the percentage of European patents published in it) depends on the technological sector considered. Thus, inventors and applicants involved in patent intelligence do not necessarily have the same needs as regards patent information, as the relevance of information sought (and therefore the need for multilingual skills) depends on the technological sector concerned.

To summarise: the European patent system tends to be more multilingual than is commonly believed, and the current language regime of the EPO does not seem to match the needs and the interests of European users (and funders) of the services of the Office.⁸ For example, Dutch- and Italian-speaking patentees use the Office's services up to the grant more frequently than their European English-speaking competitors, but they do not face the same average procedural costs and they do not receive the same quality of patent information.

5. Conclusions

This article shows that the current language regime of the EPO generates several language-related cost inequalities, both as regards the costs of access to patent protection and the distribution of information costs. Such inequalities create an artificial competitive advantage for users who are national of, or resident in, English-, French- or German-speaking countries, and also the paradox that it is today cheaper for, say, a Canadian company to patent its inventions in Europe than it is for Spanish or Dutch inventors. An English-only solution, proposed by some observers, would exacerbate (and not reduce) existing inequalities in the distribution of implicit costs. One of the reasons is that the majority of European

⁷ The increasing popularity of the Euro-PCT route is a second factor to be taken into account [10]. As admissible non-EPO languages such as Dutch, Italian and Swedish are not languages of publications of the PCT system, translations into one of the official languages of the EPO must be already available when Euro-PCT patent applications are filed with the EPO. Usually, applicants file directly the translated patent application and not the original text in the national language. Note that Euro-direct applications are more likely to be filed in national languages than Euro-PCT applications [8].

⁸ Recall that the bulk of the EPO's revenue comes from procedural fees related to the patent grant process and national renewal fees for granted patents [23].

patents granted to European applicants is still drafted in French or German.

This article presents three measures that can reduce current inequalities at the level of patent information management, that is, rationalising and harmonising the European patent information system, introducing tools to facilitate the retrieval of multilingual information, and enhancing the infrastructure for the dissemination of knowledge on the use of patent information.

Let us note, in concluding, that reducing translation costs inequalities at the filing stage could have a positive effect on the cost-effectiveness of the European patent system as a whole. Implicit language-related costs, in fact, can be viewed as a form of implicit fee [8]. As reducing fees entails an increase in the number of patent applications filed [24], a reduction of language-related costs is likely to have a positive impact on the number of European patent applications filed per year. This, in turn, could generate extra income fee that could be used precisely to support a stronger compensation scheme for the translation costs borne by European applicants whose primary language is not one of the official languages of the EPO.⁹

Acknowledgements

The financial support from the Research Executive Agency of the European Commission (Project number PIEF-GA-2012-327225) and the Swiss National Science Foundation (project PBGEP1-136158 and project PBGEP1-145655) is gratefully acknowledged. The author wishes to thank Alessia Volpe and the two anonymous reviewers for their valuable contributions and their useful remarks, and to express his gratitude to the statistics service of EPO. Any errors and all interpretations are all my own.

Appendix

Table 3

Percentages of European patents granted from 2007 to 2011 to European applicants in the three procedural languages of the EPO, by technological sector defined according to the EPO's classification.

	German	English	French
1. Basic materials chemistry	43.1%	48.9%	8.0%
2. Biotechnology	17.9%	73.4%	8.7%
3. Chemical engineering	49.2%	37.1%	13.7%
4. Environmental technology	51.8%	32.7%	15.5%
5. Food chemistry	19.6%	70.5%	9.9%
6. Macromolecular chemistry, polymers	45.4%	46.8%	7.8%
7. Materials, metallurgy	48.3%	34.5%	17.3%
8. Micro-structural and nano-technology	31.4%	36.2%	32.4%
9. Organic fine chemistry	28.7%	57.6%	13.7%
10. Pharmaceuticals	18.7%	73.2%	8.0%
11. Surface technology, coating	52.2%	35.3%	12.5%
12. Audio-visual technology	32.7%	57.4%	9.8%
13. Basic communication processes	26.9%	60.0%	13.0%
14. Computer technology	26.2%	59.3%	14.5%
15. Digital communication	18.0%	70.7%	11.4%
16. Electrical machinery, apparatus, energy	52.7%	31.7%	15.6%
17. IT methods for management	33.3%	58.8%	7.8%
18. Semiconductors	45.4%	40.1%	14.5%
19. Telecommunications	23.2%	64.7%	12.1%
20. Analysis of biological materials	26.1%	64.9%	9.0%
21. Control	54.5%	32.0%	13.5%
22. Measurement	47.5%	33.0%	19.5%

⁹ See [10] for a comparison of the advantages and drawbacks of alternative compensation schemes.

Table 3 (continued)

	German	English	French
23. Medical technology	36.9%	52.6%	10.5%
24. Optics	34.1%	50.8%	15.0%
25. Engines, pumps, turbines	57.7%	26.9%	15.4%
26. Handling	47.9%	40.6%	11.6%
27. Machine tools	64.4%	25.9%	9.7%
28. Mechanical elements	61.9%	25.0%	13.1%
29. Other special machines	46.6%	40.6%	12.8%
30. Textile and paper machines	55.6%	36.4%	8.0%
31. Thermal processes and apparatus	57.5%	33.0%	9.5%
32. Transport	55.2%	25.5%	19.2%
33. Civil engineering	51.0%	36.1%	12.9%
34. Furniture, games	44.9%	40.5%	14.7%
35. Other consumer goods	41.8%	41.4%	16.7%
36. Unclassified	20.0%	60.0%	20.0%
Total	43.8%	42.9%	13.3%

Source: EPO Statistics Database.

References

- [1] Van Pottelsberghe B, Mejer M. The London Agreement and the cost of patenting in Europe. *Eur J Law Econ* 2010;29:211–37.
- [2] Harhoff D, Hoisl K, Reichl B, Van Pottelsberghe B. Patent validation at the country level – the role of fees and translation costs. *Res Policy* 2009;38:1423–37.
- [3] Long C. Information costs in patent and copyright. *Va Law Rev* 2004;90:465–549.
- [4] Cotropia CA. Patent claim interpretation and information costs. *Lewis Clark Law Rev* 2005;9:57–91.
- [5] Fromer JC. Patent disclosure. *Iowa Law Rev* 2009;94:539–606.
- [6] European Commission. Impact assessment. Accompanying document to the proposal for a Council Regulation implementing enhanced cooperation in the area of the creation of unitary patent protection with regard to the applicable translation arrangements. SEC(2011) 482/2. Brussels: European Commission; 2011.
- [7] Pool J, McFann B. The language auction: a nondiscriminatory method of choosing official languages. In: *International Symposium on linguistic human rights*; 1991. Tallin, Estonia.
- [8] Gazzola M, Volpe A. Linguistic justice in IP policies: evaluating the language regime of the European Patent Office. *Eur J Law Econ* 2014;38:47–70.
- [9] Berger Roland. *Study on the cost of patenting*. Munich: Roland Berger Market Research; 2004.
- [10] Gazzola M. The evaluation of language regimes. Theory and application to multilingual patent organisations. Amsterdam: John Benjamins; 2014.
- [11] Ceccagnoli M, Gambardella A, Giuri P, Licht G, Marian M. Study on evaluating the knowledge economy what are patents actually worth? The value of patents for today's economy and society. Final report for lot 1. Brussels: European Commission; 2005.
- [12] Burrone E, Singh Jaiji G. Intellectual property (IP) rights and innovation in small and medium-sized enterprises. Geneva: World Intellectual Property Organisation (WIPO); 2004.
- [13] CILT. ELAN: effects on the European economy of shortages of foreign language skills in enterprise. London: CILT, the National Centre for Languages; 2006.
- [14] COTANCE. Report on language needs in business. Companies' linguistic and language related needs in Europe (Results of the CELAN Project – language strategies for competitiveness and employability). Confédération des Associations Nationales de Tanneurs et Mégissiers de la Communauté Européenne (COTANCE); 2011.
- [15] Grin F, Sfreddo C, Vaillancourt F. The economics of the multilingual workplace. London: Routledge; 2010.
- [16] Jones D, Gibson E, Shen W, Granoien N, Herzog M, Reynolds D, et al. Measuring human readability of machine generated text: three case studies in speech recognition and machine translation. *Proc IEEE Int Conf Acoust Speech Signal Process* 2005;5:1009–12.
- [17] Li H, Graesser AC, Cai Z. Comparison of Google translation with human translation. In: Eberle W, Boonthum-Denecke C, editors. *Proceedings of the Twenty-Seventh International Florida Artificial Intelligence Research Society Conference* Palo Alto, California: AAAI Press; 2014. p. 190–5.
- [18] Rossi L, Wiggins D. Applicability and application of machine translation quality metrics in the patent field. *World Pat Inf* 2013;35:115–25.
- [19] Vianès G. *Brevet européen: les enjeux de l'accord de Londres*. Paris: Secrétariat d'Etat à l'Industrie; 2001.

- [20] Bonino D, Ciaramella A, Corno F. Review of the state-of-the-art in patent information and forthcoming evolutions in intelligent patent informatics. *World Pat Inf* 2010;32:30–8.
- [21] Edfjäll C. European patent information 2007: EPO policy reformulated. *World Pat Inf* 2008;30:206–11.
- [22] Van Pottelsberghe de la Potterie B. Europe should stop taxing innovation. *World Pat Inf* 2011;33:16–22.
- [23] EPO. Annual report 2009. Munich: European Patent Office (EPO); 2010.
- [24] De Rassenfosse G, Van Pottelsberghe de la Potterie B. On the price elasticity of demand for patents. *Oxf Bull Econ Statistics* 2012;74:58–77.



Michele Gazzola is Marie Curie post-doc Research Fellow at the Institute of Public Economics, Department of Economics, of the Humboldt-University in Berlin, Germany, and member of the Center for Applied Statistics and Economics (C.A.S.E) at the same university. He is also associated researcher at the “Economics-Languages-Training” Observatory of the University of Geneva, Switzerland. His research interests include the economics of languages, language policy and planning, applied public economics, policy analysis, and intellectual property. He is the author of different publications in these areas (www.michelegazzola.com).