The European Legal Forum

Forum iuris communis Europae

Peruginelli, Ginevra

Towards a common understanding of law in a multilanguage world: The role of cross-language legal information retrieval systems

The European Legal Forum (E) 1,2-2007, 64 - 71

direct link must be established between the tax advantage concerned and the offsetting of that advantage by a particular tax levy.[12]

27. However, in the present case, not only does the Kingdom of Sweden not advance any specific arguments before the Court seeking to establish that the provisions of the IL in dispute are justified by the need to ensure coherence of the tax system concerning the deferral of taxation on capital gains, but, moreover, in its defence, it admits the failure to fulfil obligations of which it is accused. Further, it submits that, in order to fulfil its obligations under Community law, it is studying a reform of the complex rules on deferral of taxation.

28. According to the case-law, whether a Member State has failed to fulfil its obligations must be determined by reference to the situation prevailing in the Member State at the end of the period laid down in the reasoned opinion, and the Court cannot take account of any subsequent changes.[13]

29. In those circumstances, the provisions of Chapter 47 of the IL in dispute must be held to be contrary to Articles 39 EC and 43 EC.

30. Secondly, with regard to persons who are not economically active, the same conclusion applies, for the same reasons, to the complaint relating to Article 18 EC.

31. The Commission also claims that the Kingdom of Sweden has failed to fulfil its obligations under Articles 28 and 31 of the EEA Agreement, relating respectively to freedom of movement for workers and freedom of establishment.

32. It is to be noted, in the present case, that the rules prohibiting restrictions on the freedom of movement and the freedom of establishment laid down in Articles 28 and 31 of the EEA Agreement are essentially identical to those established by Articles 39 EC and 43 EC.

33. Therefore, in the light of the conclusion reached in para. 29 of the present judgment, the provisions of Chapter 47 of the IL in dispute are also contrary to Articles 28 and 31 of the EEA Agreement.

34. The Commission's action must accordingly be considered to be well founded as far as the complaint alleging infringement of the rules on freedom of movement for persons in the EEA Agreement is concerned.

35. It must therefore be held that, by adopting and maintaining in force tax rules, such as those in Chapter 47 of the IL, which make entitlement to deferral of taxation on capital gains arising from the sale of a private residential property or of a right to reside in a private cooperative building conditional on the newly-acquired residence also being on Swedish territory, the Kingdom of Sweden has failed to fulfil its obligations under Articles 18 EC, 39 EC and 43 EC and under Articles 28 and 31 of the EEA Agreement.

Free movement of capital

36. In addition, the Commission seeks a declaration from the Court that the Kingdom of Sweden has failed to fulfil its obligations under Article 56(1) EC and Article 40 of the EEA Agreement.

37. Since the provisions of the Treaty and the EEA Agreement on freedom of movement for persons preclude the contested legislation, there is no need for a separate examination of that legislation in the light of Article 56(1) EC and Article 40 of the EEA Agreement concerning the free movement of capital.[14] (...)"

[11] See, *De Lasteyrie du Saillant* (*supra* note 10), para. 49, and ECJ 7 September 2006 – C-470/04 – *N* [2006] ECR I- 07409, para. 40.

[12] See, to that effect, *Keller Holding* (*supra* note 10), para. 40 and the case-law cited.

[13] See, *inter alia*, ECJ 16 October 2003 – C-388/02 – *Commission v Ireland* [2003] ECR I-12173, para. 6.

[14] See *Commission v Portugal* (*supra* note 6), para. 45.

INFORMATION ON CURRENT ISSUES

Towards a common understanding of law in a multilanguage world: The role of cross-language legal information retrieval systems

*Ginevra Peruginelli**

1. Multilingualism in the law domain: a brief overview

Multilingualism is a phenomenon reflecting the plurality of languages used by the vast amount of communities worldwide. Referred to legal information, it is to be intended both as a de-facto situation characterized by the existence of different legal languages, and as the set of issues involved in the management of legal information across language barriers.

Internationalisation and increasing globalization of market economy 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 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.

There is no doubt that the exchange of information is largely dependent on languages, to be intended not only as a system

* *Ginevra Peruginelli, Lawyer; Assistant Researcher at Istituto di Teoria e Tecniche dell'Informazione Giuridica del Consiglio Nazionale delle Ricerche (ITTIG-CNR), Florence, Italy.*

of symbols, but also as a mean of communication[1] and as a tool for mediating between different cultures.[2]

If we consider the language of the law, such language's 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, jurisprudence and doctrine as main components of the various legal cultures.[3]

Multilingualism in the law domain is mostly unanimously perceived as a very complex issue, linked as it is to disciplines such as comparative law, linguistics, translation theory and practice. It is a highly debated topic not only among professionals and scholars of these various disciplines,[4] but also among government officials in institutional settings at national and international level.

Europe is a typical example of multi-language and multi-system environment where decisions on linguistics' policy are now receiving considerable attention. In the European Union a full multilingualism is met by providing a huge translation work of legal documentations.[5] In addition some languages such as English, French and German have a special status since the majority of the material is to be handled in these three languages. Due to economic and practical reasons a serious linguistics' policy is likely to have to choose between two opposite extremes.[6] These are represented from one side by a multilingualism embracing all European languages and being as equalitarian as possible (a very expensive solution); on the other side by the adoption of a unique language, in particular a sort of international English which is already in place in the area of international trade and in many areas of the legal doctrine.

Two major aspects have to be taken into account in dealing with cross-language legal information retrieval: one concerns the intimate link between language and law and covers the crucial issues of rendering legal terms across languages; the other refers to comparative issues, to the relationship between legal systems which, while a problem in its own, is exacerbated in a multi-language environment.

Every attempt to exchange legal knowledge among different communities and to reach a common understanding of different legal systems has inevitably to cope with the problems posed by language and systems' diversity.

2. Access to legal information across languages

Access to legal information across languages is intended as the functionality allowing the retrieval, through the use of information systems, of meaningful legal material beyond the barriers of languages and legal orders' differences.

Cross-language information retrieval (CLIR) is defined as the capability for users to retrieve material written/expressed in a language different from a query language.[7] Cross-language information retrieval is not to be simply associated with the need to find documents for people who cannot read them. In fact multilingual searchers with a limited active vocabulary, but good reading comprehension in a second language, might prefer to issue single queries into a multilingual collection in their most fluent language. Information retrieval across languages is an application of information retrieval. These two fields share the same goals and a number of information retrieval techniques for matching documents as well as for processing queries are equally useful in a multilingual application. However, cross-language retrieval differs from information retrieval in a significant way as the standard monolingual information retrieval process involves no translation component.

The requirements of a cross-language information retrieval system can be clarified by stating that in a multilingual access environment information is searched, retrieved and presented effectively, without constraints due to the different languages and scripts used in the material to be searched and in the metadata (that is bibliographic and semantic indexes to the material to be found) allowing to search such material. This implies that in creating multilingual access services, the users' native language as well as the multiplicity and richness of world-wide languages are to be accommodated, and methods have to be developed to allow users to put queries expressed in any one language and retrieve information resources independently of the language of documents and indexing.

The main approach is usually based on translating the query at retrieval time into the document (or metadata) languages, although it would be possible to translate all of the documents into the query language. Although many experiments have been carried out using query translation techniques, in the real world they pose a number of problems related to the need of contextualisation and interpretation, which are increased when dealing with legal information.[8]

---

1 *Wittgenstein, Ludwig*, Philosophical investigations, Oxford: Blackwell, 1997.

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

3 *Sacco, Rodolfo*, Language and law, in: *Pozzo, Barbara* (ed.), Ordinary language and legal language, Milano: Giuffrè, 2005, pp. 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, pp. 83-107, ISBN 8814118048.

4 *De Groot, Gérard René*, Language and law, in: Netherlands report to the fifteenth International Congress of Comparative Law, Intersentia, Antwerp/ Groningen, 1998; *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, pp. 57-65.

5 *Gallo, Giovanni*, Les linguistes juristes de la Cour de justice de la Communauté européenne, in: *Sacco, Rodolfo*, *Castellani, Luca* (eds). Les multiples langues du droit européen uniforme, Torino: L'Harmattan Italia, 1999, pp. 71-89, ISBN 8887605076.

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

7 *Lee, Kyung-Soon*, *Kageura*, *Kyo* and *Choi, Key-Sun*, Implicit ambiguity resolution using incremental clustering in cross-language information retrieval, Information processing and management, vol. 40, no. 1, 2004, pp. 145-159.

8 *Francesconi Rancesconi, Enrico* and *Peruginelli, Ginevra*, Opening the legal literature portal to multilingual access, in DC-2004: Proceedings of the International Conference on Dublin Core and Metadata Applications, October 11-14, 2004, Shanghai, pp. 37-44, http://www.slais.ubc.ca/PEOPLE/faculty/tennis/dcpapers2004/Paper-_31.pdf; *Sheridan, Paraic*, *Braschler, Martin* and *Schauble, Peter*.

Multilingual access is based on the recognition that cultural diversity is vital to the maintenance of society and that languages are a strong element of the different cultural traditions. With the increasing moves towards an integrated world and the multicultural nature of modern society and its globalisation, which is facilitated by the development of digital information and telecommunications networks, the need for multilingual information access has become more and more pressing and the issues connected with cross-language retrieval have increased in importance.

While nowadays large scale digital collections contribute to the dismantling of the geographic barriers to information access, language barriers are still critical to the effectiveness of resource sharing and world-wide common access, and their emergence as a problem is to be connected with the growing number of information databases now available on networks.[9]

It is mainly in the late 1970s that careful attention has been paid to cross-language retrieval and related tools. Software solutions have been developed and research projects and studies are under way where this functionality is often associated with disciplines like artificial intelligence and computational linguistics,[10] machine translation[11] and language engineering, natural language processing.[12]

An important aspect of multilingual access concerns strategic and management issues. These refer to the need for general consensus and recommendations to achieve multilingual functionality. Emphasis is put on the need for a paradigm shift in the information professionals' community to overcome language barriers in information retrieval. These themes are not as popular in the literature as those concerning technical and functional aspects, but are specifically addressed by some authors[13] who point out that the problem of English language dominance, initially encountered in the development of the character encoding systems, affects developments in cross-language information retrieval and that attention should expand beyond technical aspects.

It is also claimed[14] that, despite the technological developments which have occurred in the 1990s, in general digital library research and development have until recently somehow neglected the issues of multilingual querying, presentation and retrieval. Although this has changed in the recent years and progress have been made, no optimal solution is available yet.

There are two sides to the problem of multilingual access, which are of a technical and linguistic nature. These two aspects are very rarely dealt with in parallel in the literature. Issues concerning the presentation of results following a search session, the standard encoding of characters and the development of multilingual thesauri for subject headings, generally appear in strictly library related literature, while topics such as matching queries and documents, query expansion techniques and ranking methods are discussed in wider information specialists' environments.

The technical multilingual challenge is very complex. In the actual digital library environment where primary electronic resources are to be accessed, this implies interoperability of character sets used to represent texts in different languages.[15]

Concern is expressed for the issues related to both multilinguality and multiscripts[16] as there are thousands of languages, and only about two dozen scripts are currently used; however users expect them to be properly handled, and librarians and information specialists must meet their customers' requirements by asking systems' suppliers to respond to such expectations.

Furthermore, in order to realise the "global village library", access, display and storage devices have to be developed and improved to accommodate multilingual and multiscript information, but installation of communication systems are also needed to allow the transmission of different characters. These fall within the competence of the technical industry, which must be carefully advised by the information professionals' community.

In addressing issues of cross-language retrieval in a digital library environment, multilingual recognition and representation are commonly considered as essential features and in particular it is claimed how a major requirement of multilingual applications is the support of character sets and standard encoding to represent the information to be manipulated. It is also specified how HTTP protocol for the transfer of Web documents and resources provide parameters for language identification and how multilinguality is also taken

Cross-language information retrieval in a multilingual legal domain. In *Peters, Carol* and *Tanos, Costantino* (eds), Proceedings of ECDL-97, 1st European Conference on Research and Advanced Technology for Digital Libraries, Pisa, 1997, pp. 253-268.

9 *Hudon, Michèle*, Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. Knowledge Organisation, vol. 24, no. 2, 1997, pp. 84-91; *Oard, Douglas W.*, Serving users in many languages: cross-language information retrieval for digital libraries, D-Lib Magazine, December 1997, http://www.dlib.org/dlib/december97/oard/12oard.html.

10 *Michos, Stephanos*, *Stamatatos*, *Efstathios* and *Fakotakis, Nikos*, Supporting multilinguality in library automation systems using A. I., 1999, http://slt.wcl.ee.upatras.gr/papers/michos3.pdf.

11 *Kay, Martin*, Machine Translation: the disappointing past and present, in: Survey of the State of the Art in Human Language Technology, 1996, http://cslu.cse.ogi.edu/HLTsurvey/ch8node4.html#SECTION82.

12 *Fluhr, Christian*, Multilingual Information Retrieval, in: Survey of the State of the Art in: Human Language Technology, 1996, http://cslu.cse.ogi.edu/HLTsurvey/ch8node7.html.

13 *Borgman, Christine L.*, Multi-media, multi-cultural, and multi-lingual digital libraries, or how do we exchange data in 400 languages? D-Lib magazine, June 1997,

http://www.dlib.org/dlib/june97/06borgman.html; *Nardi-Ford, Loredana*, VALA Conference Sponsor report : report on 1998 VALA Conference to COLUG on 29/7/1998 at the CIAC Italian Resource Centre in Forrest,

http://archive.alia.org.au/sigs/online/reports/vala.report.html.

14 *Peters, Carol* and *Picchi, Eugenio*, Across languages, across cultures: issues in multilinguality and digital libraries, D-Lib magazine, May 1997, http://www.dlib.org/dlib/may97/peters/05peters.html.

15 *Dowlin, Kenneth E.*, Issues in offering multilingual and multiscript library services: the key to the global village library, in Automated systems for access to multilingual and multiscript materials: proceedings of the second IFLA satellite meeting, Madrid, August 18-19, 1993, München: K. G. Saur, 1994, pp. 23-26, ISBN 3598217978.

16 Clews, John, Language automation worldwide: the development of character set standards, British Library R&D report 5962, Harrogate: SESAME Computer Projects, 1988, ISBN 1870095014.

into account in HTML documents where a tag for the LANG attribute specifies the language of each HTML element. However the richness of characters, which have now reached a standard encoding system allowing a unique code (The Unicode Consortium, 2003) for each character of most of the world's written languages (about 64000 characters), can be displayed only if local applications are equipped with appropriate fonts and if base facilities in place for multilingual applications running on the World Wide Web are used, such as, for example, mark up of bi-directional text, superscripts and subscripts.

As cross-language information retrieval addresses the growing need to access large volumes of data and complex objects across language boundaries. There is consensus on identifying multilingual functionality not only as a multilingual interface dialogue, but as the ability for users to search in their preferred language and retrieve documents written or indexed in other languages.[17]

The complexity of languages as such and of languages coming into contact with one another, that is translation issues, is intimately linked to multilingual access' functionality. Gaining correspondence between languages is a major problem both when controlled and uncontrolled vocabularies are used. Vocabulary mapping, that is establishing equivalencies between terms, is a complex task due to difficulties arising from the use of different linguistic expressions for the same concept, the different degree of specificity and the presence of polysemeous terms.

In cross-language retrieval, methods have to be developed which successfully match queries against information objects over languages. While in monolingual retrieval some kind of word matching and weighting of results is the traditional way to do this, with cross-language retrieval this matching is much more complex to achieve. This implies the use of mechanisms for translating queries or documents and for disambiguating sense, which is an already crucial issue in a monolingual environment, but is considerably increased when mapping over languages.

3. Key aspects of cross-language legal information retrieval

While much attention has been given over the past years to the study and development of methodologies for accessing multilingual information in general, limited research and systems have been produced in the specific area of legal information retrieval and of related cross-language indexing and searching tools. 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 of cross-language retrieval of legal information.

The difficult task to effectively access multilingual legal material through information retrieval systems is definitively to match legal terms across languages and find adequate correspondence beyond legal system diversity, in a word to accommodate diversity in language and culture.[18] This generally implies translating from the language of the query to that of the material to be found or vice versa.

Most projects and system implementations are confined to making legislative and jurisprudential information available at national level; however in the last few years there has been a wide production of digital legal repositories which have been made available on the Internet in a variety of languages, and the need for accessing such wealth of information by a wide variety of users all over the world has rapidly increased.

As written below, the following aspects seem worthy of analysis as being closely related to the development and effectiveness of legal information retrieval systems: a) the relationship between law and language; b) comparative research of legal systems in relation to language issues.

3.1. The relation between law and language

It has been emphasized that "the law is a profession of words".[19] Many of the problems about meanings that are of concern to language specialists turn out to be of interest for legal professionals as well, and to 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.

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 common. Both, comparative jurists operating in academic environments and legal professionals are more and more 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.

Like language, law has, in its origin, development and structure, a character which is intimately linked to the history and culture of each country. A number of linguists and historians of the law have compared law and languages from the point of view of their origins as well as of their patterns of evolution and have pointed out that both law and language have, at their essence, rules which are constitutive to a system.[20]

---

[17] *Altuna, Bélen*, Consideration and requirements in designing Spanish multilingual library catalogues, in: Automated systems for access to multilingual and multiscript materials: proceedings of the second IFLA satellite meeting, Madrid, August 18-19, 1993, München: K. G. Saur, 1994, pp. 29-44, ISBN 3598217978.

[18] *Sacco, Rodolfo*, Langue et droit, in Rapports nationaux italiens au XVéme Congres International de Droit Comparé, Bristol, 1998.

[19] *Mellinkoff, David*, The Language of the Law, Boston, Toronto : Little, Brown, 1963.

[20] *Savigny, Friedrich Karl von*, Von der Poesie im Recht, Zeitschrift für geschietliche Rechtswissenshaft, 1816, 25; *Sacco, Rodolfo*, Il diritto muto, Rivista di diritto civile, 39, 1993, p. 689-702.

Moreover, there is a functional relationship between language and law as knowledge of law requires language and jurists have greatly contributed to the elaboration of legal languages for better adapting them to the needs of law within the various countries and their related communities.

In this context the transfer of legal knowledge is entrusted to written or spoken language. This leads to state that the law cannot manifest itself without language.[21]

As language is essentially a communication tool, it is successful when synchronization between the parties occurs. But while language is a means of understanding, it is also a means of misunderstanding between people belonging to different social groups or cultures. This synchronization is more than ever fundamental in the language of law[22] where equivalence of meanings across legal systems are difficult to find and assess, even when the same language is used!

In fact, legal language consists of legal terms, phrases and stable conventions and as such 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 also the languages which are proper to each category of people, who although speaking the same language, are bound to their own use of legal terminology and related concepts.. 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.[23] As legal language is culture-bound and interrelated 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 in fact not forever fixed and unchangeable, as they change when legal experience changes.24 The change of legal concepts is brought about through legal argumentation.[25] This is evidenced in the multifaceted role of language in establishing, maintaining, and changing concepts, which makes cross-system interoperability even harder to achieve.[26]

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 conceptualised in the same way in all other language systems. This is especially true for legal terminologies at a system level.[27]

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,[28] so discourse is dependent on the social context in which it is used; in other words it is shaped by that context, which is not an immutable entity.

In this view, legal discourse consists of legal rules produced within the framework of a legal system, with the double function of sustaining and reproducing the system and of changing and transforming it. All this has implications on the possibilities of a 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 transcends national boundaries.

This brings some doctrine to state that different legal practices, diverse legal languages and cultures are exposed to each other and it is likely that, for example in Europe, the national legal traditions will gradually change along with the emerging intercultural communication of legal actors, who in that way adapt themselves to the changing institutional context of law.[29]

From another perspective, similar conclusions have been drawn[30] based on the assumption that reason, as the universal human faculty of intuition by which one sees reality, is the common language of people belonging to different speech communities, but the fact that our experience of the world is bound to language, does not imply an exclusiveness of perspectives. An example clarifies this concept: even if the French legal concept of "contrat" is radically different from the English concept of "contract", due to opposite approaches to the formation of a contract, an English lawyer can understand the French concept and vice versa.

However coming to an understanding across legal languages and legal systems is a hard process and the establishment of a common understanding can only be achieved if the

---

21 *Vanderlinden, Jacques* and *Snow, Gérard* (eds), Francais juridique et science du droit, Centre international de la Common law en français, Bruxelles: Bruylant, 1995, ISBN 280270964X.

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

23 *De Groot, Gérard René*, Language and law, in: Netherlands report to the fifteenth International Congress of Comparative Law, Intersentia, Antwerp/ Groningen, 1998.

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

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

26 *Kjaer, Anne Lise*, A common legal language in Europe? In: *Van Hoecke, Mark* (ed.), Epistemology and Methodology of Comparative Law, Oxford: Hart, 2004; *Gemar, Jean Claude* (ed), Langage du droit et traduction: essays de jurilinguistique, The Language of the Law and Translation: Essays on Jurilinguistics, Montréal: Linguatech-Conseil de la langue française, 1982.

27 *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.

28 *Fairclough, Norman*, Critical discourse analysis: the critical study of language, London, New York: Longmann, 1995; *Fairclough, Norman*, Discourse and social change, Cambridge: Polity Press, 1993.

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

30 *Gadamer, Hans-Georg*, Truth and method, New York: Seabury Press, 1975.

actors base their interpretation on the tacit suppositions about the world which forms the point of departure of the communicative action.

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 and comparative law.

3.2. 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[31] 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 to 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, 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 have for a long time operated 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. Their experience reveals 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.

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 as well as intellectual property 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 comparatists are strongly concerned about the implications of the differences existing between the cultural contexts underlying the various legal languages and about the difficulties in transferring legal meanings and legal concepts from one legal system to another, even when the same language is used. A number of frequently mentioned examples are made to refer to this phenomenon, such as societé in French legal language in France, which has not the same meaning as societé in French legal language in Belgium.[32] 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.[33]

In recent years research studies have increasingly concentrated 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.[34]

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.[35]

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

In continental Europe for a number of centuries a *ius 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".[36] With the age of codification, two facts contributed to the creation of intellectual barriers between the legal systems of the several na-

---

[31] *Hoecke, Mark van*, Deep level comparative law, in: *Van Hoecke, Mark* (ed.), Epistemology and methodology of comparative law, Oxford: Hart, 2004; *Schlesinger, Rudolf B*, The past and future of comparative law, American Journal of Comparative Law, vol. 43, Summer, 1995 pp. 477-481; *Sacco, Rodolfo*, Legal formants: a dynamic approach to comparative law, The American Comparative Law, vol.1, 1991, pp. 343-358.

[32] *Vanderlinden, Jacques*, Le futur des langues du droit ou le dilemme du dernier orateur, in *Sacco, Rodolfo*, *Castellani, Luca* (eds), Les multiples langues du droit européen uniforme, Torino: L'Harmattan Italia, 1999, pp. 193-222, ISBN 8887605076.

[33] *Heutger, Viola*, A more coherent European wide legal language, European Integration online Papers (EioP), vol. 7, no 2, 2004, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=520742.

[34] *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, p. 754-767.

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

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

tions: 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 searched 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*[37], 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.

Reciprocal influences, even though not necessarily direct legal transplants, are likely to happen and the point is made that in the first decades of the twentieth century the differences between common law and civil law traditions were emphasised, whereas today, at least in a number of fields of law, the common elements are sought rather than differences stressed.[38]

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 terms and concepts to be matched across legal systems, adapting concepts of different systems and helping contextualisation 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.

4. Conclusions

Multilingualism is a phenomenon which greatly affects the efficiency of exchange of information among people in every sector of life. Cross-language retrieval systems are tools which contribute to such exchange, supporting research and world-wide communication, and helping data to be searched irrespective of the language in which information objects are expressed.

However, systems developed so far mainly concentrate on pairs rather than on multiple languages, the latter being a much more complex endeavour. In the recent years a number of experiments and discussions focused on the developments of vocabulary matching over languages as such, but cross-language retrieval is still in the experimental stage and we cannot talk about truly multilingual retrieval systems. Any single method presents some limitations and there exists no one-fits-all solution for ensuring that users effectively access information beyond language barriers.

In particular, in the law domain, interoperability is paramount. Here a vast category of users need to exchange legal information world-wide and carry out activities in a context where it is possible to reach a common understanding of law beyond language barriers. However, this requirement is hard to meet, due to the variety of languages and modes in which the legal discourse is expressed and due to the diversity of legal orders and the legal concepts on which these systems are founded.

It is a fact that research is still under way to build a strong basis for monolingual information access, and cross-language retrieval will definitely be based on advances in this area. In the specific field of legal domain, adequate techniques for disambiguation of terms, translation strategies taking into account pragmatic and cultural interpretations and, mostly important, multilingual ontologies, appear to be a very promising approach. Relevant activities are carried out in this direction for the development of legal ontologies. One example is a multilingual legal thesaurus called "Lexical Ontologies for Legal Information Sharing – LOIS project",[39] which aim is to develop a multilingual access facility for European legal databases. This will enable citizens and professional users to search for European legislation and other legal documents (such as court cases) across, at the moment, six European languages (Italian, English, German, Czech, Portuguese and Dutch). To achieve this goal the project will use formal representations of legal concepts in each language using the WordNet technique.[40] Similar concepts in different languages (synsets) will be cross linked in such a way that users can enter queries to a legal documentation base in his/her language and retrieve also documents written in different languages.

All this implies a labour-intensive work as the model of legal multilingualism is not simply to be confined to the transposition of legal concepts from one system to another, but it is likely to require a cooperative venture, an orchestration process involving all stakeholders responsible for the various legal systems as legislators, judges, legal professionals and eventually citizens in an effort towards a common understanding of law beyond language and system barriers.

[37] *Sacco, Rodolfo*, Legal Formants, in: American Journal of Comparative Law, 1991, pp. 1 ss. (I); pp. 344 ss. (II).

[38] *Orucu, Esin*, Critical comparative law: considering paradoxes for legal systems in transition, 1999, http://www.ejcl.org/41/art41-1.doc.

[39] LOIS: http://www.loisproject.org/.

[40] The technical solution that will be adopted in the LOIS project is the Princeton WordNet initiative. WordNet is a multi-lingual index that specifies lexical semantic concepts and relations across languages. As such, it can be seen as a formally specified ontology that can be used for various applications in automated text processing (e.g. correction, explanation and translation). The importance of this technique for the multilingual European context has already been recognized by the EU, which has funded a generic implementation project: EuroWordNet, http://www.illc.uva.nl/EuroWordNet/.