

## International Journal of Legal Information

the Official Journal of the International Association of Law Libraries

---

Volume 37  
Issue 3 *Winter* 2009

Article 4

---

8-1-2010

# Accessing Legal Information Across Boundaries: A New Challenge

Ginevra Peruginelli

*Consiglio Nazionale delle Ricerche - Istituto di Teoria e Tecniche dell'Informazione Giuridica*

Follow this and additional works at: <http://scholarship.law.cornell.edu/ijli>

The *International Journal of Legal Information* is produced by The International Association of Law Libraries.

---

### Recommended Citation

Peruginelli, Ginevra (2009) "Accessing Legal Information Across Boundaries: A New Challenge," *International Journal of Legal Information*: Vol. 37: Iss. 3, Article 4.

Available at: <http://scholarship.law.cornell.edu/ijli/vol37/iss3/4>

This Article is brought to you for free and open access by the Journals at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in *International Journal of Legal Information* by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact [jmp8@cornell.edu](mailto:jmp8@cornell.edu).

## **Accessing Legal Information Across Boundaries: A New Challenge**

GINEVRA PERUGINELLI\*

*Institute of Theory and Techniques of Legal Information, National Research Council, Florence (Italy)*

### **Abstract**

In the actual multilingual and multicultural environment there is a significant need, in the academic world, in the legal profession, in business settings as well as in the context of public administration services to citizens, of common understanding and exchange of legal concepts of the various legal systems. At the same time, there is a strong pressure for the preservation of their basic sense and value. Both requirements are quite difficult to meet, and they are complicated by the complexity of legal language and by the variety of modalities used to express law within the various legal systems. Unlike a number of technical and scientific disciplines where a fair correspondence exists between concepts across languages, serious difficulties arise in interpreting law across countries and languages. This is largely due to the system-bound nature of legal terminology. This paper focuses on cross-language retrieval systems' ability to facilitate access to legal information across different languages and legal orders. As such, issues are addressed relating to linguistics and translation theory, comparative law, theory of law, as well as natural language processing techniques, while some recommendations are provided with the aim to contribute to cross-language retrieval of law.

---

\* The author has a degree in law from the University of Florence and a MA/MSc Diploma in Information Science jointly awarded by the University of Northumbria, Newcastle, UK and the University of Parma, Italy. From 2003 she has been admitted to Bar of the Court of Florence as a lawyer. She holds a PhD from the University of Florence, where her project focused on cross-language information retrieval system for law. Since 2000 she works at ITTIG-CNR. She carries out her research activities in various sectors, such as standardization in legal domain and law and legal language documentation. She is the author of various publications on integrated access to legal information and metadata strategies, and the evaluation criteria for legal web sites.

## 1. The context

Internationalization and increasing globalization of market economies and social patterns of life have created a situation where the need for legal information from foreign countries and from different legal systems is greater than ever before. This requirement is not new, but it is now becoming more and more crucial and hard to meet under the pressure of the rapid and complex cross-border transactions occurring between people of different legal cultures and languages. It is no doubt that the exchange of information is largely dependent on language. It is intended not only as a system of symbols, but also as a mean of communication, a tool for mediating between different cultures.<sup>1</sup>

If we consider the language of the law, its properties have a major impact on the exchange of legal information. In fact, the language of the law is the expression of legal identities that vary according to systems and countries where different languages are used to express legislation, case law and doctrine as main components of the various legal cultures.<sup>2</sup>

Europe is a typical example of a multi-language and multi-system environment where decisions on linguistics' policy are now receiving considerable attention.<sup>3</sup> In the European Union, full multilingualism is

---

<sup>1</sup> KJAER Anne Lise. A common legal language in Europe? In Van Hoecke, Mark (ed.). *Epistemology and Methodology of Comparative Law*. Oxford : Hart, 2004. ISBN 1841134430.

<sup>2</sup> SACCO, Rodolfo. Language and law. In Pozzo, Barbara (ed.). *Ordinary language and legal language*. Milano : Giuffrè, 2005, 1-21. ISBN 8814118048; FLETCHER, George P. The language of law: common and civil. In Pozzo, Barbara (ed.). *Ordinary language and legal language*. Milano : Giuffrè, 2005, 83-107. ISBN 8814118048.

<sup>3</sup> EUROPEAN COMMISSION. Promoting Language Learning and Linguistic Diversity: An Action Plan 2004 - 2006. COM (2003) 449 def., 24 July 2003. [http://ec.europa.eu/education/doc/official/keydoc/actlang/act\\_lang\\_en.pdf](http://ec.europa.eu/education/doc/official/keydoc/actlang/act_lang_en.pdf); EUROPEAN COMMISSION. A New Framework Strategy for Multilingualism. COM (2005) 596 def., 22 November 2005. [http://ec.europa.eu/education/policies/lang/doc/com596\\_en.pdf](http://ec.europa.eu/education/policies/lang/doc/com596_en.pdf); EUROPEAN PARLIAMENT AND THE COUNCIL. Decision n. 1934/2000/EC of the of 17 July 2000 on the European Year of Languages 2001. Official Journal of the European Union L 232, 14 September 2000; EUROPEAN COMMISSION. Final report of the High Level Group on Multilingualism, 2007. [http://ec.europa.eu/education/policies/lang/doc/multireport\\_en.pdf](http://ec.europa.eu/education/policies/lang/doc/multireport_en.pdf); GROUP OF INTELLECTUALS FOR INTERCULTURAL DIALOGUE set up at the initiative of the European Commission. A rewarding challenge: how the multiplicity of languages could strengthen Europe, 2008. [http://ec.europa.eu/education/policies/lang/doc/maalouf/report\\_en.pdf](http://ec.europa.eu/education/policies/lang/doc/maalouf/report_en.pdf).

claimed by providing a huge body of translated legal documentation, though English, French and German have a special status since the majority of the material is to be handled in these three languages.<sup>4</sup>

For economic and practical reasons, serious EU linguistic policy must manage problems of communicating across a plurality of languages. Proposals for a simplified choice, at least for certain contexts and specific documentation, are advanced through that linguistic policy. There are two opposite extremes under consideration regarding multilingualism management.<sup>5</sup> These are represented by a multilingualism embracing all European languages in an effort to be as equalitarian as possible (indeed a very expensive solution!) and, on the other side, by the adoption of a unique language, in particular a sort of international English which is already in place in some fields of law and specific legal areas – such as international trade – as well as in scholarly and professional settings.

It follows that multilingualism in the legal domain is almost unanimously perceived as a very complex issue, linked as it is to disciplines like comparative law, linguistics, translation theory and practice. It is a highly debated topic, not only among professionals and scholars of these various disciplines,<sup>6</sup> but also among government officials at national and international levels. This is demonstrated by the efforts made for the preservation and management of the plurality of languages in a number of countries as a guarantee of cultural diversity. This is the case of Belgium, Switzerland, Canada, and others.

In this context, those aspects of multilingualism which are crucial for the development of cross-language legal information retrieval systems are mostly relevant. On the one hand, these issues regard the intimate link between language and law, covering the crucial questions of rendering legal

---

<sup>4</sup> Gallo, Giovanni. Les jurists linguists de la Cour de Justice des Communautés européennes. In SACCO, Rodolfo and CASTELLANI, Luca (eds). *Les multiples langues du droit européen uniforme*. Torino : L'Harmattan Italia, 1999, 71-89. ISBN 8887605076.

<sup>5</sup> Moréteau, Olivier. L'anglais pourrait-il devenir la langue juridique commune en Europe? In SACCO, Rodolfo and CASTELLANI, Luca (eds). *Les multiples langues du droit européen uniforme*. Torino : L'Harmattan Italia-Isaidat, 1999.

<sup>6</sup> Groot, Gerard René de. Language and law. In *Netherlands report to the fifteenth International Congress of Comparative Law*. Antwerp/ Groningen : Intersentia, 1998, 21-32; SACCO, Rodolfo. Riflessioni di un giurista sulla lingua (lingua del diritto uniforme e il diritto al servizio di una lingua uniforme). *Rivista di diritto civile*, vol. 42, no 1, 1996, 57-65.

terms across languages. On the other hand, the broad spectrum of comparative issues, the relationship between legal systems which, while a problem in its own, is exacerbated in a multilingual environment. Every attempt to exchange legal knowledge among various communities and to reach a common understanding of different legal systems has inevitably to cope with the problems posed by language and systems' peculiarities. Such issues have a strong impact on the development and performance of cross-language legal information retrieval systems due to the complexity associated with mapping legal concepts across languages and systems.

## **2. Key aspects of cross-language legal information retrieval**

Over the past years, the study and development of methodologies for accessing multilingual general domain information has received a lot of attention from scholars. Although research of and systems for dealing with multilingual legal information are not as developed as they are for general domain information, some progress is underway concerning retrieval systems' implementation and cross-language indexing and searching tool production.<sup>7</sup>

---

<sup>7</sup> Initiatives vary in importance and formalization and include feasibility studies, research activities on legal translation, implementations of methodologies and tools for multilingual legal information management, experiments and applications of multilingual retrieval in specific fields of law. Among linguistic resources, legal translation's tools and retrieval systems: WordNet, a linguistic resource and a lexical research tool and JurWordNet, a terminological lexicon for the legal domain; LOIS - Lexical Ontologies for Legal information Sharing, a project aiming at creating a multilingual semantic network for the law domain; Transjus, a project promoting and coordinating studies on legal translation, linguistics and comparative law; LTS - Legal Taxonomy Syllabus, a database and software developed within the European project "Uniform Terminology for European Private Law"; IATE - Interactive Terminology for Europe, the database that holds all the terminology generate by European Institutions; EUROVOC, the multilingual and polythematic thesaurus of the European Union; Jurivoc, the legal multilingual thesaurus of the Swiss Federal Court and of its courts of social law; TransSearch, a database of legal translations in French, English and Spanish available on the Web and developed in Canada by the Research Group in Computer Aided Translation; a number of linguistic and terminological tools implemented in Canada (Termium, Fiches terminologiques bijuridiques, Multilingual Legal Glossary - Vancouver Community College); GLIN - Global Legal Information Network, the public database containing statutes, regulations, judicial decisions and other complementary legal materials from countries in Africa, Asia, Europe and the Americas made available by governmental agencies and international organisations. LexALP, a project for the harmonization of legal terminology used by the institutions operating in the Alpine Convention in their four official languages.

It is worth noting that major attention has been paid by scholars and legal experts to linguistic and conceptual aspects of legal languages: these themes are undoubtedly relevant to multilingual access and can provide important insights into the subject under investigation. The difficult task of effectively accessing multilingual legal material through information retrieval systems is matching and weighing legal terms across languages.<sup>8</sup> This generally implies translating from the language of the query to that of the material to be found or vice versa, and addressing the problem of word disambiguation, as ambiguity is greatly increased when mapping across legal languages. In fact, crossing the language barrier between search requests and documents implies addressing the problems of the system-bound nature of legal terminology and devising methods to map concepts across different legal systems.

In the legal domain, users must translate their information needs in the form of legal concepts. They put these into a query which must in turn be put into technical database concepts.<sup>9</sup> Legal information retrieval requires searching both structured and unstructured content. Data contained in legal texts such as identification codes, titles, dates and authors, as well as data for version management like, for example, criteria for validity of a statute represent structured information where the semantics is clearly determined. On the other hand, unstructured information which is communicated in natural language texts, quite extensively represented in legal information sources, contains a semantics which is much more difficult to represent in simple terms. All this causes problems for the retrieval of such information. Extensive research<sup>10</sup> addresses these problems by devising approaches and techniques aimed at enhancing index representations, query languages and matching functions to better capture the meaning of the information being handled. Enhancements are represented by adopting single terms, and by explicit modelling of the relations between these terms.

Access to information content can also be improved by adapting the knowledge representation in the index to the user's perspective on the information of the database containing the documents of interest. This can be achieved by considering the task for which the user needs the information.

---

<sup>8</sup> Sacco, Rodolfo. *Langue et droit*. In *Rapports nationaux italiens au XV<sup>ème</sup> Congrès International de Droit Comparé*. Milano : Giuffrè, 1998.

<sup>9</sup> Matthijssen, Luuk. A task-based interface to legal databases. *Artificial intelligence and law*, vol. 6(1), 1998, 81-103.

<sup>10</sup> Matthijssen, Luuk. *Interfacing between lawyers and computers: an architecture for knowledge-based interfaces to legal databases*. The Hague ; Boston : Kluwer Law International, 1999.

Methods taking into account such requirements are grouped in a functionality which is commonly called "intelligent interface" to information retrieval systems. This means that a facility by which the user is presented with a view on the information in the database (a "knowledge model", as denoted by theorists working at task-based IR systems) that corresponds to the domain in which this information is used.<sup>11</sup>

With the goal of improving access to legal information, efforts are also being made to study approaches and techniques to enhance the structure of elements of information contained in legal documents. This is realized through the design of document creation tools – the so called *drafting systems* – and through the development of advanced content analysis systems capable of re-ordering unstructured information.<sup>12</sup>

The themes which appear worthy of analysis as being closely related to the development and effectiveness of legal information retrieval systems mainly concern the relationship between law and language and comparative research of legal systems in relation to language issues. This leads one to consider the strong impact of both linguistic and comparative aspects on multilingual access to law.

### 2.1 Linguistic aspects

With the well known expression "the law is a profession of words,"<sup>13</sup> one references the fact that law is expressed verbally in legislative codes, court decisions, and through the prosecution and defence of criminal cases. Many of the problems concerning *meaning* that are of concern to language specialists turn out to be of interest to legal professionals as well. These have an impact on the exchange and retrieval of legal information. In fact, information retrieval systems are based on language, as queries are matched with the documents to be searched (be them free-text or metadata) through terms.

---

<sup>11</sup> Saadoun, Adel, Ermine, Jean-Louis, Belair, Claude, Pouyot, Jean Mark. A knowledge engineering framework for intelligent retrieval of legal case studies. *Artificial Intelligence and Law*, vol. 5, no 3, 1997, 179-205.

<sup>12</sup> Moens, Marie-Francine. Improving access to legal information. In Oskamp, Arno and LoddeR, Anja (eds). *Information technology and lawyers*. Dordrecht : Springer, 2006, 119-136. ISBN 1402041454.

<sup>13</sup> Mellinkoff, David. *The Language of the Law*. Boston ; Toronto : Little, Brown, 1963; Sacco, Rodolfo. *Langue et droit*. In *Rapports nationaux italiens au XVème Congrès International de Droit Comparé*. Bristol, 1998, Milano 1998.

The relationship between language and law has since long attracted the interest of both jurists and linguistics. It is still a big source of worry and concern in our modern society where the interrelation between different legal orders is commonplace. Both comparative jurists operating in academic environments and legal professionals more and more are faced with issues and cases where disparate legal models and concepts are circulated. As these are expressed in different languages, the problem arises to cope with these languages, with the practical implications of multilingualism, as well as with its theoretical principles.

Legal language consists of legal terms, phrases and stable conventions and as such it reflects one particular legal system, but in principle the multiple languages of law are not simply the national languages which transmit the contents of one or more law or systems, but they are a concern also among speakers of the same language, that which is proper to each category of receiver. The system-specificity of legal terms makes a relevant number of legal scholars and professionals state that the language of the law is to be learnt and communicated in its close relationship with a given culture, the related country and people's history and heritage, conceived as a socially acquired pool of knowledge which represents its richness and uniqueness<sup>14</sup>. As legal language is culture-bound and intertwined with one particular society and its legal system, it is seen as the collective memory of the legal actors belonging to a given legal system.

While the dependency of legal concepts of a particular legal system is the key characteristic of legal language as a system of symbols, such concepts are claimed as not forever fixed and unchangeable, as they change when legal experience changes<sup>15</sup>. The change of legal concepts is brought about through legal argumentation<sup>16</sup>. This is evidenced in the multifaceted role of language

---

<sup>14</sup> Moreteau, Olivier. L'anglais pourrait-il devenir la langue juridique commune en Europe?. In Sacco, Rodolfo et Castellani, Luca (eds). *Les multiples langues du droit européen uniforme*. Torino : L'Harmattan Italia, 1999, 143-162. ISBN 8887605076; Groot, Gerard René de. Language and law. In *Netherlands report to the fifteenth International Congress of Comparative Law*, Antwerp/ Groningen : Intersentia, 1998, 21-32.

<sup>15</sup> Luhmann, Niklas. *Das Recht der Gesellschaft*. Frankfurt/Main : Suhrkamp, 1993. ISBN 3518581686.

<sup>16</sup> Kjaer, Anne Lise. Convergence of European legal systems : the role of languages. *Language and culture*, no 29, 2004, 125-137.



in establishing, maintaining, and changing concepts, which makes cross-system interoperability even harder to achieve<sup>17</sup>.

A further argumentation is based on the branch of linguistics known as ‘linguistic relativity’ focusing on the fact that what one language system conceptualizes in one way is not conceptualized in the same way in all other language systems. This is especially true of legal terminologies at a system level<sup>18</sup>.

As mentioned above, legal language like language in general, can be viewed both as a system of conventional symbols and as a means of communication for people belonging to a particular social group or culture. When it is viewed as discourse, the focus is on its function as a means of communication. Discourse is defined as language used in social practice, communicative practice in a particular social group<sup>19</sup>; as such it is dependent on the social context in which it is used; it is shaped by that context, which is not an immutable entity.

All this has implications on the possibilities of legal communication and, since these changes are brought about in legal discourse, it is possible to come up with a convergence of the national legal systems and their languages. In addition, today legal discourse is no longer confined to the individual national legal systems, but it transcends national boundaries.

Different legal practices, diverse legal languages and cultures are exposed to each other and it is likely that, for example in Europe, gradually the national legal traditions will change along with the emerging intercultural

---

<sup>17</sup> Kjaer, Anne Lise [2]. Convergence of European legal systems : the role of languages. *Language and culture*, no 29, 2004, 125-137; Gemar, Jean Claude (ed). *Langage du droit et traduction : essais de jurilinguistique. The Language of the Law and Translation : Essays on Jurilinguistics*. Québec : Linguattech-Conseil de la langue française, 1982. ISBN 2920342134.

<sup>18</sup> Engberb, Jan. *Statutory texts as instances of language(s): consequences and limitations of interpretation*, 2004.  
[http://www.brooklaw.edu/students/journals/bjil/bjil29iii\\_engberg.pdf](http://www.brooklaw.edu/students/journals/bjil/bjil29iii_engberg.pdf)

<sup>19</sup> Fairclough, Norman. *Critical discourse analysis : the critical study of language*. London ; New York : Longmann, 1995. ISBN: 0582219809; Fairclough, Norman. *Discourse and social change*. Cambridge, UK ; Cambridge, MA : Polity Press, 1992. ISBN: 0745606741.

communication of legal actors, who in that way adapt themselves to the changing institutional context of law<sup>20</sup>.

Incorporating these requirements in cross-language information systems requires interpretation and adaptation strategies over languages and systems which is hard to accomplish without a high expertise in linguistics as well as in legal translation.

In particular, there is a general consensus on the view that translation is a complex form of action, requiring much feeling and understanding of cultural aspects. There are lots of ongoing discussions among linguists, scholars and professionals working in various settings and disciplines, all sharing the opinion that this activity is much more than the substitution of lexical and grammatical elements between two languages<sup>21</sup>.

To this regard it is also worth mentioning a statement on translation made by an outstanding author who makes an ongoing contribution to legal translation issues: “la traduction est nécessairement une lutte. Le bon traducteur est celui qui cherche, qui se pose des questions, qui, loin de se contenter de ce qu'il a trouvé d'abord, commence par s'en méfier; il est comme le médecin scrupuleux qui, son diagnostic a été à peine posé, cherche les indices qui pourraient le conduire à le remplacer par un autre mieux fondé. En matière de traduction, on ne pourrait dire que la première idée n'est jamais la meilleure.”<sup>22</sup>

In particular, in legal translation the demands of precision are greater than in literary translation<sup>23</sup>, as what is to be carefully taken into account is not only compliance with the rules of the foreign language, but also with the rules of the foreign legal system. Further complexity is due to the fact that, although legal translation demands precision and certainty, it is bound to use

---

<sup>20</sup> Kjaer Anne Lise. A common legal language in Europe? In Van Hoecke, Mark (ed.). *Epistemology. and Methodology of Comparative Law*. Oxford : Hart, 2004. ISBN: 1841134430.

<sup>21</sup> Capellas-Espuny, Gemma. The problem of terminological equivalence in international maritime law : the case of hypothèque and mortgage in the document Final Act and International Convention on maritime liens and mortgages. *Journal of diplomatic language*, vol.1, no 4, 2004.

<sup>22</sup> Gemar, Jean-Claude. La traduction juridique et son enseignement : aspects théoriques et pratiques. In *La traduction juridique*. Numéro spécial, *Meta*, vol. 24, no 1, 1979, 35-53.

<sup>23</sup> Avalos, Francisco. Legal translations: some tips. Lecture delivered at National Language Resource Center, San Diego State University on July 24, 1998

abstractions, whose meanings derive from particular changing cultural and social contexts. These contexts generate a certain degree of ambiguity, which increases when the legal cultures and systems are vastly different from each other<sup>24</sup>. It is also claimed that law is an unstable discipline, largely indeterminate, and legal discourse is also fluctuant, with its meaning depending on the language in which it is expressed and even depending on the target audience to which it is addressed.<sup>25</sup>

In contrast to what happens with disciplines as mathematics or chemistry, where there is an objective extra-linguistic reference, legal realities are conceived as the result of legal discourse which creates its own reality from different or shared historic traditions, in one or several languages, and which cannot coincide with the concepts of analysis or can only coincide partially when they focus on a common international legal phenomenon.

However, several different opinions are expressed denying the special status of legal translation and arguments are offered on equally specific disciplines, as, for example, astrophysics, where the target text must have effects in the special subject area<sup>26</sup>. Furthermore, the characteristics of the law as a system-bound discipline are typical of other subject areas like religion and political science where the notion of system is an inherent feature.

But on the whole, the opinion is widely shared that legal translators are more rigidly bound to specialized knowledge than the translators of everyday language or humanities<sup>27</sup> and that finding out terminological equivalence between terms is a serious problem when comparable concepts do not exist in the legal systems expressed by the languages to be mapped. In this context the danger of ambiguity and miscomprehension is considerable.

A field where the significance of legal translation is evident in many respects is international law. Since the right of States to communicate in their own language has been accepted, translation has become more important than ever in this field of law. Yet, very little attention has been devoted to language

---

<sup>24</sup> Rotman, Edgardo. The Inherent Problems of Legal Translation: Theoretical Aspects. *Indiana International and Comparative Law Review*, vol. 6, no 1, 1995, 187-196.

<sup>25</sup> Harvey, Malcolm: What's So Special About Legal Translation? *Meta*, XLVII, 2, 2002.

<sup>26</sup> Id.

<sup>27</sup> Gizbert-Studnicki, Tomasz. Is an Empirical Theory of the Language of the Law Possible? In Ziembinski, Zygmunt (ed.). *Polish Contributions to the Theory and Philosophy of Law*. Amsterdam : Bodopi, 1987, 99-114.

in international law and there is still the danger of the existing communication gap among nations.

Translation also matters greatly for international law in the area of international organizations. For example, the plurality of languages in the enlarged European Union creates a serious challenge to communication, which is the task of the large translation service of the Union in its effort to make materials available through legal translation.

For its operation, legal translation implies both a comparative study of the different legal systems and an awareness of the problems created by the absence of equivalents<sup>28</sup>. This means that legal translators must be familiar with the legal culture of the target language in order to reformulate an equivalent meaning through what they judge to be the most appropriate linguistic and legal expressions. In fact a particular concept in a legal system may have no counterpart in other systems, or a particular concept may exist in two different systems but may refer to different realities. In other words, law lacks a common knowledge base or “universal operative referents”<sup>29</sup>, which makes it very difficult to find equivalents for culture-bound terms, especially those concerning legal concepts, procedures and institutions.

On a practical level, one main problem legal translators are faced with is the poor quality of legal dictionaries, which severely hinders the possibility of conveying the meaning of the source legal language into the target one. According to Groot and Laer<sup>30</sup> who extensively analyse the problem of functional equivalence of legal terms, only very few dictionaries are reliable tools. These authors propose requirements and desiderata of these important instruments, as, for example, the indication of the degree of equivalence and the provision of alternatives according to area of law, system and use.

---

<sup>28</sup> Capellas-Espuny, Gemma. The Problem of Terminological Equivalence in International Maritime Law : The Case of Hypothèque and Mortgage in the Document Final Act and International Convention on Maritime Liens and Mortgages. *Journal of diplomatic language*, vol.1, no 4, 2004.

<sup>29</sup> Pelage, Jacques. La traductologie face au droit. In *Proceedings of the International Colloquium University of Geneva, February 17-19, 2000*. Berne: Association Suisse des Traducteurs, Terminologues et Interprètes (ASTTI), Ecole de traduction et d'interprétation de Genève (ETI), 2000, 125-131.

<sup>30</sup> Groot, Gerard René de, and Laer, Conrad J.P. van. The Dubious Quality of Legal Dictionaries. *International Journal of Legal Information*, vol. 34, no 1 (Spring 2006), 65-86.

The comparison of law and translation studies shows differences of perspective, especially in the application of the rule of equal authority which has not been used in translation studies, but also reveals similarities, both in the weaknesses attributed to translation (secondary nature of translation, authority of the original text). However, in spite of these common views, law and translation studies continue to develop separately even though it would plainly be advantageous if they shared some of their theoretical approaches<sup>31</sup>.

Legal translation is an essential function for cross-language retrieval systems. One major question concerns the translation strategy to be adopted in order to ensure that users access legal information independently of the language used in a query. As described above, legal translation mainly refers to texts, whereas in cross-language retrieval what mainly matters is handling single units of information or a combination of them as searched by users.

There are different approaches to legal translation which can be used for accessing multilingual legal information. The approach of fidelity to the letter of the original document, that is the strict adherence to the original, has long lasted over the centuries. Little by little the method of simple linguistic equivalence has given way to a target-oriented translation adopting a functionalist approach, where non formal correspondence between source and target text is to be sought, but to the equivalent legal effects principle.<sup>32</sup>

Despite the functionalist approach has received much attention, so far linguistic fidelity is still a popular approach, being recommended by the United Nations instructions for translators<sup>33</sup>. Furthermore, a narrow view of fidelity to the original text is favoured by Court interpreters<sup>34</sup>.

Other methods of translation as borrowing (translation procedure whereby the translator uses a word or expression from the source text in the target text) and creation of neologisms are scarcely significant in cross-language information retrieval, where the objective is to find materials in any language irrespective of the language used in searching.

---

<sup>31</sup> Lavoie Judith. *Droit et Traductologie: Convergence et Divergence*. In J.C. G mar and N. Kasirer (eds). *Jurilinguistique entre langues et droits*. Bruxelles: Bruylant, 2005. ISBN 2894001967.

<sup>32</sup>  ar evi , Susan. *New Approach to LEgal Translation*. The Hague ; Boston : Kluwer Law International, 1997. ISBN 9041104011.

<sup>33</sup> Harvey, Malcolm. *What's So Special About Legal Translation?*, *op. cit.*

<sup>34</sup> Morris, Marshall (ed.). *Translation and the Law*. Amsterdam ; Philadelphia: Benjamins, 1995. ISBN 9027231834.

In the final part of this contribution further thoughts and possible solutions are presented on this fundamental aspect.

## 2.2 *Comparative aspects*

The problems raised by multilingualism are strictly connected with those related to the variety and diversity of legal systems and as such to comparative law. Far from the opinion that pursuing comparisons may be limited to descriptive translations or summaries of foreign law, a number of comparatists<sup>35</sup> express their doubts about the possibility of a real comparison of legal systems. This does not mean ignoring that comparative research has reached very good results in putting scholars and legal professionals work together in comparative projects, launching harmonization activities and, at European level, having codes drafted as well as directives to be fitted with the legal concepts and structure of the member States.

Retrieval systems to legal information across different legal systems represent a practical approach to the confrontation and exchange of legal cultures; since comparison involves observation and explanation of similarities and differences, comparative research can give a major contribution to the development of these information systems. In fact, the implementation of retrieval functionalities implies taking into account and properly managing the peculiarities of legal concepts across systems, while handling the variety of languages used to express these various concepts, and addressing the terminological issues of representing the various legal cultures.

A glance to worldwide legal orders shows that several countries long since operate in a multi-system and multilingual environment: Canada, Switzerland, Belgium, Spain are only some examples of this, not to mention the case of Europe, with its 27 countries participating in the European Union, with their respective systems, languages and families of law. The experience shows that this pluralism is managed using different methods and practices, based on translation, interpretation, adaptations of legal terms, and in a number of cases on multi-language drafting.

---

<sup>35</sup> Hoecke, Mark Van. *Deep Level Comparative Law*. In Van Hoecke, Mark (ed.). *Epistemology and Methodology of Comparative law*. Oxford : Hart, 2004. ISBN 1841134430; Schlesinger, Rudolf B. *The Past and Future of Comparative Law*. *American Journal of Comparative Law*, vol. 43, Summer, 1995, 477-481; SACCO, Rodolfo. *Legal Formants: a Dynamic Approach to Comparative Law*. *The American Comparative Law*, vol. 1, 1991, 343-358.

Multilingualism and comparison among systems are often addressed as a joint main issue in cooperative efforts promoting harmonization activities for the creation of uniform law in various areas (at European level efforts have mainly been made in contract, private and trade law). It is a matter of fact that the direct implications of comparing and possibly integrating different legal concepts and structures are intimately linked to language issues.

Many corporatists are strongly concerned on the implications of the differences existing between the cultural contexts underlying the various legal languages and on the difficulties in transferring legal meanings and legal concepts from one legal system to another, even when the same language is used<sup>36</sup>. A number of frequently mentioned examples are made in the legal literature to refer to this phenomenon, such as *société* in French legal language in France, which has not the same meaning as *société* in French legal language in Belgium<sup>37</sup>. Similarly, *Besitz* means factual possession for a German; however, an Austrian lawyer understands *Besitz* as the factual possession including the *animus domini*, that is *Innehabung*. So even German speaking lawyers from Austria, Germany, Liechtenstein and Switzerland will not understand automatically each other's concept-based legal terminology<sup>38</sup>.

In recent years research studies increasingly concentrate on the relationship between legal language and comparative analysis of different legal orders. This topic, mainly debated in conferences, is often tackled from the point of view of the validity and performance of legal translation and of the analogy between legal translation and legal interpretation. In this direction many are the initiatives aimed at laying the foundation for a common frame of reference and at promoting, for example at European level, a pan-European terminology<sup>39</sup>.

---

<sup>36</sup> Kjaer Anne Lise. A Common Legal Language in Europe? In Van Hoecke, Mark (ed.). *Epistemology. and Methodology of Comparative Law*. Oxford : Hart, 2004. ISBN: 1841134430.

<sup>37</sup> Vanderlinden, Jacques. Le Future des Langues du Droit ou le Dilemma du Dernier Orateur. In Sacco, Rodolfo and Castellani, Luca (eds). *Les Multiples Langues du Droit Européen Uniforme*. Torino : L'Harmattan Italia, 1999, 193-222. ISBN 8887605076.

<sup>38</sup> Heutger, Viola. A More Coherent European Wide Legal Language. European Integration Online Papers (EioP), Vol. 7, no 2, 2004. <http://eiop.or.at/eiop/pdf/2004-002.pdf>.

<sup>39</sup> Pozzo, Barbara. Harmonisation of European Contract Law and the Need of Creating a Common Terminology. *European Review of Private Law*, Vol. 11, no 6, 2003, 754-767.

In addressing the issues related to the development of systems and tools for accessing legal information across legal systems, consideration is to be given to the methods employed in the comparative process of legal systems: integrative as opposed to contrastive<sup>40</sup>.

A brief historical outline is given of the two approaches with special reference to Europe, because such approaches are likely to influence the cross-system retrieval techniques adopted in the implementation of retrieval systems.

In the continental Europe for a number of centuries a *jus commune* emerged which did not mean an entirely uniform law, but certainly a set of shared formative elements of the law, which are called by Sacco "legal formants"<sup>41</sup>. With the age of codification, two facts contributed to the creation of intellectual barriers between the legal systems of the several nations: the abandonment of Latin and the adoption of national codes in each country's national language. This introduced a contrastive approach in the practice of comparative law and law professionals treated the national systems as real foreign law. It is only under the actual influence of trans-national exchange and increasing cross-border transactions in every sector of life, that a common core of legal systems has started to be sought and an integrative comparison has newly emerged among legal scholars.

The actual debate among comparative scholars is extremely rich and complex. It is claimed that original innovation in law is very small and borrowing and imitation is of central importance in understanding the course of legal change. But the focus is also on divergences in the peculiarities of common and civil law systems, namely in their formants, system's principles and rules, manner of reasoning of lawyers and use of authorities guiding them in legal questions. However, the possibility for fruitful convergence and mutual understanding is envisaged and encouraged<sup>42</sup>.

The convergence or divergence approaches mentioned above are key elements for implementing multilingual retrieval tools and services: according to the chosen approach, the methods followed in these systems will facilitate

---

<sup>40</sup> Schlesinger, Rudolf B. The Past and Future of Comparative Law. *American Journal of Comparative Law*, vol. 43, Summer, 1995, 477-481.

<sup>41</sup> Sacco, Rodolfo. Legal Formants : A Dynamic Approach to Comparative Law. *The American Comparative Law*, vol. 1, 1991, 343-358.

<sup>42</sup> Schlesinger, Rudolf B. The Past and Future of Comparative Law. *American Journal of Comparative Law*, vol. 43, Summer, 1995, 477-481.



terms and concepts to be matched across legal systems, adapting concepts of different systems and helping contextualization so to approach the most likely similar concept in the target language and system. In a more restrictive approach, only broad correspondences will be established, focused on broad concepts which are likely to be commonly understood by a variety of users.

### **3. Putting it into practice: a feasibility study for accessing multilingual legal literature**

The issues related to multilingual access to legal information and to methods for cross-language retrieval in the law domain are at the basis of the development of a feasibility study undertaken by the Institute of Theory and Techniques of Legal Information (ITTIG) for the implementation of a multilingual portal to legal literature<sup>43</sup>. The portal's aim is the provision of a single access point to distributed legal doctrine resources through the exploitation of rich metadata and the development of tools for the discovery, selection and use of relevant legal material. Connected to this, methodologies and techniques have been studied to address the specific question of cross-language retrieval of legal information resources. In this context consideration has been given to a number of issues like criteria for term equivalence to be established for matching queries to documents, methods and techniques for legal translation.

Following a survey on legal users' requirements and attitudes<sup>44</sup> in accessing legal information, efforts have been made to design the portal's system as well as the tools for generating and capturing metadata of structured and semi-structured web documents. In fact, the retrieval of legal doctrine requires high quality indexing, as well as appropriate searching methods and tools in order to ensure effective access of such documents by diverse legal user communities.

The feasibility study's focus is on two distinct requirements. These consist of: a) opening up the system to a wide user community, including foreign patrons, who must be offered the possibility to access legal material in their native language; b) providing multilingual access to foreign legal

---

<sup>43</sup> Francesconi, Enrico and Peruginelli, Ginevra. Access to Italian Legal Literature: Integration Between Structured Repositories and Web Documents. In *DC-03: Proceedings of the International Conference on Dublin Core and Metadata for e-Communities*, September 28 - October 2, 2003, Seattle, Washington, 99–107.

<sup>44</sup> Peruginelli, Ginevra. Understanding Information-Seeking Behaviour and the Needs of Italian Legal Users in Accessing Legal Literature. In: "Informatica e diritto", vol. XXX, no. 1-2, 2004, 281-302.

resources. These objectives are based on the belief that the development of strategies and tools for information access regardless of geographic and language barriers is a key factor to truly global sharing of legal knowledge, so as to make it possible for legal research and legal profession to progress according to the requirements of a modern society.

A two-phase approach has been planned for the implementation of the portal's multilingual access functionality. Firstly, an analysis has been carried out on topics as cross-language retrieval methods and techniques, their application to legal material, management of multilingual metadata. Secondly, a practical approach to the retrieval of multilingual legal doctrine has been envisaged. Based on the features of a federated system, documents coming from structured repositories and from the web are qualified using the Dublin Core (DC) metadata set<sup>45</sup> in its XML version and harvested using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).<sup>46</sup> In order to ensure multilingual access, query translation and word disambiguation as well as recognition and processing of characters (their presentation, arrangement, and transfer) have been identified as suitable methods for cross-language retrieval.

It is the system-bound nature of legal terminology that comes into play, meaning by that a close link between a term expressed in one language, referred to a given concept, and the specific meaning of such concept in a particular legal system. As a consequence, when a concept does not exist in the target legal system, simply translating terms is misleading for users.

During the design phase of the portal it has been soon realized that firstly account had to be taken of the fact that the whole process of interaction between legal languages occurring in cross-language retrieval is one of seeking subsidiary solutions. All these matters are definitely not technical in

---

<sup>45</sup> Dublin Core metadata set is a scheme including, in its 'unqualified' version, 15 basic fields identifying the main elements of electronic information resources in disparate domains. It has been designed to allow exchange and interoperability across data repositories and application systems. It is widely used all over the world and mapping (conversion) procedures to and from this format to different metadata formats are commonplace in the information arena, being easily implementable. For more information: Dublin Core Metadata Element Set. <http://dublincore.org/documents/dces/>.

<sup>46</sup> Open Archives Initiative Protocol for Metadata Harvesting, <http://www.openarchives.org/OAI/openarchivesprotocol.html>; Francesconi, Enrico and Peruginelli, Ginevra. Access to Italian Legal Literature: Integration Between Structured Repositories and Web Documents, *op. cit.*

nature as the research for equivalence implies both a comparative study of the different legal systems and a good knowledge of technical legal terminology.

To face the difficulty of establishing equivalence of the legal concepts of the various legal systems, a compromise had to be adopted, in an effort to favour integration of diverse legal cultures, while respecting each national legal system. When legal dictionaries, multilingual thesauri and other comparative tools are not available or adequate, it was decided that pragmatic choices are to be made as, for example, identifying a common ground, namely common legal concepts and facts which, although not perfectly coinciding with those belonging to other systems, are conceptually close. In this case users could, once the retrieved material has been examined, perceive the differences and peculiarities which make these resources unique. It is worth noting that this method does not necessarily lead to noise or unsuccessful searches, but allows for a first-phase search in context, useful to give evidence of the existence or non-existence of a specific concept in other legal systems.

In the feasibility study the identification of categories of law (i.e. trade law, constitutional law, criminal law) can help in effectively retrieving legal information indexed according to these categories. Mapping between law categories is necessary to reach proper contextualisation of the query in the diverse legal systems. An example illustrates the need for such mapping. The concepts related to property rights, like land law, property questions on insolvency, intellectual property, etc. according to the UK law belong to the field of property law, whereas in the Italian legal system these legal facts are regulated respectively by agricultural law, private law and industrial law.

In order to make the query more focused, the potential users of the portal may choose a legal category of the legal system expressed in the language of the query.

At operative level the user is required to choose a legal system, so to implicitly identify a language for queries, and a legal category, in terms of the dc:subject element<sup>47</sup>, thus implicitly identifying the right translations of possible ambiguous words. This is part of the functionalities to be offered by the user interface of the system which has to cope with ambiguous words in query language leading to multiple translations in a target language, each corresponding to a legal category in the target legal system (i.e. the Italian

---

<sup>47</sup> DC.subject is one of the 15 elements of the Dublin Core metadata scheme. The others are: contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, title, type.

word *dolo* can be translated into English either as *fraud* or as *malice*, respectively belonging to private law and criminal law). As the right sense of an ambiguous word in query language can be obtained only by word contextualization, providing the context by specifying the legal category, methods can be devised for mapping a legal category in the query legal system to the correspondent legal category in the target legal system.

From a technical point of view not every field has to be translated. In fact, Dublin Core metadata can be divided into query language-dependent and query language-independent metadata. For example the dc:title element is query language-independent since the title of a document has to be queried in its native language, independently from the query-language. Therefore, only the content of query language-dependent metadata has to be translated. While in a multilanguage domain-general environment the dc:subject element is usually query language-independent<sup>48</sup>, within a multi-language legal domain this is not true. For this reason the dc:subject element has to be translated by mapping its values from a legal system to different target ones. Also the content of the dc:description element (with its qualifiers, such as “Abstract”) is query language-dependent. It is a widely used access point and the information contained is often expressed using a semi-technical language; therefore in the portal functions the dc:description element has been held as being as important to translate as dc:subject. The content of the dc:subject and dc:description elements, submitted in a native language, are translated in a “pivot” language, that is English. Then, from the “pivot” language, the query is translated again to the other languages of the portal. At the end of the process, the right translations of ambiguous words can be obtained, and as many different queries as target languages used by the portal can be dispatched to the different language indexes.

The core of the system is definitely the query translation module. The approach is essentially based on line dictionaries although their limitations are known concerning the provision of alternative translations, together with the low quality of the legal dictionaries themselves which is a serious constraint. Specific choices are proposed within the portal to face these problems and some examples are given of methods for use.

---

<sup>48</sup> Lee, Wonsook, Sugimoto, Shigeo, Nagamori, Mitsuharu, Sakaguchi, Tetsuo, Tabata, Koichi. A Subject Gateway in Multiple Languages: A Prototype Development and Lessons Learned. In *DC-03: Proceedings of the International Conference on Dublin Core and Metadata for e- Communities*, September 28 - October 2, 2003, Seattle, Washington, 59-66.

As regards queries of just one word, the translation is selected which is more relevant to the domain (legally characterized translations). If the word does not belong to the law domain, the alternative translation is more likely to be selected. To multiwords queries the methodology “Cross Language Delegated Search”<sup>49</sup> is applied which has proved highly performant in the context of the international evaluation campaign CLEF 2006.<sup>50</sup> Such a methodology is based on semantic proximity of different translation set.

In this project an effort is made to analyse the set of issues concerning distributed access to legal multilingual information resources, for which unfortunately no ready-made solutions are available. Methods still need to be tested for adoption in a context where careful account is to be taken of the specificity of the material, due to its semantic content and diverse metadata used for its description, all this adding complexity in gaining unified access to multilingual legal information.

#### **4. Final remarks and recommendations**

On the grounds of research studies and information retrieval systems developed so far, it can be assumed that at the moment a specific model suitable for any type of application and domain does not exist and we are far away from an ideal approach to be adopted in view of overcoming the language barriers arising in information retrieval. A relevant number of issues are involved in cross-language retrieval. One important factor is the diverse availability of linguistic resources and tools like quality bilingual dictionaries, multilingual thesauri, parallel and comparable corpora, ontologies.

The amount and characteristics of these tools vary according to domain and languages involved. An additional relevant variable is the actual possibility to access sophisticated technologies for the development and management of information retrieval systems in combination with language processing methodologies. As a consequence, a large variety of techniques and methods have been adopted in the systems developed so far.

Despite the labour intensive and enthusiastic activity carried out by a large number of researchers, computer people and linguists all over the world,

---

<sup>49</sup> Dini, Luca and Curtoni, Paolo. CELI participation at CLEF 2006: Cross-Language Delegated Search. [http://www.clefcampaign.org/2006/working\\_notes/workingnotes2006-CurtoniCLEF2006.pdf](http://www.clefcampaign.org/2006/working_notes/workingnotes2006-CurtoniCLEF2006.pdf).

<sup>50</sup> CROSS-LANGUAGE EVALUATION FORUM – CLEF is an outstanding forum promoting research and development activities in multilingual information access. <http://www.clef-campaign.org/>.

and beyond the five-year action plan recommended in an important international meeting held in 2002<sup>51</sup> within the Association for Computing Machinery-Special Interest Group on Information Retrieval, reluctance still persists, especially by industry, to develop multilingual information retrieval systems in the form of consolidated services. The generic search engines available today on the Internet and the commercial firms already on the market with qualified information services in specific domains limit their functionalities to a multilingual search interface, while ensuring search results by simply automatically matching query terms and documents, either considered as full text or metadata describing them. Therefore, scientific research on multilingual access has not reached adequate maturity and the market size of such services is not sufficiently clear.

Furthermore, although the following opinions do not reflect the views of a large community on the subject, there is a debate showing some reservations about the need, following a user's query put in a given language, to retrieve documents in a language which is not familiar to him/her that he/she cannot master<sup>52</sup>. It is also pointed out that expert users, those who are able to make full use of information in a language other than their own, most probably will search such information in this latter language. However, it is a matter of fact that today users are multilingual and diverse: quite often a certain type of information exists only in poorly known languages or even ignored by users who nevertheless strongly need it; they have no other way to search it than using their own language and further ask for help to have it translated and interpreted.

In general, it can be stated that the critical aspects involved in the development of multilingual systems essentially concern three factors: 1) the relation between the domain dealt with in the application and the functionality which can be expected in searching for such information; 2) the relevance of lexical tools oriented to handling multiple languages; 3) the availability of financial resources needed for implementing complex systems, which often require a combination of techniques for multilingual retrieval.

---

<sup>51</sup> XXV International Annual Conference ACM/SIGIR (Association for Computing Machinery, Special Interest Group on Information Retrieval): Research and Development in Information Retrieval, Tampere, August 11-15, 2002. Gey, Fredric, Kando, Noriko, Peters, Carol. Cross-language information retrieval: a research roadmap. ACM SIGIR Forum, vol. 36, no 2, 2002, 72-80.

<sup>52</sup> Moulinier, Isabelle and Schilder, Frank. What is the Future of Multi-Lingual Information Access? In Working Notes of the New Directions in Multilingual Information Access Workshop at SIGIR 2006, Seattle, August 6-11, 2006.

As regards the first factor, it is worth pointing out that there are many experiments, assessed in official settings as CLEF<sup>53</sup> and TREC<sup>54</sup>, which deal with general collections material, as for example news in journals. In this area combined techniques are adopted by using lexicons and dictionaries which do not give rise to serious problems in creating correspondences among terms, due to the information processed, which is non-specialist in nature. Furthermore, in this case users do not require an extremely high level of precision, but it is enough for them to be provided with a functionality allowing them to understand search results produced according to a decent, not necessarily perfect, translation. In other thematic knowledge areas, as for example medicine and law, for which some multilingual retrieval applications have been developed, precision in search and retrieval are of main importance. As regards medical terminology, differences in concepts' definition are not the majority of cases as concepts are likely to be similar across languages. On the contrary, in the law domain, it is common knowledge that concepts are system-bound, depending on the system they belong to, and risks can occur in translation and transposition.

The second factor, concerning availability of linguistic resources and lexical tools of various types, is an essential condition for comparing a query put by the user and the information contained in documents. In multilingual information systems dealing with law usually techniques are used that allow translating query terms, and the terminology useful to match a query against documents is extracted from textual corpora. The collection to be searched is rarely translated for the purpose of its consultation, as legal documents are likely to be better interpreted in their original language.

Shortage of lexical sources which could help to establish correspondences among languages is still a reality for a great amount of languages considered as minor, but spoken by millions of people: this is definitely an obstacle to multilingual information access. Although the Web is an eclectic communication tool not only for the variety of content and

---

<sup>53</sup> The purpose of the Cross-Language Evaluation Forum is to support global digital library applications by developing an infrastructure for the testing, tuning and evaluation of information retrieval systems operating on European languages in both monolingual and cross-language contexts.

<sup>54</sup> TEXT RETRIEVAL CONFERENCE – TREC. <http://trec.nist.gov/>. The objective of the Text RETrieval Conference (TREC), co-sponsored by the National Institute of Standards and Technology (NIST) and U.S. Department of Defence is to support research within the information retrieval community by providing the infrastructure necessary for large-scale evaluation of text retrieval methodologies.

linguistic styles, but also for the many languages represented, it is still not adequately exploited for the production and processing of tools allowing to establish relationships among different languages. The dominance of English is a fact, and this phenomenon is partly the reason of the scarce interest, or even the impossibility at practical and economic level, to develop information retrieval systems in less-represented languages.

As regards the domain of law, the poor quality of law dictionaries, the amount of work needed for exploiting parallel corpora and the lack of availability of effective machine translation commercial system for legal information are all factors which severely hinder the development of effective cross-language information retrieval systems to law material.

The third factor having a major impact on multilingual systems' functionality is of an economic nature. It is a fact that major search engines have not ventured in true cross-language information retrieval services also due to the high cost of development of these systems. On the grounds of applications developed so far, there is evidence that a combination of different techniques is needed. These consist, for example, in query expansion and relevance methodologies, to be operated also through users' interaction. To be effective, these techniques have to be adapted and properly fitted to the various applications and, if implemented in large scale systems, they raise development costs considerably.

Most probably the future of multilingual systems is conditioned not only to plurilingual resources like textual and parallel corpora and thesauri, but also to advances and progress in the processing and exploitation of monolingual tools, which should be fairly structured and analytical to ensure concordances and functional relations with other languages' terminology. As a consequence, the contribution of the science of linguistics is extremely important for building effective cross-language retrieval systems and this is particularly true in the law domain.

On the basis of these considerations, a multilingual model in the law domain requires an orchestration process in which all responsible actors are involved, those operating in the wide environment of the diverse legal orders: legislators, judges, legal professionals, scholars, linguists and also citizens. The challenge is not to choose a given communication language, rather to find a way to make linguistic and cultural diversities coexist in harmony.

In this context multilingual legal information retrieval systems do represent the necessary tools to encourage multilingualism in the law domain



and have the chance to make it effective. By examining a number of linguistic tools, projects and information systems, there is evidence that various approaches are developed for their implementation; further major factors are that comparative aspects are not always addressed in real terms, and the focus is mainly on purely linguistic aspects. It is a matter of fact that there is no true debate on the specific topic of legal cross-language information retrieval, while research studies and discussions among linguists and jurists on legal translation are frequent and abundant.

Certainly many aspects of this specific type of translation, which are crucial to transposition of texts from one language to another, are also relevant to the purpose of multilingual retrieval systems, whose functionality is based upon the possibility to match users' queries and documents. However, the requirements for text translation are not identical to those related to the processing of query translation in a retrieval system, so that different criteria have to be adopted to establish a correspondence among concepts.

Leaving untouched that precision in legal language is an essential feature due to the inevitable negative implications of inexactness and inaccuracy in every action where law is involved, it can be stated that in multilingual retrieval systems, where comparison of legal concepts comes into play, solutions can be acceptable that are likely to be less rigid as compared to legal text translation, whose incorrect formulation can have extremely serious implications.

Less rigour does not imply abandoning the principle of clarity and precision which is necessary in comparing legal concepts of different legal orders, rather it means adopting a flexible approach. From legal information retrieval systems it is expected to ensure knowledge communication by providing really relevant documents, even coming to a compromise. This means, for example, in terms of comparison and therefore in translation, accepting a more general concept as compared to the more specific original one, and also systematically adopting term disambiguation techniques, so as to present more results and interact with users. In fact users can be asked to provide feedback by checking translation alternatives presented following a search in a given language, as well as to select the context. However, in real life research studies and investigations on users' attitudes to interact with the application in view of improving search results are scarce, and user-system interaction is still quite a complex issue which needs further exploration.

The development of multilingual legal information systems offers the opportunity to address the questions concerning the relationship between

language and law and to look at legal translation from a particular point of view, oriented to the provision of qualified information on pertinent legal material all over the world. In this direction translation strategies are to be adopted allowing such essential objective to be achieved, while excluding techniques like creation of neologisms, calco and linguistic loans which are of no value for cross-language retrieval. Differently from these, the functional equivalence approach, to be intended as accomplishing the same general function as that of the source concept, appears to be the most effective method, following a query expressed in any language, to compare concepts and present users with relevant documents in different languages. In particular, in the search for correspondence, a theme which is highly debated in translation theory and practice is the applicability and validity of the principle of functional equivalence, which in the legal domain is called legal equivalence<sup>55</sup>, to mean the consideration of equal legal effects that a translated text will have in a target culture expressed in a given language. This represents a real challenge to the traditional "to the letter" approach to legal translation which was valid until the beginning of the last century when multi-ethnic countries as for example Canada, India, Switzerland and Belgium had to face the problem of multilingual legislation<sup>56</sup>.

In view of preparing to develop systems and tools allowing users to retrieve and make use of legal information resources made available through institutional and commercial services, it is necessary to start a joint activity among actors with quite different skills: jurists, linguists, jurilinguists, translators and researchers in the field of new technologies, as well as carry out large scale comparative studies on several fields of law. Through linguistic-conceptual correspondence definition activities carried out in cooperation among institutional organisations committed to indexing and delivering of legal material world-wide, it will be possible to set up experimental applications, pilot projects and systems capable of fostering awareness and understanding of different countries' legal concepts. In particular the following actions are recommended to enhance awareness of the essential components of multilingual legal information retrieval and to develop tools supporting the related systems:

1. investigation on the state of the art of linguistic-conceptual tools in the law domain, available at national and international level;

---

<sup>55</sup> G mar, Jean-Claude. Traduire ou l'art d'interpr ter, Tome I et II. Sainte-Foy : Presse de l'Universit  de Qu bec, 1996.

<sup>56</sup> The question mainly addressed by legal theorists and comparatists concerns whether the concept of legal equivalence is applicable to all types of texts, genres and sub-genres in legal translation.

2. analysis of techniques and concept retrieval (rather than string retrieval) oriented methodologies based on conceptual information retrieval models;
3. systematic development of ontologies as concept definition systems allowing legal entities to be represented in various legal systems and languages. A large scale development activity (although quite expensive) would facilitate the collection of a critical mass of data to create possible correspondences, both lexical and semantic, among concepts of the various legal orders;
4. systematic identification of large scale parallel and comparable legal text corpora useful to create lexicons and thesauri;
5. surveys in different service settings, both institutional and commercial, on legal users, evaluating their requirements in multilingual legal research, attitudes as well as critical issues, while assessing their availability to interact with retrieval systems in view of improving search results.

It is my conviction that only a strong collaboration among different countries' institutions having similar skills and responsibilities can contribute to the setting up of multilingual access services capable to foster shared awareness and understanding across countries, systems and languages for the benefit of world citizens.

## References

Avalos, Francisco. Legal translations: some tips. Lecture delivered at National Language Resource Center, San Diego State University on July 24, 1998.

Capellas-Espuny, Gemma. The problem of terminological equivalence in international maritime law: the case of hypothèque and mortgage in the document Final Act and International Convention on maritime liens and mortgages. *Journal of diplomatic language*, vol.1, no 4, 2004.

Dini, Luca and Curtoni, Paolo. CELI participation at CLEF 2006: Cross-language delegated search. [http://www.clefcampaign.org/2006/working\\_notes-/workingnotes2006/CurtoniCLEF2006.pdf](http://www.clefcampaign.org/2006/working_notes-/workingnotes2006/CurtoniCLEF2006.pdf).

Engerb, Jan. Statutory texts as instances of language(s): consequences and limitations of interpretation, 2004. [http://www.brook-law.edu/students/journals/bjil/bjil29iii\\_engberg.pdf](http://www.brook-law.edu/students/journals/bjil/bjil29iii_engberg.pdf).

Fairclough, Norman. *Critical discourse analysis: the critical study of language*. London ; New York : Longmann, 1995. ISBN: 0582219809;  
Fairclough, Norman. *Discourse and social change*. Cambridge, UK ; Cambridge, MA : Polity Press, 1992. ISBN: 0745606741.

Fletcher, George P. The language of law: common and civil. In Pozzo, Barbara (ed.). *Ordinary language and legal language*. Milano : Giuffrè, 2005, 83-107. ISBN 8814118048.

Francesconi, Enrico and Peruginelli, Ginevra. Access to Italian legal literature: integration between structured repositories and web documents. In *DC-03: Proceedings of the International Conference on Dublin Core and Metadata for e-Communities*, September 28 - October 2, 2003, Seattle, Washington, 99–107.

Gallo, Giovanni. Les jurists linguists de la Cour de Justice des Communautés européennes. In Sacco, Rodolfo and Castellani, Luca (eds). *Les multiples langues du droit européen uniforme*. Torino : L'Harmattan Italia, 1999, 71-89. ISBN 8887605076.

Gémar, Jean Claude (ed). *Langage du droit et traduction : essais de jurilinguistique. The Language of the Law and Translation : Essays on Jurilinguistics*. Québec : Linguatex-Conseil de la langue française, 1982. ISBN 2920342134.

Gémar, Jean-Claude. La traduction juridique et son enseignement : aspects théoriques et pratiques. In *La traduction juridique*. Numéro spécial, *Meta*, vol. 24, no 1, 1979, 35-53.

Gémar, Jean-Claude. *Traduire ou l'art d'interpréter*, Tome I et II. Sainte-Foy: Presse de l'Université de Québec, 1996.

Gey, Fredric, KANDO, Noriko, PETERS, Carol. Cross-language information retrieval: a research roadmap. *ACM SIGIR Forum*, vol. 36, no 2, 2002, 72-80.

Gizbert-Studnicki, Tomasz. Is an empirical theory of the language of the law possible?. In Ziembinski, Zygmunt (ed.). *Polish contributions to the theory and philosophy of law*. Amsterdam : Bodopi, 1987, 99-114.

Groot, Gerard René de and Laer, Conrad J.P. van. The dubious quality of legal dictionaries. *International journal of legal information*, vol. 34, no 1 (Spring 2006), 65-86.

Groot, Gerard René de. Language and law. In *Netherlands report to the fifteenth International Congress of Comparative Law*. Antwerp/ Groningen: Intersentia, 1998, 21-32.

Harvey, Malcolm: What so special about legal translation? *Meta*, XLVII, 2, 2002.

Heutger, Viola. A more coherent European wide legal language. *European Integration online Papers (EioP)*, Vol. 7, no 2, 2004. <http://eiop.or.at/eiop/pdf/2004-002.pdf>.

Hoecke, Mark Van. Deep level comparative law. In VAN Hoecke, Mark (ed.). *Epistemology and methodology of comparative law*. Oxford: Hart, 2004. ISBN 1841134430.

Kjaer Anne Lise. A common legal language in Europe? In Van Hoecke, Mark (ed.). *Epistemology and Methodology of Comparative Law*. Oxford: Hart, 2004. ISBN 1841134430.

Kjaer, Anne Lise. Convergence of European legal systems : the role of languages. *Language and culture*, no 29, 2004, 125-137.

Lavoie Judith. Droit et traductologie: convergence et divergence. In J.C. Gémar and N. Kasirer (eds). *Jurilinguistique entre langues et droits*. Bruxelles: Bruylant, 2005. ISBN 2894001967.

Lee, Wonsook, Sugimoto, Shigeo, Nagamori, Mitsuharu, Sakaguchi, Tetsuo, Tabata, Koichi. A Subject Gateway in Multiple Languages: a prototype Development and Lessons Learned. In *DC-03: Proceedings of the International Conference on Dublin Core and Metadata for e- Communities*, September 28 - October 2, 2003, Seattle, Washington, 59-66.

Luhmann, Niklas. *Das Recht der Gesellschaft*. Frankfurt/Main: Suhrkamp, 1993. ISBN 3518581686.

Matthijssen, Luuk. A task-based interface to legal databases. *Artificial intelligence and law*, vol. 6(1), 1998, 81-103.

Matthijssen, Luuk. *Interfacing between lawyers and computers: an architecture for knowledge-based interfaces to legal databases*. The Hague; Boston: Kluwer Law International, 1999.

Mellinkoff, David. *The Language of the Law*. Boston ; Toronto : Little, Brown, 1963; Sacco, Rodolfo. *Lingue et droit*. In *Rapports nationaux italiens au XVème Congrès International de Droit Comparé*. Bristol, 1998, Milano 1998.

Moens, Marie-Francine. *Improving access to legal information*. In Oskamp, Arno and Lodder, Anja (eds). *Information technology and lawyers*. Dordrecht: Springer, 2006, 119-136. ISBN 1402041454.

Moréteau, Olivier. *L'anglais pourrait-il devenir la langue juridique commune en Europe?* In SACCO, Rodolfo and CASTELLANI, Luca (eds). *Les multiples langues du droit européen uniforme*. Torino: L'Harmattan Italia-Isaidat, 1999, 143-162. ISBN 8887605076.

Morris, Marshall (ed.). *Translation and the law*. Amsterdam; Philadelphia: Benjamins, 1995. ISBN 9027231834.

Moulinier, Isabelle and Schilder, Frank. *What is the future of multilingual information access?* In Working Notes of the New Directions in Multilingual Information Access Workshop at SIGIR 2006, Seattle, August 6-11, 2006.

Pelage, Jacques. *La traductologie face au droit*. In *Proceedings of the International Colloquium University of Geneva, February 17-19, 2000*. Berne: Association Suisse des Traducteurs, Terminologues et Interprètes (ASTTI), Ecole de traduction et d'interprétation de Genève (ETI), 2000, 125-131.

Peruginelli, Ginevra. *Understanding information-seeking behavior and the needs of Italian legal users in accessing legal literature*. In: "Informatica e diritto", vol. XXX, no. 1-2, 2004, 281-302.

Pozzo, Barbara. *Harmonisation of European contract law and the need of creating a common terminology*. *European review of private law*, Vol. 11, no 6, 2003, 754-767.

Rotman, Edgardo. The inherent problems of legal translation: theoretical aspects. *Indiana International and Comparative Law Review*, vol. 6, no 1, 1995, 187-196.

Saadoun, Adel, Ermine, Jean-Louis, Belair, Claude, Pouyot, Jean Mark. A knowledge engineering framework for intelligent retrieval of legal case studies. *Artificial Intelligence and Law*, vol. 5, no 3, 1997, 179-205.

Sacco, Rodolfo. Langue et droit. In *Rapports nationaux italiens au XVème Congres International de Droit Comparé*. Milano : Giuffrè, 1998.

Sacco, Rodolfo. Language and law. In Pozzo, Barbara (ed.). *Ordinary language and legal language*. Milano : Giuffrè, 2005, 1-21. ISBN 8814118048.

Sacco, Rodolfo. Legal formants : a dynamic approach to comparative law. *The American Comparative Law*, vol. 1, 1991, 343-358.

Sacco, Rodolfo. Riflessioni di un giurista sulla lingua (lingua del diritto uniforme e il diritto al servizio di una lingua uniforme). *Rivista di diritto civile*, vol. 42, no 1, 1996, 57-65.

Šarčević, Susan. *New approach to legal translation*. The Hague ; Boston : Kluwer law International, 1997. ISBN 9041104011.

Schlesinger, Rudolf B. The past and future of comparative law. *American Journal of Comparative Law*, vol. 43, Summer, 1995, 477-481.

Vanderlinden, Jacques. Le future des langues du droit ou le dilemma du dernier orateur. In Sacco, Rodolfo and Castellani, Luca (eds). *Les multiples langues du droit européen uniforme*. Torino: L'Harmattan Italia, 1999, 193-222. ISBN 8887605076.