

10th International "Law via the Internet" Conference

Durban, South Africa, 26 - 27 November 2009

A multilingual approach for promoting worldwide open access to law

Enrico Francesconi
Ginevra Peruginelli
Mario Ragona

Institute of Legal Information Theory and Techniques of the National Research Council (ITTIG-CNR), Florence, Italy

1. A glance to cross-language legal information retrieval

Internationalization 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 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. Nowadays improving knowledge of the law is imperative for the benefit of a wide category of users: legal professionals, businessmen, scholars and laymen, interacting as they are in an environment of increasing cross-border relations where powerful information and communication technologies are available, greatly facilitating access to legal material despite its specificity and complexity. However, language barriers still hinders the exchange of legal information and universal access to law. There is no doubt that the exchange of information is largely dependent on language, to be intended not only as a system of symbols, but also as a means of communication and thus as a tool for mediating between different cultures. If we consider the language of the law, we notice that the properties of such language have a major impact on the exchange of legal information.

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, due to the system-bound nature of legal terminology. In fact, each legal order is situated within a complex social and political framework originating from the history, traditions and habits of a particular community.

Multilingualism in the law domain is almost unanimously perceived as a very complex issue. It is a highly debated topic not only among professionals and scholars of comparative law, linguistics, translation theory and practice (de Groot 1998), but also among government officials in institutional settings at national and international level, as 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.

Europe as whole in particular is a typical example of multi-language and multi-system environment where decisions on linguistic policy are now receiving considerable attention (Gallo 1999). The European Union subscribes to full multilingualism in its 23 official languages, causing a huge amount of translation work for legal documentation. Linguistic policy is faced with the economic and practical problems of handling so many languages. Hence, there is a certain pressure to simplify, at least for certain contexts and specific types of documentation. Two opposite

positions are held concerning the management of multilingualism (Moréteau 1999). These are represented by a multilingualism embracing all European languages and being as egalitarian as possible on the one extreme, and on the other by the adoption of one single language, in particular a sort of international English which is already used in some fields of law and specific legal areas such as international trade as well as in the scholarly and professional literature.

While much attention has been given over the past years to the study and development of methodologies for accessing multilingual information in general, research has been limited to the specific areas of legal information retrieval and of related cross-language indexing and search tools. As concerns the law domain, it is worth noting that major attention has been paid by scholars and legal experts to linguistic and conceptual aspects of the legal languages: these themes are undoubtedly relevant to multilingual access and can provide important insights into this topic.

The difficulty in effectively accessing multilingual legal material through information retrieval systems definitively lies in matching and weighing of legal terms across languages. This generally requires translating from the language of the query to that of the material to be found or vice versa and addressing the issue of word disambiguation, which is greatly increased when mapping over legal languages. In fact, crossing the language barrier between search requests and documents calls for finding solutions to the problems caused by the system-bound nature of legal terminology, and devising methods to map concepts across different legal systems.

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 made available over the Internet in a variety of languages. Hence, the need for accessing such a wealth of information by a wide variety of users all over the world has rapidly increased.

2. Key issues for cross-language legal information retrieval systems

In developing cross-language legal information retrieval systems two specific aspects of multilingualism have to be carefully taken into account: on the one hand the intimate link between language and law, covering the crucial issues of conveying legal concepts across languages and, on the other, the relationship between legal systems which, while a problem in its own, is exacerbated in a multi-language environment. Therefore, the aspects which seem worthy of analysis as being closely related to the development and effectiveness of legal information retrieval systems are essentially of a linguistic, translational and comparative law nature. These are briefly examined below.

2.1 Linguistic issues

It has been said more than once that "the law is a profession of words." (Mellinkoff 1963). Many of the problems about meaning that are of concern to language specialists turn out to be of interest to legal professionals too, 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 they free-text or metadata) through strings.

While the dependency of legal concepts of a particular legal system is the key characteristic of legal language as a system of signs, such concepts are claimed to be not forever fixed and unchangeable, as they change when legal experience changes (Luhmann 1993). The change of legal concepts is brought about through legal argumentation (Kjaer 2004b); this is evidenced in the multifaceted role of language in establishing, maintaining, and changing concepts, which makes cross-system interoperability even harder to achieve (Gémar 1982).

A further argumentation is based on the branch of linguistics known as ‘linguistic relativity’, focusing on the fact that what one language system conceptualises in one way is not conceptualised in the same way in all other language systems. This is especially true of legal terminologies at system level.

Coming to an understanding across legal languages and legal systems is definitely a hard process and establishing share knowledge can only be achieved if the actors base their interpretation on the same tacit suppositions about the world, which forms the point of departure for 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.

2.2 Legal translation issues

There is a general consensus 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 (Capellas-Espuny 2004). As regards the area of law, the relation between law and language can significantly broaden the scope of legal translation theory. In fact, while it can be assessed that everyday language already implies a formalised way of communication, legal language introduces a supplementary system of formalisation (Prakken & Sartor 1995). Differences between legal language and ordinary language lie, in the view of a number of authors (Phillips 2003, Pozzo 2005), in the use of words and in the connotations they bear when they are incorporated in legal speech.

Legal realities cannot coincide or can coincide only partially when they focus on a common international legal phenomenon. The opinion is widely shared that legal translators are more rigidly bound to specialized knowledge than the translators of everyday language or humanities (Gizbert-Studnicki 1987) 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.

On a practical level legal translation requires both a comparative study of the different legal systems and an awareness of the problems created by the absence of equivalents (Capellas-Espuny 2004). This means that legal translators must be familiar with the legal culture of the target context in order to reformulate an equivalent meaning through what they judge to be the most appropriate linguistic and legal expression. In fact, a particular concept in a legal system may have no counterpart in other systems, or a specific 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” (Pelage 2000), which makes it very difficult to find equivalents for culture-bound terms, especially those concerning legal concepts, procedures and institutions (Harvey 2002).

Addressing the problem of functional equivalence is a major task of legal translators, who need to find an equivalent in the target language legal system for the terms of the source language legal system. Because of the system-specificity of legal terms, full equivalence can only be achieved between the languages of the same legal systems, as is the case when translating within a bilingual or multilingual legal system, such as Belgium, Finland, Switzerland and - to some extent - Canada (Gémar 1988).

At operative level it is worth noting that one main problem legal translators are faced with is the poor benefit of legal dictionaries, which often fail to meet their expectations for conveying the meaning of the source legal language into the target one. According to de Groot and van Laer (2006), who extensively analyse the problem of functional equivalence of legal terms, only very

few dictionaries are reliable tools. These authors propose to indicate the degree of equivalence and to provide alternatives according to area of law, legal system and use.

2.3 Comparative law issues

The problems raised by multilingualism are strictly connected to 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 (Hoecke 2004, Sacco 1991) express their doubts about the possibility of a real comparison of legal systems. Many comparatists are strongly concerned with the implications of the differences existing between the cultural contexts underlying the various legal languages and with the difficulties in transferring legal meanings and legal concepts from one legal system to another, even when the same language is used (Kjaer 2004a). A number of frequently mentioned examples are made to refer to this phenomenon, such as *société* in the French legal language in France, which has not the same meaning as *société* in the French legal language in Belgium. 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 legal terminology (Heutger 2004).

In recent years research studies increasingly concentrate on the relationship between legal language and the 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. Many are the initiatives aimed at laying the foundation for a common frame of reference and at promoting, for example at European Union level, a pan-European terminology. (Pozzo 2003).

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 (Schlesinger 1955). Such approaches are likely to influence the cross-system retrieval techniques adopted in the implementation of retrieval systems with special reference to Europe. In fact 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 mentioned as "legal formants".

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 searched and an integrative comparison has newly emerged among legal scholars. The actual debate among comparatists 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.

The convergence or divergence approaches mentioned above are key elements for implementing multilingual retrieval tools and services: according to the chosen approach, the methods adopted in these systems will facilitate terms and concepts to be matched across legal systems, adapting

concepts of different systems and helping contextualization so as 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. The ITTIG Project for promoting worldwide multilingual open access: the thesauri alignment project

This overview shows that much progress has been made in research work and related applications and a number of international activities are in place aimed at evaluating systems and tools for cross-language retrieval, while playing an important role in stimulating system development. A major factor towards successful retrieval across languages is the use of adequate lexical resources. Electronic bilingual dictionaries are not the proper tool for this purpose, requiring some kind of pre-processing before they can be profitably used in automatic systems. Specialized resources such as thesauri seem to be the best solution for cross-language retrieval applications. Good document indexing and satisfactory retrieval in single collections is in fact based on the availability of high quality thesauri. Effective cross-language and cross-collection retrieval depends on interoperability among thesauri. This means using a particular thesaurus for users' query and mapping it to thesauri in other languages, to more specialized vocabularies, or to different versions of a thesaurus, in order to retrieve documents in different languages, consistently with the original query.

In this context the Institute of Legal Information Theory and Techniques of the National Research Council (ITTIG-CNR) intends to develop solutions to match national legal terminologies at conceptual level. Currently no effective semantic services exist dealing with cross-lingual legal information indexing and retrieval, preserving the basic sense and value of legal concepts across languages and legal orders.

The proposed project intends to develop a collaborative Web platform around national thesauri capable to provide automatic, language independent facilities to support conceptual mapping and language localization services for mono/poly-lingual thesauri, using a pivot thesaurus.

Starting from a multilingual thesaurus as pivot resource, the system will allow to load a new thesaurus, developed by a local editorial staff for specific document indexing purposes, and it will provide a semi-automatic conceptual mapping between the pivot thesaurus and the loaded one. Mapping proposals will be afterwards validated by human experts in a collaborative Web environment.

This service will also contribute to collaborative enrichment and maintenance of the multilingual pivot thesaurus. This will be done by domain experts who can propose and negotiate the inclusion of new concepts, coming from local thesauri, in the pivot one, as well as describe and localize such new concepts in different languages.

This achievement would have essential positive effects for a wide category of legal users. In particular, through linguistic-conceptual correspondence definition activities it could be possible to set up experimental applications and systems capable of fostering awareness and understanding of different countries legal concepts, in a word facilitating worldwide open access to law.

3.1 Automatic Thesaurus Mapping

The proposed environment will be endowed with specific implementation of language-independent algorithms able to provide semi-automatic mapping between thesaural concepts.

Such algorithms have been developed and tested in a recent experience in thesauri/ontology mapping recently carried out in favour of the Office for Official Publications of the European Communities (OPOCE) will be exploited (Invitation to tender N. 10118 Eurovoc studies in the frame of a Restricted Procedure following the Call for Expression of Interests ("AMI") published in the OJS N. 2005/S 195-191899, field N. 2: Multilingual thesauri).

A methodological framework for thesauri semantic mapping based on thesauri schema information, as well as a specific approach within such framework on the OPOCE case-study, have been proposed and widely tested (Francesconi et al., 2008). In this context a methodology has been proposed to describe thesaural term semantics and to provide automatic mapping between thesaural concepts, to be validated by humans. Term mapping between thesauri is a process which aims at matching term meaning (term semantics) rather than finding their lexical equivalences. Term semantics is conveyed not only by its morphological characteristics, but also by the context in which the term is used as well as by the relations with other terms. Therefore in this tender semantics of a thesaurus term has been proposed to be represented, according to an ascending degree of expressiveness, by information that can be drawn from thesauri terms and their mutual relationships, in particular:

- 1) its Lexical Manifestation: a string of characters normalized according to pre-processing steps;
- 2) its Lexical Context: a term vector of binary entries representing terms in relation;
- 3) its Lexical Network: a directed graph where nodes are terms and the labeled edges are semantically characterized relations between terms included in a specific thesaurus.

Specific language-independent methods to measure the similarity between homogeneous representations of terms semantics have been given and their efficacy widely tested.

A ranking function is able to provide a similarity measure between terms; when extended to a set of target terms such a function may provide a matching order among such terms.

For Lexical Manifestations, a similarity measure based on the Levenshtein distance (the number of insertions, deletions, and substitutions of characters required to transform one string into another, normalized by the length of the longest string ($dist_{lev}(s_i, s_j)$) applied on pre-processed strings (s_i, s_j), is defined as $sim_{lev}(s_i, s_j) = 1 - dist_{lev}(s_i, s_j)$, where $sim_{lev} \in [0, 1]$.

For Lexical Contexts (binary vectors) the similarity can be measured as the correlation between such vectors, quantified as the cosine of the angle between these two vectors.

For Lexical Networks (directed graphs) the most frequently addressed graph similarity measure (Sowa, 1984) (Montes-y-Gomez et al., 2001) is the *Graph Edit Distance*, namely the minimum number of nodes and edges deletions/insertions/substitutions to transform a graph g_1 into a graph g_2 . Because of computational complexity we have considered three variants of the Graph Edit Distance expressing 1) how many concepts two graphs g_1 and g_2 have in common (Rasmussen, 1992); 2) how similar the relations between the same concepts in both graphs are; 3) how to combine the previous measures (*Graph similarity*) (Cai et al., 2004).

Candidate terms of the target thesaurus are ranked according to the similarity measure values $sim \in [0, 1]$ and a semantics is assigned to the mapping relation using proper heuristic thresholds ($T1, T2 \in [0, 1]$)

- | | | |
|----|-----------------|----------------------------------------------|
| if | $sim < T1$ | => No Match |
| if | $T1 < sim < T2$ | => Partial Match (broadMatch or narrowMatch) |
| if | $T2 < sim$ | => Exact Match |

Term relations are then expressed using RDF SKOS Mapping standard.

In particular a collection (“gold standard”) of 624 relations between thesauri concepts have been identified by experts, including 346 exact match relations. On this data set, experiments showed good performances of the automatic mapping algorithms in identifying matching concepts with respect to human predictions (an average of around 80% matching concepts identified by the automatic tools with respect to the “gold standard”) (for details see Francesconi et al., 2008).

3.2 The Thesaurus Alignment Environment

The automatic mapping algorithms will be conceived to support the activities of an editorial staff of experts, cooperating in the identification of matching concepts by a collaborative Web platform conceived as a thesaurus alignment environment.

In this framework:

1. the experts, using the editorial environment, will be able to choose a concept in the source thesaurus;
2. the system will be able to retrieve a number of possible matching concepts as well as ranking them according to the match probability (in terms of similarity measures);
3. the experts will be able to
 - confirm the system suggestion (matching concepts and related relation);
 - select an alternative among those suggested by the system;
 - browse the thesaurus to search for matching concepts not identified by the system.

In this framework an effective hint to the experts is the system ability to retrieve the most part of matching concepts in the minimum number of predictions suggested by the system, so to reduce the number of global browsing of a target thesaurus to search for a matching concept.

This way the experts, for the most part of the mapping requests, will be required to select a proposed matching concept and, in case, to confirm or specify the type of relation proposed by the system. Otherwise the experts will browse the whole thesaurus to select a different matching concept.

Using the platform collaborative facilities the editorial staff, responsible for the local thesaurus, can cooperate in providing the most satisfying thesaurus match and localization.

New or more specific local thesauri concepts, not included in the pivot resource, can be highlighted and proposed for inclusion by the local thesaurus editorial staff. Using the collaborative Web platform, a community of interest, endowed with specific privileges, can manage the proposed concepts in a sort of semantic wiki, providing definitions in the native thesaurus language as well as in a pivot language (English), establishing proper relationships with respect to existing concepts, localize them in different languages and finally, accept them for inclusion in the pivot resource.

4. Conclusion

Access to information regardless of language barriers is a key factor for effective global sharing of knowledge. This need has created a convergence of interests from various research communities around cross-language retrieval, which is becoming more and more popular between information specialists. Referred to legal information, cross-language retrieval 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. 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. 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, a joint activity among actors with quite different skills is recommended. Jurists, linguists, jurilinguists, translators should commit themselves to cooperate with researchers in the field of new technologies and promote large-scale comparative studies on several fields of law. In particular, through linguistic-conceptual correspondence definition activities carried out in cooperation among institutional organisations committed to indexing and delivering legal material world-wide, it can be possible to set up experimental applications, pilot projects and systems capable of fostering awareness and understanding of different countries' legal concepts, thus facilitating worldwide open access to law material.

In order for legal material to be accessible worldwide, interoperability is paramount, but so far there are not many solutions for cross-language legal information access and no truly multilingual information system is available yet. Systems developed so far greatly concentrate on pairs rather than on multiple languages, the latter being a much more complex endeavour. While lots of experiments and discussions focus on the developments of vocabulary matching over languages as such, a strong need is felt to study some kind of interlingual mechanism allowing the transfer of meanings between languages. This would enable matching to equivalent terms in all languages and realizing cross-language retrieval in any of the language combinations. Furthermore, linguistic tools and multilingual systems developed so far for the law domain mostly concentrated on a specific theme or field of law: this approach is sound as these systems are likely to provide insights for larger applications and developments. In this sense, the ITTIG Project is a pioneer project within Europe, trying to find solutions to meet national legal terminologies at semantic and conceptual level.

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: legislators, judges, legal professionals and eventually citizens in an effort towards a common understanding of law beyond language and system barriers.

The service for thesauri alignment will strongly enhance accessibility and diffusion of legal information across languages and legal orders, thus contributing to provide European citizens with a better multilevel awareness and understanding of EU and national laws.

Bibliography

- Cai Wen-Tao / Wang Sheng-Rui and Jiang Qing-Shan (2004): "Address extraction: a graph matching and ontology-based approach to conceptual information retrieval", in: Proceedings of the Third International Conference on Machine Learning and Cybernetics, pp. 1571–1576.
- Cappellas-Espuny, Gemma (2004): "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", in: Journal of diplomatic language, 1:4
- de Groot, Gerard René (1996): "Het Vertalen van Juridische Informatie", in: Preadvies Nederlandse Vereniging voor Rechtsvergelijking, 53, Deventer: Kluwer, 1-77. ISBN 9026829949
- de Groot, Gerard René / van Laer, Conrad J.P. (2006): "The dubious quality of legal dictionaries", in: International journal of legal information, 34:1 (Spring), 65-86
- Francesconi, Enrico / Faro, Sebastiano / Marinai, Elisabetta (2008): "Thesauri Alignment for EU eGovernment Services: A Methodological Framework", in: Proceedings of the JURIX Conference: The Twenty-First Annual Conference, E. Francesconi, G. Sartor, and D. Tiscornia (Eds.), pp. 73-77.
- Gallo, Giovanni (1999): "Les linguiste juristes de la Cour de Justice des Communautés européennes", in: Sacco, Rodolfo/ Castellani, Luca (eds): Les multiples langues du droit européen uniforme. Torino: L'Harmattan Italia, 71-89
- Gémar, Jean Claude (1982): "Langage du droit et traduction: essais de jurilinguistique", in: Gémar, Jean Claude (ed): The Language of the Law and Translation: Essays on Jurilinguistics. Québec: Linguattech-Conseil de la langue française
- Gémar, Jean Claude (1988): "Le Traducteur Juridique ou l'interprète du Language du Droit", in: Nekeman, Paul (ed.): "Translation, our Future/La Traduction, Notre Avenir", Proceedings of the XIth World Congress of FIT, Maastricht, 422-430.
- Gizbert-Studnicki, Tomasz (1987): "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, 99-114
- Harvey, Malcolm (2002): "What's so special about legal translation?", in: *Meta*, 47:2, 177-185
- Heutger, Viola (2004): "A more coherent European wide legal language", in: *European Integration online Papers (EioP)*, 7:2, <http://eiop.or.at/eiop/pdf/2004-002.pdf>
- Hoecke, Mark Van (2004): *Deep level comparative law*. In Van Hoecke, Mark (ed): *Epistemology and methodology of comparative law*. Oxford: Hart. ISBN 1841134430
- Kjaer Anne Lise (2004a): "A common legal language in Europe?", in: Van Hoecke, Mark (ed.): "Epistemology and Methodology of Comparative Law", Oxford: Hart.
- Kjaer Anne Lise (2004b): "Convergence of European legal systems: the role of languages", in: *Language and culture*, 29, 125-137
- Luhmann, Niklas (1993): "Das Recht der Gesellschaft", Frankfurt/Main: Suhrkamp
- Mellinkoff, David (1963): "The Language of the Law", Boston; Toronto: Little, Brown
- Montes-y-Gomez, Manuel / Gelhukh, Alexander / Lopez-Lopez, Aurelio and Baeza-Yates, Ricardo (2001): "Flexible Comparison of Conceptual Graphs", *Lecture Notes in Computer Science 2113*, Springer-Verlag.
- Moreteau, Olivier (1999): "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, 143-162.
- Pelage, Jacques (2000): "La traductologie face au droit", in: *Proceedings of the International Colloquium University of Geneva, February 17-19*. Berne: Association Suisse des Traducteurs, Terminologues et Interprètes (ASTTI), Ecole de traduction et d'interprétation de Genève (ETI), 125-131.
- Phillips, Alfred (2003): "Lawyer's language: how and why legal language is different", London; New York: Routledge.
- Pozzo, Barbara (2003): "Harmonisation of European contract law and the need of creating a common terminology", in: *European review of private law*, 11:6, 754-767.
- Pozzo, Barbara (ed.) (2005): "Ordinary language and legal language", Milano: Giuffrè.
- Prakken, Henry/ Sartor, Giovanni (1995): "On the relation between legal language and legal argument: assumptions, applicability and dynamic priorities", in: *Proceedings of the Fifth International Conference on Artificial Intelligence and Law*. New York: ACM, 1-10.
- Rasmussen, Edie (1992): "Clustering algorithms", in: *Information Retrieval: Data Structures & Algorithms* (W. B. Frakes and R. Baeza-Yates, eds.).
- Sacco, Rodolfo (1991): "Legal formants: a dynamic approach to comparative law", in: *The American Comparative Law*, 1, 343-358
- Schlesinger, Rudolf B. (1995): "The past and future of comparative law", in: *American Journal of Comparative Law*, 43, Summer, 477-481
- Sowa, John F. (1984): "Conceptual Structures: Information Processing", in: *Mind and Machine*. Addison-Wesley