Ágnes Hajdu Barát — University of Szeged, Hungary

Knowledge Organization in the Cross-cultural and Multicultural Society

Abstract

This paper explores the changing role of knowledge organization in the representation of information retrieval languages in the Cross-cultural and Multicultural Society. The paper studies the concept of Cross-cultural and Multicultural Society — the relationship between language and culture. Globalization and the Internet bring cultures, intellectual products and users closer to one another, but it shows their differences, too. With respect to KO the basic problem is: How can users from different backgrounds use the similar interfaces similar? The author introduces a Hungarian solution what maps the different topical terms, subjects, UDC codes on conceptual level.

1. Introduction

Nowadays a worldwide network infrastructure is rapidly moving information across national boundaries, but we often forget about the linguistic and cultural barriers, which make information retrieval more complicated. The solutions can be:

- the usage of multilingual thesauri
- multilingual subject headings
- the adaptation and usage of classification systems, which are not based on language, as UDC
- machine translation or machine-aided translation.

There are more and more emphases put in constructing a worldwide network of information. Information retrieval has become global, although a "global language" has not been developed yet. To some extent English can be called such, because it is used worldwide in different fields of research in development. Next to globalization another process can be observed as well: localization. Only those products which use the local language can be successful in a local market. Nowadays product designers use the local language and try to accommodate the special features of the local culture and customs. Local languages have become more and more important, although the network has stayed global. We can reach Norwegian, Korean or Dutch documents from Hungary, and Hungarian documents are available also from the furthest part of the world.

2. Language and culture
2.1. Language

European Union gives attention to evaluating policy measures for minority languages in Europe. There is a growing awareness of the importance of Europe's linguistic diversity. "The yearly reports of European Centre for Minority apply in particular to minority language policies." (Grin, 2000) The report in 2000 said:

> Linguistic diversity has long been recognised as a defining characteristic of Europe, as an element conducive to the assertion of its identity, and as a condition for its democratic development; it is now also increasingly recognised as an asset enhancing creativity in all domains of social life, including social cohesion and economic performance. (Grin, 2000)

As we mentioned, English is one of the most important languages in global information retrieval, not the only one. Sometimes it seems like the only one used language in international and scientific communication. Researchers who investigate the development tendencies of the Internet say that nowadays the number of users who use English as a mother tongue is smaller that the non-English ones (Japanese, Chinese, Spanish). Other

research proved that the number of Non-English Web pages has overtaken the number of English Web pages.

That is why there is a growing need for cross-language retrieval systems, which help to overcome the languages barriers. A cross-language retrieval system is needed for those who can only search in one language, or for those who can speak more than one language, but his/her information retrieval is easier and more effective in one language.

Some typical cross-language retrieval tasks:
- search a monolingual collection in a language that the user cannot read
- retrieve information from multilingual collection using a query in a single language
- select images from a collection indexed with free text captions in an unfamiliar language
- locate documents in a multilingual collection of scanned page images.

2.2. Culture not only language

Culture is peculiar to Homo sapiens, together with material objects used as an integral part of this behavior. Thus, culture includes language, ideas, beliefs, customs, codes, institutions, tools, techniques, and works of art, rituals, and ceremonies, among other elements. There are 6000 languages in the world, which means almost 6000 different cultures, and means everyday problems in finding information in a databases or digital library originate from different cultures.

Hungary is a small country, but there are 13 recognized *communities* of ethnic minorities traditionally (Roma, German, Slovak, Croatian, Romanian, Serbian, Slovene, Armenian, Greek, Bulgarian, Polish, Ruthenian, and Ukrainian). Naturally there is much more nationality, but these raised minorities have special national ethnic culture in Hungary and own school systems, radio programs etc.

Human sociality is saturated with meaning (Holtgraves & Kashima, 2008, 73). The question of meaning making and meaning exchange — how people engage in meaningful thought and action in social contexts — is a central question of social cognition. (Kashima, 2000).

Definition of culture

Culture can be defined from many perspectives, according to the social science one is involved with. One of the well known is the definition by Kluckhohn.

> He clearly distinguishes culture from the limited concepts of ordinary language, history and literature. The anthropological term designates those aspects of the total human environment, tangible and intangible, which have been created by men. A 'culture' refers to the distinctive way of life of a group of people, their complete 'design for a living'. Culture seems to be the master concept of American anthropologists. For ethnologists, folklorists, anthropological linguists, archaeologists and social anthropologists, culture is always a point of departure or a point of reference if not invariably the point of emphasis. (Kluckhohn, 1951; Heidrich, 2002, 26)

Kroeber and Kluckhohn gave a very complex definition of culture:

> Culture consists of patterns, explicit and implicit of and for behaviour acquired and transmitted by symbols, constituting the distinctive achievement of human groups, including their embodiment in artefacts; the essential core of culture consists of traditional (i.e., historically derived and selected) ideas and especially their attached values; culture systems may, on the one hand, be considered as products of action, on the other, as conditioning elements in a future action. (Kroeber and Kluckhohn, 1952)

2.3. Cross-cultural or cross-national

The differences between cultures are rare questions, usually the studies focus on the languages in the field of knowledge organization. Susanna Keränen studied this problem in her PhD dissertation (Keränen, 2002). She has examined equivalence of meaning, methods and dimension for measuring them. She presented many case studies in her practice under created an English and Finnish multilingual thesaurus. There are three problems of research in cross-cultural studies must be considered:

1. The term 'culture' has been used in so many ways that academics fail to arrive at a consensual definition;
2. The distinction between cultural and national boundaries is problematic with the consequence that 'nation' has been used as a synonym for 'culture';
3. The measurement of the impact of cultural attributes on organisational functioning is problematic due to the lack of definition clarity. (Clark, 1994)

Thus the other problem apart from the different aspects of cultural definitions is that 'culture' is often substituted for 'nation'. The two terms are often used interchangeably. 'Nation' is invariably used as a synonym for 'culture' with the consequence that national distinctiveness is interpreted as cultural differences. Therefore, according to Bhagat and McQuaid (1982), *"what are called cross-cultural differences are really only crossnational differences".*

Clark is correct in saying that it is an oversimplification to argue that cultural boundaries correspond to national (political) borders, since no nation is so pure that all its members share a common set of cultural factors. Any nation is a patchwork of different and unique subjective cultures. Thus, whereas two nations may share a common language, climate, political system and religion, differences in the mixture of their subjective cultures (subcultures?) will result in distinctive belief systems, norms, values and cognitive maps. A national culture therefore reflects the unique interaction between, and a combination of, a set of subjective cultures. There are many examples of these subjective (sub)cultures, which are clearly recognizable as regional cultures in Belgium, Canada, Germany, the UK, etc. (Heidrich, 2002, 27)

There are more levels and sources of Multiculturalism: personal age, geography, national affiliation, religious affiliations, language, tradition, historical elements, climate, economy, racial mix, political system, gender, modernization and many other intangibles. So if we are looking for any general knowledge organizational solution we should know the problem is very complex and a simple translation does not give a correct solution. Therefore the result on conceptual level can give sufficient accomplishment.

Features of culture:
- social product — it is based on mutual effect and interaction
- learned cognition and knowledge — socialization, system of cognition and norm
- individual — technical and intellectual appearance, what interpret the world by language
- different meaning — it is based on cultural experience

Propositions of culture:
- Culture is learned
- Culture derives from the biological, environmental, psychological, and historical components of human existence
- Culture is structured
- Culture is divided into aspects
- Culture is dynamic
- Culture is variable

- Culture exhibits regularities that permit its analysis by scientific methods
- Culture is the instrument whereby the individual adjusts to his total setting, and gains the means for creative expression. (Herskovits, 1940)

Situation of intercultural communication:

- Tourist
- Official visiting, scholars, delegation
- Immigrants and migrants
- Multinational workplace
- Minority, ethnical groups, subculture
- Internet users etc.
- People in the globalization

3. Equivalence in knowledge organization and digital environment

3.1. Equivalence

The basis of the problem is equivalence between worlds and concepts in multicultural and multilingual environment. We should define a threshold value, the equivalence level for KO solutions (multilingual thesauri or subject headings, the adaptation and usage of classification systems and machine translation or machine-aided translation). We should try to anticipate and decide what kind of topical term will most likely be used in our information retrieval system.

In International Organization for Standardization's standards Documentation — Guidelines for the establishment and development of multilingual thesauri (ISO, 1985, p. 8) equivalence is divided into exact equivalence, inexact equivalence, partial equivalence, single-to-multiple equivalence and nonequivalence (Keränen, 2002, 797).

3.2. Digital environment

More and more digital libraries are designed and available all around the world. These libraries are produced in national languages, so they can be the target of cross-cultural and cross language information retrieval. The designer and content providers and users of the digital libraries can have different cultural backgrounds, so these libraries are cross-cultural in at least three levels. Cross cultural usability issues of digital libraries are an important area for social informatics. The misinterpretation of colors, form, symbols, metaphors, language and use are the most common cross-cultural design failures.

For number of years the focus of the research on digital libraries has been to solve the technical problems of digitalization and development. Usability has seemed a relatively minor concern. Although there were researchers and developers, who have investigated this cross-cultural issue considering the cultural aspects of designing of global information systems in a deeper way. This line of work can be grouped into four overlapping categories:

- cross-cultural usability studies
- internationalization of industrial products
- internationalization of software
- international web design.

Elke Ducker deal with cross-cultural usability, revealed in one of her researches the color preferences of students with different cultural backgrounds can differ dramatically. She pointed to the fact the usage of digital library and its metaphors are not unambiguous for

everybody. Computing metaphors have become an integral part of information systems design, yet they are deeply rooted in cultural practices.

Nowadays there are several projects in Europe which goal is to help and support the digitalization projects of the European countries, and to deal with the problem of cross-language retrieval in these digital libraries. Such projects are: Minerva and Michael, MACS (Multilingual Access to Subjects), Calimera, Madiera project (Multilingual Access to Data Infrastructures of the European Research Area) etc.

4. Multilingual thesauri

One of the solutions of helping the cross-language information retrieval can be the usage of multilingual thesauri. A multilingual thesaurus is more than just "putting together" several monolingual thesauri. A true multilingual thesaurus offers full conceptual and terminological inventories for each language represented; most importantly, it presents a fully developed thesaurus structure (i.e. all semantic relations of equivalence, hierarchy, and association/affinity) in each one of the languages of the thesaurus.

One of the main requirements of designing multilingual thesauri is that the different versions should be developed simultaneously. A multilingual thesaurus should not be a translation of an existing monolingual thesaurus. The practice although is not equivalent with the theory, because there are several multilingual thesauri which were developed by translation because of economic reasons.

Promising accomplishment is the Eurovoc thesaurus. The Eurovoc thesaurus is published in the official languages of the European Community (Bulgarian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Slovak, Slovene, Romanian, Spanish and Swedish) and one another language (Croatian). All these languages have equal status: each descriptor in one language necessarily matches a descriptor in each of the other languages. However, there is no equivalence between the non-descriptors in the various languages, as the richness of the vocabulary in each language varies from field to field (Eurovoc, 2008).

5. General Information Retrieval Language Dictionary — MÁTrIkSz

The usage of classification systems that are not based on languages can be another solution for the problem of cross-language information search, although there are several problems that have to be solved. Significant advantages are the conceptual level, the absence of linguistic signs among others, but the designers should observe, place and adopt the classification concepts to not only a foreign conceptual structure but establish connections with different subjects and topical terms.

The Mátriksz (*General Information Retrieval Language Dictionary)* project was finished by 2000 in the Széchényi National Library (Hungary), and there are some characteristics:

- It is the largest Hungarian Information Retrieval Language Dictionary on natural languages (about 28,000 lexical terms by 2002).
- It was made with the help of modern technology, by computer and it is used on electronic supporting, concerning informational network, Internet, too. Its software is RELEX.
- It is possible to develop and keep on the Internet dynamically.
- It can work with MARC format.

- It has very easy sentence construction and syntax.
- It is of high level, efficient, open, flexible, easily used, with rich vocabulary and a clear structure scheme.
- It has involved some of the existing Hungarian thesauri since 1990. Their numbers are over 20.
- It can be used by automatic classification and indexing systems, too.
- It can involve and the UDC index of the new Hungarian edition.
- The weakness of these thesauri was the few bibliographic records, which are really used in these systems. So the MÁTrIkSz project was a good opportunity for the KÖZTAURUSZ and the OSZK thesaurus. (Hajdu Barát, 2003)

The concept of project was: to give the possibility of the common search of the topical terms (subject, descriptors, UDC codes etc.) in some big libraries' catalogues and databases with different structures and information retrieval methods.

The MÁTrIkSz is the co-operating system of the participant information seeking methods and databases. The KÖZTAURUSZ/OSZK thesaurus took a prominent role among them, but the participants are its equals. They are independent, and have the chance of common pursuit. There is a suitable user interface to search each participant dictionary and to utilize different topical terms, for example UDC terms, subjects, descriptors.

The MÁTrIkSz has its own information retrieval language dictionary, too, which is independent of the other bibliographical databases, but searches those systems. There is a common online index and its records have a local identification. The result is not only a bibliographic record, but there are contents of topical terms. The different classification systems and their bibliographical databases appear in a unified and homogeneous environment. Each database keeps their own descriptors, terms and if they have any similar expressions with different morphological forms, they make the KÖZTAURUSZ's phrases appear in 750 field *(equivalent)* of the MARC record. Hits would come from all bibliographic databases, presenting diverse expressions of participant libraries without any analogy. Hits would come from only the own database (Bánki, 2002, p. 35).

The project assisted the new medium edition of UDC index. The adaptation of UDC MRF relating the KÖZTAURUSZ/OSZK thesaurus and the UDC index were built in the thesauri. The UDC codes became searchable in the MÁTrIkSz system, and they give the basis of searching in multicultural and multilingual environment, too, because the UDC codes and the descriptors are in conceptual level and the UDC independent of language.

6. Conclusion — Concentration to knowledge on the conceptual level

This paper aims at exploring the changes of knowledge organization in the role and possibility of representation of information retrieval languages in the Cross-cultural and Multicultural Society. Over the last decades there have been moves towards unification and standardization of bibliographic systems all over the world. This means the traditional classifications (for example UDC) and thesauri are now being stretched to cover cultural and linguistic artifacts and concepts would be quite different and broader from those originally intended.

References

Bánki, Zsolt István (2002) *Általános Információkereső-nyelvi szótár elméleti kérdései és gyakorlati megvalósítása*. Budapest: ELTE BTK, 43 p.

Bhagat, R. S. & McQuaid, S. J. (1982) The Role of Subjective Culture in Organisations: A Review and Directions for Future Research. *Journal of Applied Psychology Monograph*, 67, 5.

Bognárné Lovász Katalin (2005) *Kultúrákon és nyelveken átívelő többnyelvű információkeresés*. Budapest, ELTE, 28 p.

Clark, T. (1994) International Human Resource Management. Perspectives, Problems, Polycentrism (Paper presented to BUIRA Annual Conference, Worcester College, Oxford).

Duncker, Elke (2002) Cross-Cultural Usability of the Library Metaphor. *JCDL '02*, July 13–17, pp. 223–230 Portland, Oregon. http://portal.acm.org/citation.cfm?id=544269. (Accessed 11 Nov. 2007.)

Eurovoc (2008) *Eurovoc Thesaurus*. http://europa.eu/eurovoc/sg/sga_doc/eurovoc_dif!SERVEUR/menu!prod!MENU?langue=HU. (Accessed 11 Nov. 2007.)

Grin, François (2000) Evaluating policy measures for minority languages in Europe: towards effective, cost-effective and democratic implementation (European Centre for Minority Issues report) Flensburg: ECMI. http://www.ecmi.de/34/2004/10/07/Evaluating-Policy-Measures-for-Minority-Languages-in-Europe.php. (Accessed 11 Nov. 2007.)

Hajdu Barát, Ágnes (2003) General Information Retrieval Language Dictionary in the Széchényi National Library (Hungary). In: *Tendencias de Investigación en Organización del Conocimiento — Trends in Knowledge Organization Research*. Salamanca, Universidad de Salamanca, p. 463–468.

Heidrich Balázs (2002) Business as unusual the role of national cultural backround in corporate life. *European Integration Studies*, Miskolc, 1(2), pp. 25–36.

Herskovits, M. J. (1940) *Man and His Works*.

Holtgraves, Thomas M. & Kashima, Yoshihisa (2008) Language, Meaning, and Social Cognition. *Personality and Social Psychology Review* 12; pp. 73–94. http://psr.sagepub.com/cgi/content/abstract/12/1/73. (Accessed 11 Nov. 2007.)

Kashima, Yoshihisa (2000) Recovering Bartlett's social psychology of cultural dynamics. *European Journal of Social Psychology*, 30, pp. 383–403.

Keränen, Susanna (2002) Content Management — Concept and Indexing Term Equivalence in a Multilingual Thesaurus. In: *Informing Science InSITE — "Where Parallels Intersect"* June, pp. 795–809. http://web.abo.fi/~skeranen/content.html. (Accessed 11 Nov. 2007.)

Kluckhohn, F. (1951) *The Study of Culture*. Stanford University Press.

Kroeber, A. L. & Kluckhohn, F. (1952) Culture: A Critical View of Concepts and Definitions. In: *Peabody Museum Papers*, vol. 47, no. 1. Cambridge, Mass.: Harvard University.