*Journal of the*
*International Institute for Terminology Research*
*– IITF –*

# TERMINOLOGY SCIENCE & RESEARCH

***Vol. 17 (2006)***

Publisher:
Address:
Redaktion: International Institute for Terminology Research (IITF)
Secretary General: Christer Laurén

**CONTENTS**

## Foreword

Volume 17 of the Journal Terminology Science and Research contains three articles, all of them from Scandinavian terminologists. We find that they demonstrate the diversity of terminology research and practice in Scandinavia. Moreover, they reflect the rapid development which takes place in terminology science, whose range of applications is steadily increasing.

We are two editors who cooperate in compiling and preparing the journal: Nina Pilke (University of Vaasa) and Birthe Toft (University of Southern Denmark). Please submit articles to the editorial board via one of our e-mail addresses (see below).

As was mentioned in the 2005 issue, the Board of the International Institute for Terminology Research has decided to publish future volumes of the Journal of Terminology Science and Research electronically via IITF's portal and to publish selected articles in book form with intervals of a couple of years. The portal with the guidelines for authors etc. can be found under the websites of the University of Vaasa, see http://www.uwasa.fi/hut/svenska/iitf/presentation.html.

Vasa and Kolding, December 2006

Nina Pilke
nina.pilke@uwasa.fi

Birthe Toft
toft@sitkom.sdu.dk

**Asta Høy**
**Authorized Translator, PhD**
**Consultant of the National Board of Health**

# COMING TO TERMS WITH SNOMED CT® TERMS: LINGUISTIC AND TERMINOLOGICAL ISSUES RELATED TO THE TRANSLATION INTO DANISH

Acknowledments: Thank you to SST staff members Ulrich Andersen, Camilla Wiberg Danielsen, Janni Lerche, and Ulla Magdal Petersen for useful advice as to form and content of this article.

## INTRODUCTION

For many years, the Danish medical terminology, as represented in text books, articles, encyclopedia, dictionaries, and classifications, has been marked by discrepancies regarding the choice of terms and morphological and syntactical features. The growing use of electronic aids and information technologies has brought about a recognition of the need for terminological consensus as far as general, commonly used medical concepts are concerned.

This recognition led to two important initiatives at the beginning of the present century:

- the establishment of a National Health Terminology Committee (Det Nationale Begrebsråd) whose main task consists in defining a number of top-level concepts within areas such as administration, medication, and quality management, and
- the launching of the SUNDTERM Project the aim of which is to establish a common health terminology which may be used for the Basic Structure for Electronic Health Records (GEPJ) and which allows for mapping of terms and concepts to existing classifications in Danish like the national versions of ICD, ICF, ICPC, nursing diagnoses, etc. (The Ministry of the Interior and Health, 2003).

The translation of Systematized Nomenclature of Medicine Clinical Terms, SNOMED CT®, into Danish, carried out by the Danish National Board of Health (SST), forms part of the SUNDTERM Project. This article describes and discusses the bases and principles on which the huge SUNDTERM translation task is based as well as some important issues and problems which have arisen in connection with the translation of – so far – almost 150,000 medical terms.

The author of the article has been involved in the SST project as a language and terminology consultant since October 2003.

## SNOMED CT® – A MULTIHIERARCHICAL TERMINOLOGY

This vast terminology system was created by combining SNOMED Reference Terminology (SNOMED RT) and Clinical Terms Version 3 (CTV3), formerly known as the Read Codes, which was created on behalf of U.K. Department of Health and is a Crown copyright. SNOMED CT®, first released in January 2003, is a registered trademark of the College of American Pathologists (CAP). It comprises 350,000 concepts represented by some 1,000,000 terms, i.e. "preferred terms" and synonyms. As explained by Spackman and Reynoso (Spackman et al. 2004), it is "a terminological resource designed to be implemented in software applications to represent clinically relevant information reliably and reproducibly" (further information concerning the history of SNOMED CT® may be found in the introductory sections of their article). SNOMED CT® is an ontology-like terminology organized in a number of generic (IS-A) hierarchies, each of which covers a semantic field like Finding, Procedure, Body Structure, Substance, Infective Agent, etc. Terms representing concepts related to non-medical information needed for patients' health records are included, too – they may be found in hierarchies like Environment and Geographical location, Event, or Physical object. Each concept is linked to one or more concepts in other hierarchies by means of semantic relations, and a concept may have more than one parent concept (IS-A relation).

The semantic relations make up a structured definition of the SNOMED concept – as opposed to a textual definition: each defining or qualifying characteristic of the concept is represented by a relation composed of an attribute + a value. Thanks to the impressive number of concepts and the possibility of combining concepts, SNOMED CT® contains the vast majority of the terms and concepts needed for the recording of fundamental information in patients' health records.

Thus, two concepts like *transhepatic portogram* (belonging to the Procedure hierarchy) and *drug-induced dermatosis* (belonging to the Finding hierarchy)1 have several relations: they both have more than one parent concept (IS-A relations) and both point to a concept in the Body Structure hierarchy by means of a semantic relation composed of the attribute SITE + a value chosen among the Body Structure concepts. As for transhepatic portogram, the method used to obtain the portogram is described by pointing to another concept belonging to the Procedure hierarchy: imaging; and particular characteristics of the concept are described by means of two relations: the approach used and the intent of the procedure. Drug-induced dermatosis has two other relations – one reflecting the fact that the condition manifests itself in a change of the structure of a body tissue, and one reflecting the cause of the condition:

*1. transhepatic portogram*
IS-A abdominal angiografi
IS-A venography
IS-A procedure by approach
APPROACH transhepatic approach
METHOD imaging
HAS INTENT diagnostic
PROCEDURE SITE portal vein structure

*2. drug-induced dermatosis*
IS-A skin lesion
IS-A drug-induced lesion
ASSOCIATED MORPHOLOGY morphologically abnormal structure
CAUSATIVE AGENT pharmaceutical/biological product
FINDING SITE skin structure.

The corresponding textual definitions of these two concepts could be:
*1. abdominal angiography imaging of the portal vein for diagnostic purposes performed by means of transhepatic approach*
*2. skin lesion resulting in morphological alterations and caused by a pharmaceutical or biological product.*

In the case of complex concepts involving for instance several conditions of different aethiological origins, a principle known as role grouping is applied to the pattern of the semantic relations between the hierarchical concepts involved (Spackman et al. 2002). Explaining the principles of role grouping would exceed the scope of this article; what should be stressed, however, is that role grouping provides an extremely useful source of information for all project participants in cases where a term may be interpreted in several ways.

The fact that concepts are defined by certain relations and inter-linked in a highly structured way means that once the terminology has been implemented and coordinated with the GEPJ system, it will be possible to extract selected data from a large number of health records for statistic, economic, or even diagnostic purposes. Relations may be patterns for extraction of relevant information associated with a certain concept. It would, for instance, be possible to find out how many patients in region X had their appendix removed in year Y by method Z by setting specific search criteria for the different information categories contained in the health records. Or it would be possible to find out which one of several medication regimes was, within a given period of time, the most successful as a cure against pneumonia caused by a specific bacteria.

## Remarks concerning ontological and epistemological issues

When faced with SNOMED CT® terms in a translation context, it is useful to bear in mind that the terminology is not 100% consistent and that errors or inconsistencies do occur on the word as well as the

system level: although SNOMED CT® may be regarded as a domain ontology, it does not always meet formal ontological requirements. This problem has been pointed out by several researchers dealing with medical informatics and SNOMED CT® modeling (Spackman et al. 2002 & 2004, Bodenreider et al. 2004), and it is one of the issues which need to be dealt with by the coming SNOMED Standards Development Organisation, SDO. (For information on the foundation and organisation of this international organisation, consult http://www.hiww.org/smcs2006/programme.html.)

Indeed, some SNOMED CT® terms represent concepts whose place in the hierarchy is the result of classifications of concepts based on conventions, presumptions or habits rather than on scientific laws. An illustrative example, mentioned by Bodenreider et al (Bodenreider et al. 2004)) in their article on the "ontology-epistemology divide", is Gram negative bacteria versus Gram positive bacteria which have been named, not after some intrinsic characteristics, but after extrinsic characteristics, i.e. according to their reaction when exposed to crystal violet dye. Similar examples of concepts violating true ontological principles are phantom limb syndrome without pain, in which the specification without pain does not indicate a reality of the phantom limb itself, but has been added in order to indicate that the normal presumption, i.e. that the phantom limb syndrome is accompanied by pain, does not apply in this case. And in the case of possible thrombus, which is, in the ontological sense, a non-existing entity, the modifier refers implicitely to the criteria on which a diagnosis of thrombus would be based.

As long as the classification principles or the epistemological background are universal, equivalence may be expected between concepts in English and Danish. Indeed, in most cases, the ontology-epistemology issue does not represent a problem to the translation as such. However, misplacement of concepts does occur and may complicate the translation. For example, the concept fairly heavy drinker (Finding) can be found in the sub-hierarchy Finding relating to alcohol drinking behavior (Finding). If one considers the semantics of the term fairly heavy drinker, it becomes clear that the principle of top-to-bottom IS-A relations has been violated: the English term ought to have been fairly heavy drinking habits (Finding). In cases like this one, there is no doubt that the translation should respect the principle of concept based translation rather than term based translation. However, in less evident cases, where the actual meaning of the term cannot be detected by consulting the concept relation patterns, the general principle adopted within the SUNDTERM project has been to respect the English term while making a note of the problem. Corrections and improvements of the structure of the SNOMED terminology will be a continuous process, and the Danish terminology will have to be updated accordingly.

## GETTING THE SUNDTERM PROJECT STARTED

### Basic principles

The basic approach of the translation project is pragmatic-functionalist: the aim of the project, i.e. establishing a terminology functioning as a background system for the electronic health record is constantly kept in mind, and an effort is made to provide terms which reflect the underlying concepts and are understandable and psychologically acceptable to the clinician.

The overall approach has been one of close collaboration between specialists within medicine and/or informatics, and linguists/terminologists. As pointed out by numerous professionals and terminologists, interdisciplinary collaboration is crucial in terminology work (Infoterm 2005). On the one hand, a translation based solely on linguistic, morphological- syntactical analysis might result in a seemingly correct term which would not after all represent the concept in question, or which would not be used by professionals. On the other hand, for pedagogical (and normative) reasons, a certain compliance with linguistic, systematic, and orthographic principles is necessary in order to avoid confusion and ensure practical applicability of the terminology. In practice, this means that a set of basic principles to secure consistency are followed, but in case of serious conflicts with daily clinical language the clinical use prevails.

### Initial studies and organisation

During 2003, the Danish National Board of Health (SST) carried out a series of studies of SNOMED CT®. Eventually, it was deemed applicable as a basis for the above mentioned GEPJ, which is developed in

preparation for exchange of clinical information in Denmark. A trial translation of about 40,000 terms, followed by clinical testing, was carried out during 2004, and in the spring of 2005, the SUNDTERM Project as such was adopted and launched.

The SUNDTERM translation project has been divided into two phases. Phase I comprises translation of the terms representing the concepts and initial approval of terms in order to ensure that the terms are understandable and possibly also acceptable to the clinicians. Phase II, which has been initiated in 2006, consists in reviews by clinicians, including the addition of synonyms as well as clinical testing of the terminology.

Contracts were signed with a translation bureau (Intertext), work flows were defined, and a web-based IT system to support work flow processes was developed by a software company (CareCom A/S). All specialists involved in the translation process participated in a seminar in which the SNOMED CT® structure as well as the aims, principles, methods, translation tools, work flows, timetables, etc. of the project, were presented and debated.

An Editorial Board consisting of about 10 persons was formed; apart from an advisory and quality management function, the main tasks of the Editorial Board comprise the updating of documents containing linguistic/terminological guidelines for the translators and the SST reviewers as well as suggesting solutions for particular translation problems. The board is composed of medical doctors and other health professionals, some of whom have a masters' degree in health informatics, as well as professionals with an educational background in LSP (Language for Specific Purposes) translation, linguistics, and terminology. The Editorial Board meets at regular intervals; the majority of its members are SST employees some are consultants and two represent the translation bureau.

During its first six months of Phase I, the Editorial Board carried out discussions with various specialists in order to refine and revise the general guidelines and to ensure its awareness of current tendencies regarding e.g. information search techniques and computational linguistic strategies, as well as of principles applied in other projects involving medical terminology. These meetings included a presentation of term extraction tools by ass. prof. Lotte Weilgaard from the University of Southern Denmark, consultations with prof. Bodil Nistrup Madsen, DANTERMcentret and Copenhagen Business School, terminology consultant of the National Health Terminology Committee, and discussions regarding syntax and morphology with MD, lic.scient. Søren Nørby, editor of the latest edition of the Danish cyclopedic medical dictionary (Klinisk ordbog), and Jørgen Schack, researcher at the Danish Language Council.

### Organising the translation process

A special web-application with restricted access was elaborated by the Danish software company CareCom A/S. The basic ideas behind the application were that

- all participants were to work on a common web-based platform
- each concept was to pass through a work flow – from translation via review to approval
- all authorized users should have access to all information needed about each concept as well as to various sources provided by the SST (i.e. general guidelines and translation suggestions provided by the Editorial Board, a number of textbooks in electronic form, and useful references of high validity).

The people involved in the SUNDTERM translation process, i.e. translators, SST staff, and consultants, were assigned log-ins and access authority according to their role in the translation process: translators may translate only, not validate the terms, and reviewers cannot change suggested translation, but must either accept the term or send it to the Editorial Board with a remark/question/suggested change. In this way the Editorial Board functions as the ultimate instance of review.

## GUIDELINES AND IMPORTANT ISSUES REGARDING THE TRANSLATION OF SNOMED CT®

As mentioned above, a number of guidelines, or rules, were set up before the translation process was initiated. Apart from the effort made to ensure a certain level of consciousness amongst the participants

regarding the fundamental principles of terminology work, this included the formulation of a number of (linguistic) rules of syntax, morphology, and orthography to which everybody could refer and with which they were to comply. The message regarding general principles of terminology work was passed on to the participants at a seminar/workshop in the spring of 2005, and the linguistic guidelines, already established in connection with the pilot project in 2004, were reviewed and updated.

## Terminological principles

Depending on the task in question, certain generally accepted term requirements must be complied with in order to ensure terms of high quality. In the present project , interdisciplinary cooperation and reviews as well as the need to meet specific term requirements were particularly important. Term requirements at issue in medical LSP were thoroughly discussed in my PhD dissertation of 1998 (Høy 1998); for the SUNDTERM translation project, the following major requirements were adopted:

- **unambiguity** (a term having the status of "preferred term" must not refer to more than one concept in the hierarchy in question)
- **linguistic correctness** (national syntactic and orthographic rules must be complied with)
- **motivation** (immediately understandable terms, i.e. terms reflecting the characteristics of the underlying concept, should be preferred)
- **international recognizability** (terms based on Latin and Greek word elements should be preferred)
- **psychological acceptability** (clinicians' habits should be taken into account whenever possible)
- **systematism & consistency** (similar morphological and syntactical solutions should be sought for terms covering semantically similar concepts).

Unfortunately, these requirements will often be in conflict. Psychological acceptability tends to be an obstacle to compliance with several other principles:
Commonly used and accepted eponyms such as Apgar score or Down syndrome are at odds with the wish for motivation; the general clinical preference for Danish (or even English!) terms is in conflict with international recognizability or linguistic correctness; and random habits of using noun+noun and adjective+noun constructions respectively for terms challenge systematicity & consistency (such as "hjertelyd" = heart sound as opposed to "kardiel revaskularisering" = heart revascularization).

For these reasons, not all SNOMED terms translated can be expected to meet all term requirements. In fact, changing a term because it does not meet one requirement may result in inconsistency in respect of another requirement. Therefore, the Editorial Board has a difficult task trying to comply with the principle of making general as well as individual decisions which are justifiable not only from a terminological and linguistic, but also from a psychological and clinical point of view.

## Linguistic principles

Because of the inevitably normative nature of the Danish terminology, defining a set of linguistic guidelines, including syntactic, morphological, and orthographic rules, was also crucial. On a term status level, it was decided to give preference to so-called hybrids (i.e. terms based on Greek and Latin morphemes but adapted to Danish) which are widely used by clinicians, but to accept the use of purely Greek or Latin terms for certain diagnoses, and in particular for anatomical concepts; purely Danish and English terms are to be accepted only when an alternative choice would be incompatible with current clinical use. On a term corpus level, recommendations as defined by the Danish Language Council are generally followed, but a number of specific morphological and orthographic rules were specified for the construction of hybrids, since no rules pre-existed in this area and many different orthographic variants could be found in the medical literature.

In short, the guidelines laid down include directions regarding:

- the level of validity ascribed to already approved sources of information (national special dictionaries, textbooks, specific homepages, etc.)
- the extent to which foreign (mainly English) terms may be accepted as such
- the choice of term type – purely Latin/Greek terms, hybrids, or purely Danish terms
- morphological and orthographical principles for the formation of hybrids
- the handling of eponyms, abbreviations, formula, symbols, punctuation, etc.
- orthographical and syntactical solutions in cases in which several solutions may be correct
- problems originating in the multi- and poly-hierarchical, ontology-like structure of SNOMED CT©.

## Translation techniques

The translation of complex terms presented to the translator in an extra-textual setting requires a rather high level of domain insight on the part of the translator in order to ensure that any target language word or phrase chosen be an exact representation of the source language concept. On several occasions, the attention of the SUNDTERM project participants has been drawn to the necessity of considering the meaning of the term, i.e. of applying a concept based translation approach and, in case of any doubt, of consulting the hierarchical position and the semantic relations of the concept in question. Within a functionalist translation approach, various techniques may be applied as described by Molina et al (Molina et al. 2000). To some extent, techniques like borrowing or literal translation may be recommended as long as concept equivalence is ensured: the resulting target language terms are internationally recognizable and often psychologically acceptable to clinicians, and they make it possible to conform with the structure of SNOMED CT®. But several more genuinely functionalist techniques are very often used, for instance transposition, amplification/description, and established equivalents. To a large extent, the translation methodology and solutions resemble the ones in prominence in connection with the translation into Spanish of SNOMED CT® as described by Reynoso et al (Reynoso et al. 2000).

The main translation techniques applied are represented in the table below.

| **Translation technique** | **English term** | **Danish term** |
|---|---|---|
| borrowing | *cardiac output* | *cardiac output* |
| calque or literal translation | *closed fracture of metacarpal bone* | *lukket fraktur af metakarpalknogle* |
| transposition (change of grammatical category) | *disability affecting daily living* | *funktionsnedsættelse der påvirker daglig livsførelse* |
| amplification/ description | *battered wife* | *hustru der har været udsat for vold i hjemmet* |
| established equivalent | *product of conception* | *befrugtet æg* |

## Equivalence and false friends

As mentioned earlier, questions related to terminology structure rarely play an important role in translation as such, since most ontological and epistemological patterns are similar in the source and target language. The important criterion of successful communication, as pointed out by Reynoso et al (Renynoso et al. 2000) in their article about the translation of SNOMED CT® into Spanish, will be met: "the same symbol should evoke similar concepts in different people's minds". In most cases, it is possible – with the appropriate linguistic and professional insight – to find a target language term which would match the source language term. Thus, in principle the translation of the above mentioned Gram negative bacteria and Gram positive bacteria, phantom limb syndrome without pain, and possible thrombus should not be too problematic.

However, sometimes the epistemological issue is of a social or cultural nature, and in such cases, the translation becomes difficult because of the lack of concept equivalence: Some IS-A relations of SNOMED CT® represent an arbitrary or national (American) conception of the world which may not necessarily be acceptable on a worldwide scale. For example, the concept person in the healthcare environment, which is the top-level concept (parent concept) of an entire SNOMED CT® sub-hierarchy, encompasses a vast number of – American – job titles. Although it would normally be possible to obtain a reasonably

motivated Danish term for each concept, the IS-A relations between these concepts are not based on a truly ontological principle. From an epistemological perspective, the term-concept-entity relationship is not static, and the translated Danish term with its underlying concept may be completely useless because of socially or professionally based cultural differences.

Even concepts which are seemingly equivalent may not be so after all: the English term typhus and the Danish term tyfus are so-called "false friends"; the correct translations would be English typhus = Danish plettyfus and English typhoid fever = Danish tyfus. Similar problems arise when a term covers a broader concept in one language than in the other. This is the case with the English words substance and drug which are constituent parts of a number of terms: translating substance into substans in Danish would in some cases be OK, in others it would be nonsense. As for drug, a holonym with a broad semantic coverage in English, there is no similar word nor concept in Danish; the solution is to find the meronym in Danish which may best represent the concept in each individual case.

## Concluding remarks

Apart from their necessity to the electronic health record from a data modelling point of view, the IS-A and the semantic relations may sometimes be the only means of detecting the concept behind a given term. Therefore, they are extremely valuable sources of information to the translators, reviewers, and editors in the translation process. Very often, the translator's uncertainty originates in ontological or epistemological inconsistencies or in the lack of concept equivalence as a result of cultural differences. This does of course present a problem, but within the SUNDTERM project framework, the major concern is to provide a Danish term for each SNOMED CT® concept. The systematic or ontological improvements which must inevitably be made to the structure of the core version of SNOMED CT® will have to be applied to the national version later. Anyhow, any national version of SNOMED CT® will need to be adapted to local conditions: deletions, changes, and additions of concepts are inevitable. Therefore, all terms are currently translated notwithstanding the "would-never-be-used" status of the underlying concepts in a Danish setting.

## THE SUNDTERM TRANSLATION WORK FLOW

The work flow may be described as follows:

1. A so-called batch of SNOMED CT® concepts to be translated is selected by SST. Each batch forms a subset containing approximately 10,000 concepts which belong to a certain medical specialty. The translation bureau is granted a period of one month for translating a batch.

2. The translation bureau distributes the terms among 5-6 translators who have specialized in the translation of medical LSP.

3. After translation, the term is reviewed by another translator. At this stage, the term may be returned to the translator for further investigation or it may be accepted. The reviewer may also choose to consult a subject matter expert (SME). If no solution is found and/or the concept or the term present a particular problem, the reviewer may send it directly to the Editorial Board.
4. A term accepted by the reviewer of the translation bureau is taken over by a SST reviewer. These reviewers all have a medical or paramedical background and/or linguistic training. They may accept a term, in which case it is "approved", or they may return it to the translation bureau if they deem it incorrect or not compliant with the linguistic/terminological guidelines, or they may pass it on to the Editorial Board for further verification or investigation.

5. The Editorial Board may either accept the term, find a solution, or return it to the translation bureau – in case it is evident that the term does not cover the meaning of the concept or if the guidelines or work flow have not been complied with.

Fig. 1 below shows the work flow chart; the minimum cycle for any term is shown by means of bold-faced arrows:

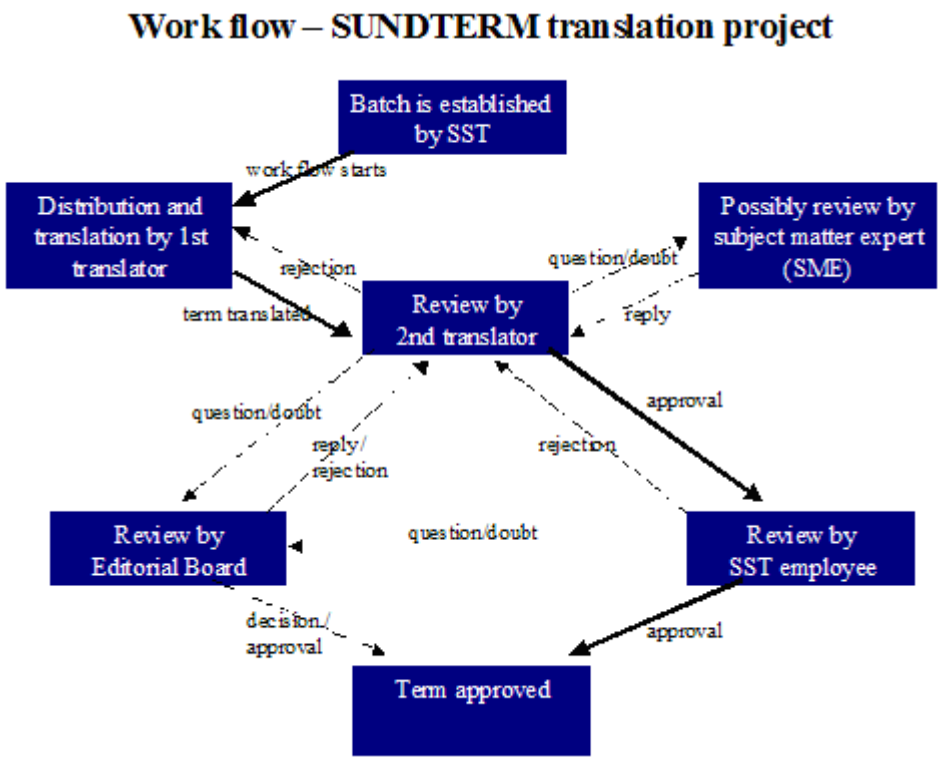


Fig. 1

Fig. 2 shows an example of the user interface in the case of a concept having worked its way through the work flow:

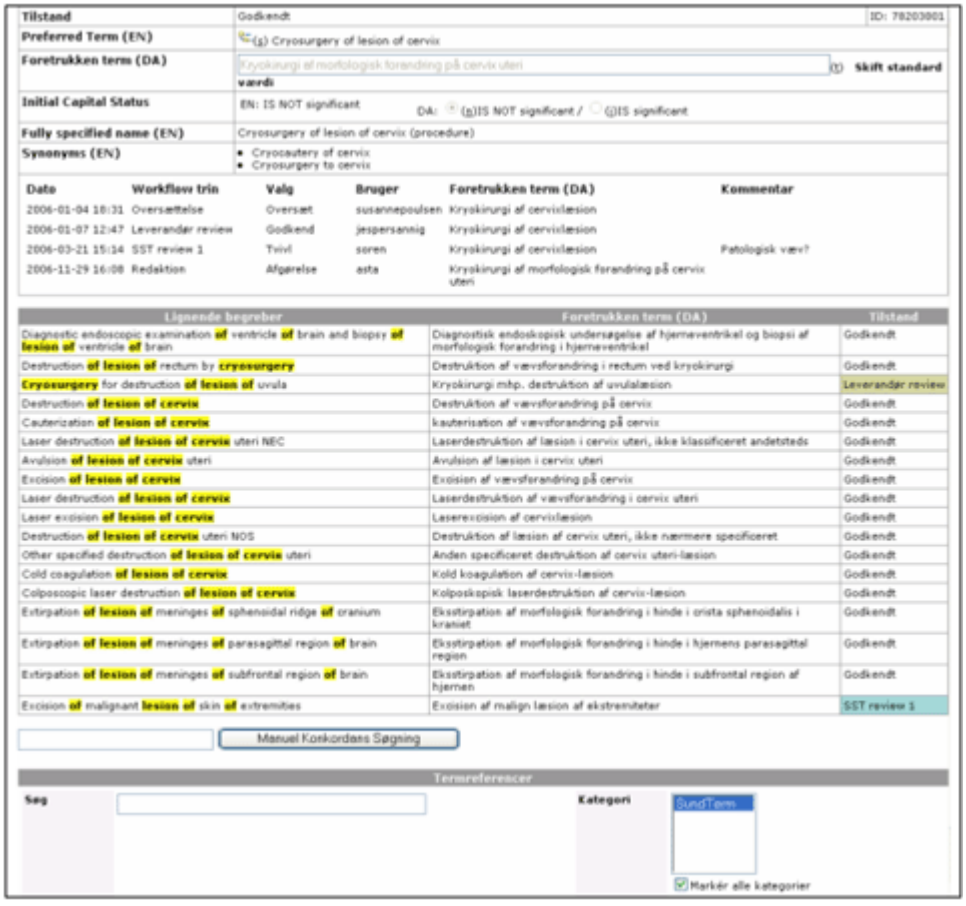

Fig.2

The original translation of this term was approved by the local reviewer and sent on to the SST. The SST reviewer made a comment on the compound ”cervixlæsion” and sent the term on to the Editorial Board. This led to a discussion of principles regarding the translation of lesion in the sense of tissue alterations not resulting from injuries, and the term was finally changed and approved. At the same time, guidelines regarding the translation of the English term lesion were registered in a document of decisions of principle of the Editorial Board. This document is accessible to all those who have access to the application.

## CONCLUSION

Translating SNOMED CT® into another language than English is a time consuming, resource consuming, and complex task which demands close collaboration between skilled specialists within medicine, linguistics, terminology, information technology, and project management. Preparatory work, introducing all participants to the basic structure of SNOMED CT®, elaboration of efficient and transparent tools, definition of roles and work flow, as well as specific terminological and linguistic recommendations and rules, are important quality criteria.

In spite of all the efforts, it cannot be expected for a national version of SNOMED CT® to become a totally consistent terminology. On the one hand, inherent ontological inconsistencies in SNOMED CT® will inevitably be transferred, and on the other hand, national habits, needs, and wishes in regard to a medical terminology and its use make it necessary to reach compromises on the terminological-linguistic level.

With the SNOMED Standards Development Organisation (SDO) which is being founded now in late 2006, it may be expected that considerable structural improvements will be made to the terminology over the coming years. Yet many of the issues described and discussed in the present article will remain questions and challenges which have to be overcome in a translation context. Hopefully, the experience derived from the Danish translation project, still in progress, may be of help to those who consider translating the terminology into their own national language.

[1] Most term examples in this article have been gleaned from the July 2005 version of SNOMED CT®, so that some of them may not occur in the present version.

BIBLIOGRAPHY

BODENREIDER, O., SMITH, B., BURGUN, A. (2004). The Ontology-Epistemology Divide: A Case Study in Medical Terminology. In: Achille Varzi and Laure Vieu (eds.), Proceedings of FOIS 2004, International Conference on Formal Ontology and Information Systems, Turin, 4-6 November 2004 (http://ontology.buffalo.edu/medo/Onto_Epist.pdf).

HØY, A. (1998). Det medicinske fagsprogs terminologi - en patient med behov for behandling? (PhD dissertation), Jelling.

INFOTERM (2005). Guidelines for Terminology Policies, UNESCO, Paris (CI-2005/WS4).

MOLINA, L., ALBIR, A.H. (2000). Translation Techniques Revisited: A Dynamic and Functionalist Approach, Meta, XLVII(4), 497-512.

REYNOSO, G.A., MARCH, A.D., BERRA, C.M., Strobietto RT, Barani M, Iubatti M et al (2000). Development of the Spanish Version of the Systematized Nomenclature of medicine: Methodology and Main Issues, Prc AMIA Symp; 694-8.

SPACKMAN, K.A., DIONNE, R., MAYS, E., WEIS, J. (2002). Role grouping as an extension to the description logic of Ontylog, motivated by concept modeling in SNOMED, Proc AMIA Symp., 712-6.

SPACKMAN, K.A., REYNOSO, G. (2004). Examining SNOMED from the perspective of formal ontological principles: Some preliminary analysis and observations. In: Hahn U (ed), Proceedings of the KR 2004 Workshop on Formal Biomedical Knowledge Representation. Whistler, BC, Canada, 72-80.

The Ministery of the Interior and Health (2003). National IT Strategy 2003-2007 for the Danish Health Care Service, The Ministery of the Interior and Health, 46 (http://www.sst.dk/publ/Publ2004/National_IT_strategy.pdf).

**Marita Kristiansen**
**Department of Professional and Intercultural Communications**
**Norwegian School of Economics and Business Administration**

# A TERMINOLOGICAL APPROACH TO MULTI-DISCIPLINARY DOMAINS AND DISCIPLINARY AUTONOMY

Abstract

*Today's society is characterised by an increasing specialisation and the emergence of new disciplines. To be able to understand and mediate the content of these areas, it is important to be able to delineate them. Within this setting this article presents the main background and results from my doctoral study (Kristiansen 2004). The study addresses the question of what constitutes a scientific discipline and when is it possible to say that a particular discipline is autonomous. Secondly, the study investigates how such disciplinary autonomy can be determined.*

*The tradition of borrowing concepts as found in many social sciences, results in highly multi-disciplinary subject fields. It has therefore been an aim of the study to investigate the nature of social sciences in particular, with focus on their concepts. A major aim has been to see whether methods of conceptual analysis as described in terminology theory can offer a useful means of evaluating the autonomy status of a given discipline.*

*The point of departure for the conceptual analysis has been terminological methods as described in among others Laurén et al. 1997, which have been developed to enable an investigation of conceptual changes in the social sciences.*

## 1 INTRODUCTION

As stated in Sager (1990: 16), no individual or group of individuals possesses the totality of the knowledge structure of a community, which is why we divide knowledge into disciplines. This does not have to result in easily distinguishable units, but rather in overlapping subject fields. In an emerging discipline the conceptual basis is likely to be muddled and misunderstandings in communication situations may occur. Hence it would be fruitful if the relative autonomy of a discipline could be assessed.

In the social sciences there is a tradition of borrowing concepts, something which results in highly multi-disciplinary subject fields. In my doctoral thesis (Kristiansen 2004), the attempt has been to outline how the origins of one such discipline, namely organisational behaviour (also called OB in specialist texts) can be traced in other, more traditional disciplines, which I have called parent disciplines. These parent disciplines include psychology, sociology, social psychology, anthropology and political science.

OB is a relatively new branch of knowledge, and at present no common view of how it should be delineated seems to exist. In Robbins (1996) it is defined as:

> a field of study that investigates the impact that individuals, groups, and structure have on behavior within organizations, for the purpose of applying such knowledge toward improving an organization's effectiveness (ibid.: 10).

The central unit of study is behaviour, expressed by individuals acting on their own or in interaction with others in groups. The discipline studies how this behaviour influences the performance of the organisation. An organisation may be defined as a "consciously co-ordinated social unit, composed of two or more people, that functions on a relatively continuous basis to achieve a common goal or set of goals" (op. cit.: 5). The preoccupation with performance and the need to control this performance are what distinguish organisations from other forms of social arrangements.

Thematically, OB is frequently divided into four subareas, (Northcraft and Neale 1994; Robbins 1996): individual behaviour, group behaviour, organisational dynamics and organisation system. The satellite

system in Figure 1 shows 'organisational behaviour' as the parent node, surrounded by four sister nodes which correspond to the OB subareas.

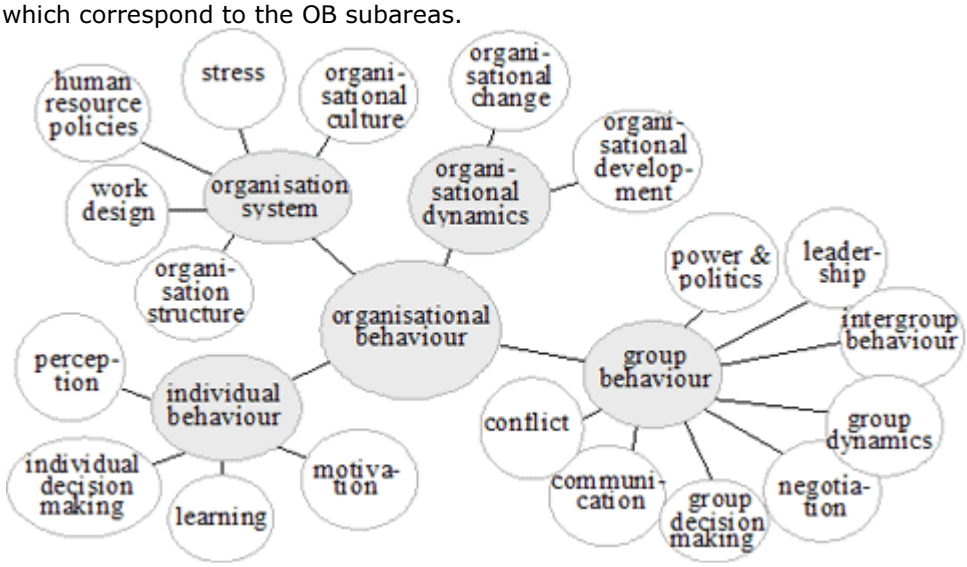


**Figure 1.** A satellite system of the OB subareas

The subareas of the discipline have been further subdivided into smaller areas (or subordinate concepts), here rendered as the smallest nodes. This division of OB subareas will be the point of departure for the investigation of conceptual changes in OB and the assessment of the autonomy status of OB.

Since behaviour is the most central unit of research of OB, it has been natural for OB researchers to search for the answers in theories explaining human conduct. This may explain the tendency in OB to borrow concepts from other social sciences, i.e., the parent disciplines. The fact that a discipline is considered to be a parent discipline of OB does not imply that OB has originated directly in that discipline, but that OB to a great extent has borrowed from the discipline and incorporated the borrowed concepts in its own conceptual apparatus.

All the parent disciplines belong to the group of social sciences and have thus human behaviour as the main object of their studies. However, the various parent disciplines differ in what characteristics of behaviours they study and in what settings the behaviours take place. For instance, behaviour may be studied as either individual behaviour or group behaviour. Furthermore, it is possible to study how group behaviour may influence individual behaviour. Finally, behaviour may be analysed in a political environment or a cultural context.

In my study the attempt have been to take into account these various angles from which behaviour may be viewed, including psychology, whose influence on OB may be traced back to the 1950s, followed by social psychology, sociology, anthropology and, finally, political science, which is the discipline that has most recently begun to assert its influence on OB, dating back to the 1970s.

The idea to the project emerged as a result of a terminological analysis, of motivation concepts in OB, which was the focus of my cand.merc thesis (Kristiansen 1997). When analysing OB concepts in this previous study, it soon became evident that I was dealing with a domain which was highly multi-disciplinary in nature. In fact, several disciplines seemed to have offered theories and concepts to the foundation of OB. In addition, OB appeared to be closely related to several other disciplines such as management, economics and strategy.

Within this setting, two major questions have set the frame for my study. The first one relates to the concept of disciplinary autonomy: When is it possible to say that a particular discipline is autonomous? The second question involves LSP research and in particular terminological analysis: Can a conceptual analysis of the concepts of a discipline be used as a means to assess the autonomy status of that discipline?

To find an answer to the first question of disciplinary autonomy, philosophy of science has been taken as a point of departure, in which general criteria for scientific activities as discussed by among others Kuhn (1970), have been addressed to delineate a set of characteristics typical for what Kuhn describes as the

qualitative nature of normal science. Secondly, the concept of scientific concept has been discussed to provide a basis from which to analyse conceptual changes by means of terminological methods. Thirdly, OB has been delineated in terms of the scientific foundations on which it is based, including a delimitation of its parent disciplines and their subareas. Finally, the study comprises an analysis of three central concepts which have been borrowed from the parent disciplines, namely 'motivation', 'group dynamics' and 'group decision making', including their conceptual clusters. The attempt has been to investigate whether the content of the borrowed concepts have changed compared with the content of their respective parent concepts.

In this article I will outline the overall purposes of the thesis. Next, the applied method will be presented, including a discussion of what I understand by the concept of disciplinary autonomy and the relation between the conceptual analysis and the research questions. Finally, the main results will be summed up, and some implications for future research will be presented.

## 2 METHODOLOGY

The study has had two overall purposes, namely to investigate when it is possible to say that a particular discipline is autonomous and, secondly, to see whether conceptual analysis as described in LSP research can be used as a means to assess the autonomy status of a given discipline. These two overall questions have different methodological implications. Whereas the first question is of a more philosophical nature, the second is of a more empirical one. The methodology developed in the study has therefore been two-fold.

### 2.1 How to investigate disciplinary autonomy

To anyone familiar with the social sciences and the high degree of interrelatedness that exists between the various disciplines, it should be obvious that it is difficult to find objective criteria for whether a discipline is autonomous. Hence autonomy must be seen as a relative characteristic of a discipline.

In order to analyse the autonomy status of OB, it has first been necessary to stipulate some criteria that could be used to delineate domains as disciplines and to distinguish between interrelated disciplines. To do so, I have turned to philosophy of science which has been taken as a point of departure. Based mainly on Kuhn's description of the qualitative nature of normal science (1970), it has been possible to determine a set of characteristics that should be present in a domain for it to qualify as a scholarly discipline. These criteria are of both a sociological and an epistemological nature.

In short, sociological characteristics include the existence of research groups who can be identified with a common paradigm or paradigms, associations, common communication channels, and regular events where researchers meet, such as conferences and meetings. In addition, associations, internet sites and university level courses which teach the disciplines will add to the sociology of a discipline. The sociological criteria may be described as being of an external nature since they are what can be observed by outsiders, such as e.g. web sites and journals. Disciplinary analyses based on sociological criteria are described in for instance Jakobsen (1980) and Antia (2001), in which citation analyses have been applied to investigate who cites who in marketing and terminology, respectively.

Epistemological characteristics include a separable research object, separate methods for empirical investigations, an independent theory development and a common conceptual apparatus, including terminology. A previous epistemological study which has investigated the various schools within terminology science is Laurén and Picht 1993. In their article "Vergleich der terminologischen Schulen" (1993), they conclude, among other things, that there are divergent views on how theoretical issues are solved and thus how theories are developed depending on whether the departure point is a linguistic orientation or an inter- and transdisciplinary orientation.

Based on the sociological and epistemological criteria, a working definition of 'disciplinary autonomy' has been established, which reads as follows:

> a discipline's relative independence from other related disciplines, when it comes to both "sociological" and "epistemological" disciplinary characteristics

Although investigations based on sociological criteria certainly indicate the existence of autonomous disciplines, they say little about how the knowledge presented as a discipline differs from other related disciplines, and therefore whether we in fact may argue that the discipline is autonomous. Especially when disciplines are interdisciplinary by nature, sociological investigation may not disclose to what extent a discipline present knowledge that can be considered as an autonomous whole, independent of other related disciplines, relatively speaking. To add to the confusion, identical terms may be found across disciplinary borders. Whether the terms represent the same knowledge in the related disciplines will therefore have to be investigated more closely. And this I believe can be investigated by conceptual analysis, since the concepts of a discipline will constitute the central knowledge units in that discipline.

There have been several attempts to reconstruct the scientific foundation of disciplines by means of other methods than conceptual analysis. A widely used method is citation analysis as mentioned above (see also Garfield 1979; Jakobsen and Grønhaug 1993). Such analyses may be applied to construct a historical mapping of a discipline to investigate how it is related to other disciplines and how such relations may change during a period of time. A conceptual analysis, however, will result in a description of very central epistemological characteristics of a discipline since the concepts of the discipline will constitute the core of its knowledge.

## 2.2 Terminological analysis

In the analysis I have focused mainly on the last epistemological criterion of a discipline, namely that a discipline should have a common conceptual apparatus, including terminology. Methods of conceptual analysis have been applied, as developed in terminology theory, to compare OB concepts with their corresponding parent concepts. This analysis will provide an answer to the second overall question of whether terminological methods can be a useful means to delineate disciplines. Such a comparison will also disclose whether any changes have taken place in the adapted OB concepts or not, and also whether new concepts have been formed. The belief has therefore been that the analysis will provide an answer to whether OB is an autonomous discipline.

Through the conceptual analysis the attempt has been to find an answer to three more specific research questions:

- What are the scientific foundations of OB?
- Which intrinsic changes have taken place in the borrowed concepts?
- Which extrinsic changes have taken place in the borrowed concepts?

By the scientific foundation I understand the parent disciplines from which OB seems to have borrowed concepts. At this stage in the analysis the various parent disciplines have been seen as concepts themselves and analysed based on how they are defined and delineated by their subject specialists. The motivation for raising this first question has been the belief that the parent disciplines have influenced OB and its subareas by lending their concepts to the discipline.

The first superordinate level of the analysis has focused on the influence the parent disciplines have had on OB when it comes to the scope and purpose of research, indicated by the first two levels in Figure 2.

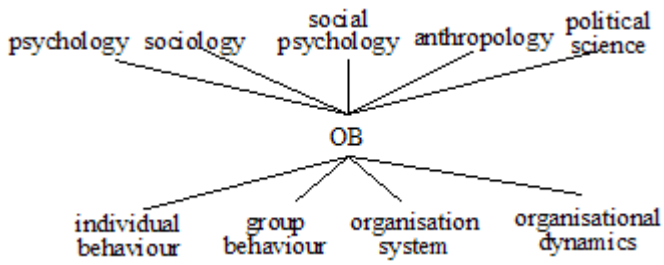


**Figure 2.** Superordinate and subordinate levels of analysis

The second, subordinate level of the conceptual analysis has focused on the influence the parent discipline concepts have had on the four subareas of OB, indicated by the middle and lowest level in Figure 2. This stage of the analysis has also involved an analysis of what I have described as intrinsic

changes, i.e. changes in the intensions of the concepts and any term changes that may have taken place. At this stage the selected concepts of 'motivation', 'group', 'group dynamics', and 'group decision making' have been analysed to investigate to what extent any changes in the borrowed concepts have taken place compared with their respective parent concepts. A second aspect has been to investigate what influence such changes may have had on the autonomy status of OB.

In the comparison of OB and its parent disciplines, at least three different sources of conceptual change have been investigated, namely separate concept formation, the borrowing of concepts from parent disciplines, and the production of new concepts by developing borrowed concepts. Similarly, concepts may just as well have been deleted in the process of defining the new, emerging discipline. The point of departure for the analysis has been a taxonomy of conceptual change which has been developed based on Thagard 1992 (Kristiansen 2004: 92ff)

The investigation of possible intrinsic changes has involved the testing of two hypotheses, which have turned out to be valid, although for different groups of concepts: For some subareas and concepts it seemed likely that conceptual changes have taken place in the intension of the concepts as they have been borrowed from the parent disciplines and adapted in OB. However, my pilot studies have indicated that for other concepts it might be the case that the concepts have been borrowed without significant changes being made to their intensions, i.e., most of the essential characteristics of the parent disciplines' concepts have been retained.

The third and final stage of the analysis has focused on the relations that exist between the various OB concepts compared with those of the corresponding concepts in the parent disciplines. This analysis has been motivated by the belief that conceptual changes take place in the conceptual extensions as they are borrowed from the parent disciplines and adapted in OB. The purpose of this part of the analysis has therefore been to find an answer to the third research question of whether any extrinsic changes have taken place in the borrowed concepts. The term **extrinsic changes** thus relates to changes that have taken place "outside" a concept, i.e. in the way a given concept is related to other concepts.

The choice of the terms **intrinsic** and **extrinsic** to describe the two types of conceptual change has not been made without hesitation. At first I wanted to use the terms **intensional** and **extensional**, since these are already used in terminology theory in relation to concept characteristics to distinguish between inherent characteristics and characteristics that are used to indicate conceptual relations. Consequently, relational change has also been considered instead of extrinsic changes. Since I wanted to include possible **term changes** as something that is associated closely with the concept itself, I wanted, however, to apply a term that could take this into account. The use of **intensional** would in that case have been misleading. The long tradition for dichotomies in terminology theory, i.e., classifying things into two opposed parts or subclasses has therefore led me to choose two other terms, namely **intrinsic** and **extrinsic changes**, where intrinsic includes both intensional changes and term changes.

## 2.3 Corpus material

In order to carry out an empirical investigation of the exchange of concepts between the parent disciplines and OB, specialist texts have been used as corpora. The corpus material has consisted of textbooks on OB and its parents disciplines aimed at students at university level. In addition, the material has been supplemented with research articles which discuss the concepts that have been analysed.

For a textbook to be useful to students who do not have a comprehensive overview of all the discussions and controversies currently taking place within the various research milieus, substantial selection and exclusion of the theories and concepts of a discipline are necessary. Thus it has been assumed in the present study that textbooks generally describe theories and relations among concepts which are no longer disputable. The conceptual descriptions given in the selected textbooks have therefore been taken as a valid delineation of the concepts in question. In contrast, specialist journals issue for instance research articles on recent developments of OB subjects and will therefore introduce new concepts that have emerged, and which are not described in the textbooks yet. Research articles thus add a dynamic dimension to the analysis since such articles very often represent new research in a field. They may also comprise controversies and areas which may be highly debated and even refuted in specialist milieus. The addition of articles has therefore been made to provide a more detailed description of concepts than what is provided in the textbooks.

In addition to the written corpora texts, discipline specialists have been consulted for a professional evaluation of the results (Kaufmann 2004, personal communication). This evaluation may be seen as representing an "extended" corpus since it has provided additional evidence which could not be immediately extracted from the written text corpus.

# 3 MAIN RESULTS

## 3.1 The scientific foundations of OB

To assess the scientific foundations of OB, the five parent disciplines – psychology, sociology, social psychology, anthropology and political science – have been delineated and compared with OB, and the subareas of the various disciplines have been investigated.

Based on the conceptual analysis, the selected parent disciplines and OB may be described as a family of disciplines, with characteristics which in many instances are very similar, both when OB and the parent disciplines are compared, but also when the parent disciplines are compared with each other. It is evident that the scope of research in the disciplines is very similar, human behaviour being one vital common characteristic. This is supported in the corpus material, for instance Northcraft and Neale (1994), in which it is argued that:

> The study of people's behavior in organizations – organizational behavior – is one of a family of scientific areas of study known as the behavioral sciences. The behavioral sciences all share both a scientific orientation and a focus on human behavior as the object of their study (ibid: 18).

The defining characteristic of OB as a discipline which focuses on 'behaviour within organisations', as opposed to for example 'any behaviour', however, indicates a special level of research for the discipline. The comparison of OB and psychology may be used as an example.

OB can be described as:

> **(1) organisational behaviour (OB)**
> A field of study that investigates the impact that individuals, groups, and structure have on behavior within organizations, for the purpose of applying such knowledge toward improving an organization's effectiveness (Robbins 1996: 10).

Based on the above, the research object of OB is human behaviour, including both individual and group behaviour, in addition to organisational structure. The overall research question is: What impact do individuals, group and structure have on behaviour in an organisation? Finally, the purpose of posing such a question is to improve the effectiveness of an organisation.

Psychology is the oldest parent discipline, usually considered to be dating back to 1879, the year that W. Wundt established the first formal psychology research laboratory (Bernstein et al. 1997: 6). However, its roots can be traced through centuries of history in philosophy and science. A common way of describing the discipline is to see it as:

> **(2) psychology (P)**
> […] the scientific study of behavior and its causes. Behavior is used in its broadest sense to include anything that a human or animal can do, including both observable behavior and inner mental and physiological processes […] The basic goals of psychological research and applications are to describe, understand (explain), predict and control (influence) behavior (Passer and Smith 2001: 36).

This involves measuring, explaining and sometimes changing the behaviour of individuals. Literally, psychology means the "study of the mind". A central issue is the relationship of mind and body, due to its derivation from philosophy and physiology. The explanation of psychology indicates that the discipline has both human and animal behaviour as its research objects. The research questions are centred round concepts such as 'learning', 'cognition', 'intelligence', 'motivation', 'emotion', 'perception' and

'personality', which are all concepts that explain individual behaviour, with the purpose of describing, understanding, predicting and controlling behaviour and mental processes (ibid.).

When comparing the scope of psychology with that of OB it has become evident that there is a strong overlap between the two disciplines. In fact, it may be argued that there is a generic relationship between them, where psychology is hierarchically above OB. The fact that the scope of psychology is much wider than that of OB, in that it includes all human beings (and even animals) and not only people in organisations, supports this view. So does the fact that the research questions of psychology are more general, and do not specify a special focus on structural influence, which OB does, as well as the characteristic "organisational structure influence" as a classifying characteristic of OB. Finally, the emphasis on goal achievement represents a delimiting characteristic of OB. However, the presence of other parent disciplines in the conceptual apparatus of OB indicates that the relationship between psychology and OB must be of a polyhierarchical nature.

Although OB and psychology have highly overlapping research interests, the purpose of OB is, nevertheless, more restricted than what is the case for psychology. Thus the scope of research is only corresponding in some areas. For instance, OB does not only study individual behaviour, but also the interaction among individuals and groups. Furthermore, its knowledge base includes theory on the whole organisation system. These areas do not seem to be very central in psychology.

As for family members, OB and the parent disciplines are not identical, and the degree of relationship varies among the disciplines. Perhaps the idea of family resemblance can be used to illustrate this (Wittgenstein 1963). Similarly to the members of a family, OB is more closely related to some than others. Based on the findings of the conceptual analysis, psychology and social psychology stand out as being the parent disciplines with which OB has the closest ties. The analysis indicates that a further analysis of the concepts found in the three disciplines will show a very close conceptual interrelationship. If describing the degree of similarity along a scale, the third most similar parent discipline is sociology. Anthropology and political science will have to be positioned further apart from OB in respect of similarity, political science being the parent discipline with least resemblance.

Furthermore, the delineation of the parent disciplines shows that psychology has mainly contributed to knowledge of individual behaviour, whereas the other disciplines have offered knowledge of group, organisational, institutional, cultural and political behaviour in OB. In other words, the parent disciplinescan be said to have contributed to the micro and macro level of analysis, respectively. This is illustrated in Figure 3 below, which indicates the conceptual links between the parent disciplines and OB.

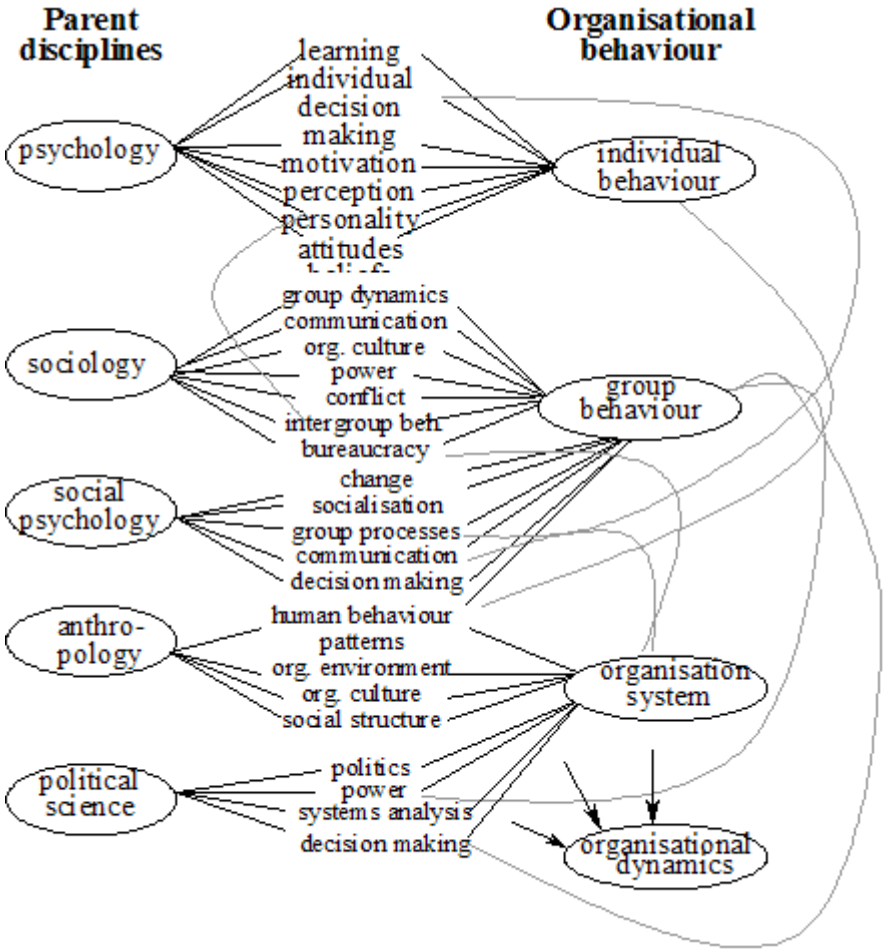


Figure 3. The interrelatedness between social sciences

The four elliptic shapes to the right-hand side in Figure 3 represent the various OB subareas. The concepts connecting the parent disciplines and the OB subareas, indicated with straight lines, are concepts which are shared between the disciplines. This survey does not, however, show all concepts which OB has borrowed, the total number is much higher. The motivation for the selection of concepts presented in the figure has been to show those which are the most relevant to OB, and which were to be analysed in the thesis. To an anthropologist, for instance, obvious relations will therefore be missing in the figure. Furthermore, the relations between the disciplines and concepts are much more complicated than indicated in the figure, which is illustrated by the curved lines. The true relationships between the disciplines and their concepts are rather of a multi-dimensional nature, which results in a complicated network of similarities overlapping and criss-crossing to use the words of Wittgenstein (ibid: 32, English version).

The analysis also shows that the parent disciplines are highly interrelated. The three concepts of 'perception', 'attitudes' and 'beliefs', for instance, are studied in psychology. These concepts are also studied indirectly in sociology and social psychology. In these two disciplines the focus will, however, be on how social factors and social groups, respectively, influence individuals. Naturally, such interdisciplinary relations complicate the analysis greatly.

To conclude, the analysis has shown that OB has at least five parent disciplines from which it has borrowed concepts, as illustrated in Figure 3. Consequently, these disciplines have influenced the OB knowledge base. Even though it is possible to outline how these parent disciplines have influenced OB, it is important to remember that few clear-cut distinctions can be made between the various traditional social sciences. Nevertheless, a number of concepts which have been borrowed can be outlined: 'motivation' and 'motivation theory', 'group dynamics' and finally 'group decision making', including the concepts of 'groupthink' and 'group polarisation'. These concepts have formed the point of departure for the investigation of the autonomous status of OB.

## 3.2 Intrinsic conceptual changes in OB

A common characteristic of many of the analysed concepts is that they have been borrowed from the parent disciplines with few or no changes being made to their intensions, i.e., intrinsic conceptual changes cannot be said to be a prominent characteristic of OB. The conceptual analysis has also shown that a number of concepts not found in the respective parent discipline(s), have been found in OB (Kristiansen 2004: 179, 196, 205).

Thus, the most typical characteristic of the OB concepts is that no intrinsic changes take place in the borrowed concepts. OB concepts such as 'motivation', 'intrinsic motivation', 'group', 'norm', 'role', 'group decision making' and 'group cohesiveness' all share the same, or almost the same intension as the corresponding concept in the parent discipline(s). A general conclusion is therefore that few intensional changes in the form of the addition or deletion of characteristics have been found in the OB concepts.

There is, however, one exception which is prominent in many OB concepts, and that is the additional characteristic in their intensions which positions the concepts in an organisational setting in relation to for instance organisational goals, work groups, or similar delimitations. Such changes, which can be grouped as characteristic additions, may be seen as a generic subordination of the concepts in the parent discipline(s) and not as a sign of an emerging new discipline. However, the analysis has shown that the concepts described in the OB corpus texts are related to each other in a new way and that a new conceptual system of special knowledge has emerged.

The changes that have taken place are to a great extent characteristics that reflect the scope of the various disciplines, and which are added to fit the purpose of the discipline in question. The OB concept of 'conflict', for instance, describes possible differences that may take place between "organization members", whereas the sociology concept of 'conflict' involves differences that include any argument between "individuals or groups in society" as such.

> **(3) conflict (OB)**
> Differences among the perceptions, beliefs, and goals of organization members (Northcraft and Neale 1994: 693).
> **(4) conflict (S)**
> Antagonism between individuals or groups in society (Giddens 1997:581).

In addition, the conceptual analysis has disclosed instances of characteristics being deleted from the description of OB concepts compared with the corresponding concepts in the parent disciplines.

The study has not been designed to analyse the expressions used to denote the OB concepts in any detail; nevertheless, some findings have been found during analysis. In general, the denotations of OB concepts are the same as for the corresponding parent discipline concepts. Conceptual change in the form of term change is therefore not a typical characteristic of the emergence of OB. On the contrary, the same terms are applied in several disciplines for seemingly identical concepts. However, there is a tendency to add a premodifying expression in the denotations of the OB concepts, such as ***organisational*** **decision making,** ***organisational*** **culture,** ***work*** **group,** ***work*** **role**, and ***organisational*** **behaviour**. All these terms refer to concepts which have a more general denotation in their parent disciplines, such as **culture** and **social group** in sociology, or **behaviour** in psychology.

## 3.3 Extrinsic conceptual changes in OB

Since many concepts are very similar in OB and the parent disciplines, the proposed disciplinary autonomy of OB may be questioned. The final stage of the conceptual analysis has therefore been to investigate whether there is anything else that distinguishes the disciplines.

According to the belief that conceptual changes take place in the extension of the borrowed concepts, the answer to this question should be sought in the conceptual structures of the disciplines, i.e., how the concepts are related to each other.

In my thesis I have discussed and outlined several types of conceptual change which have been grouped as system reorganisation (Kristiansen 2004: 95ff). System reorganisation involves a change in the

existing conceptual relations of a discipline or its subareas. The conceptual analysis has identified a number of concepts which have been reorganised in the conceptual systems. An example may be the conceptual change that has taken place in connection with the concepts of 'group polarisation', 'risky shift' and 'cautious shift' in the OB subarea of 'group decision making'.

According to Blackwell (1998: 493), social psychology research in the 1960s suggested that individual group member decisions in a potentially risky situation were, on average, less risky than the final decision of the group as a whole. This was termed the risky shift phenomenon:

> **(5) risky shift (social psychology)**
> The process by which a group's initial average position becomes riskier following group interaction (adapted from Smith and Mackie 000: 592). An early term for group polarization, coined before research indicated polarization also occurs toward conservative positions (Deaux, Dane and Wrightsman 1993: 417).

Subsequent research has indicated that the shift to risk is, in fact, a shift to extremity. Groups shift away from a neutral point beyond the average of the decisions initially favoured by individuals in the groups; in other words, shifts to caution as well as risk occur. This has resulted in the formation of two distinct concepts, i.e., 'risky shift' and 'cautious shift', where formerly only the first was recognised. The theoretical development has also resulted in the birth of a new superordinate concept, namely 'group polarisation'.

> **(6) group polarization (social psychology)**
> The process by which a group's initial average position becomes more extreme following group interaction (Smith and Mackie 2000: 592).

The intension of 'risky shift' (OB) has not changed; however, this concept only describes one of the two tendencies (i.e. 'group polarisation') that may result from group discussions.

> **(7) risky shift (OB)**
> Tendency of a group as a whole and each member to be more willing to accept greater levels of risk after a group discussion than prior to it (Northcraft and Neale 1994: 704).

The other tendency is described in the concept of 'cautious shift' (OB), which together with 'risky shift' make up the extension of the new superordinate concept of 'group polarisation' (OB).

> **(8) cautious shift (OB)**
> Tendency of a group as a whole and each member to be less willing to accept risk after a group discussion than prior to it (Northcraft and Neale 1994: 692).

The concept of 'risky shift' has therefore undergone an intrinsic change (both an intensional change and a term change), whereas 'cautious shift' has emerged as a new concept. The changes have also led to the subordination of the two concepts of 'risky shift' and 'cautious shift', as illustrated in Figure 4 below. These changes have been adopted in OB as they have become accepted knowledge in social psychology and are not changes which have taken place due to the emergence of OB.

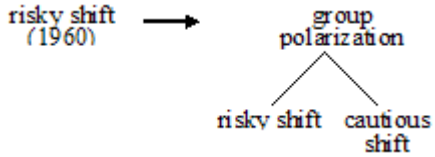


Figure 4. A simple system reorganisation - differentiation

The most prominent relational changes found in OB are the results of the disciplines having different perspectives or aims in their attempts to explain phenomena. Based on the analysis, it may therefore be concluded that a high number of concepts found in the parent disciplines have been borrowed and adapted in the conceptual structures in OB to provide answers that are relevant for an OB researcher who investigates the impact that individuals, groups, and structure have on behavior within organisations, for the purpose of applying such knowledge toward improving an organisation's effectiveness (Robbins 1996: 10).

## 3.4 Disciplinary autonomy in OB

The social sciences, being relatively young disciplines, are characterised by a conceptual apparatus and terminology which are not fully developed. This belief has been supported by the conceptual analysis of OB concepts. The conceptual changes in the textbook corpus show few signs of autonomous theoretical developments in OB. The investigated research articles, however, show that much current research takes place which has already led to conceptual changes and which may in future lead to further changes in OB. This will strengthen the disciplinary autonomy of OB.

Even though the epistemological characteristics of emerging disciplines are less established than those of more mature disciplines, emerging disciplines may still be relatively autonomous. The analysis has shown that this appears to be true for OB, which has developed a separate research object, discipline-specific subareas, concepts and terminology in addition to new conceptual relations and consequent subsystems.

Although the sociological characteristics of OB have not been investigated in the thesis, a number of sources of sociological evidence which indicate the existence of OB have been applied in the conceptual analysis to provide conceptual descriptions. Such sources of evidence include the corpus material as well as several web sites of discipline-related discussions. There are also a number of conference and organisation sites which show that OB communities have been established. These sources support the epistemological analysis.

Taking into account the relatively short life cycle of OB, its close relations to the group of administrative disciplines, it is difficult to see how OB will continue to develop. However, based on the criteria established in this study and the conceptual analysis which has been based on these criteria, I have concluded that OB can be described as an autonomous discipline.

## 3.5 Terminological methods as a means to assess disciplinary autonomy

The terminological analysis has been able to provide few absolute "truths" about the intensions and extensions of OB concepts or the concepts of the parent disciplines. However, what the analysis has done is to offer valuable insight into the interrelatedness of social sciences and their concepts. The use of terminological methods has demonstrated that concepts are relatively vaguely described in textbooks. This lies in the nature of the disciplines; however, it may also be a motivation for a subject specialist to consider whether more specific descriptions should be given to avoid confusion.

Conceptual analysis has traditionally been used to analyse concepts with their present content and in their local environment. What this study has demonstrated is how a conceptual analysis can be applied for the purpose of carrying out a comparative analysis of scientific constituents in parent disciplines. Furthermore, by allowing a time dimension in the analysis, the investigation has been able to illustrate how the application of terminological methods can be extended to include an analysis of how a discipline has developed.

The analysis has also demonstrated that concepts have a life cycle and that concepts can be analysed as dynamic elements and not only as static constructs. However, in the conceptual analysis, the actual investigation has required an analysis of concepts as static elements at a given time in the sense that it has been based on written texts published at a specific time. The dynamic aspect of the analysis has been confined to the investigation of how the concepts have developed from being concepts in the parent disciplines to being borrowed concepts in OB. This comparison has been based on the conceptual structures in the OB subareas.

In my view the analysis of conceptual relations by means of terminological methods has proven to be very useful to describe disciplines with highly overlapping concepts. This includes not only the analysis of the relationship between parent disciplines and the disciplines they lend concepts to, but also disciplines that are emerging within the same period of time and which are also highly related and interdisciplinary.

To conclude, terminological methods seem to offer a useful tool for clarifying relations and family recemblances among disciplines that are clustered together in the continuum of the scientific landscape.

## 3.6 Usability of the corpus material

The study has shown that the selected corpus material has rendered less information than I initially assumed. The textbooks are relatively vague in their conceptual descriptions, something which may be due to the inherent indeterminacy of the disciplines. It may, however, be questioned whether the textbooks focus enough on conceptual descriptions.

The research articles that were added to the corpus to provide a more detailed theoretical description than the textbooks, in general provided a practical, or real-life application of theoretical concepts. This made the research articles less suitable for the conceptual investigations than I first assumed. Another point is that although they may define themselves as being e.g. OB journals, it is quite clear that the point of view is often that of a social science discourse community. This raises the question of whether the borders between the disciplines are as vague as they appear from the articles, or whether the vagueness resides in the way these disciplinary borders are described in the articles. Instead of shedding a brighter light on the disciplinary autonomy of OB, the inclusion of such texts in the corpus material has rather strengthened my belief that the social sciences make up a family of highly related and overlapping disciplines of which autonomy is not a striking characteristic.

# 4 CONCLUSION

The conceptual analysis has given valid support for the hypothesis that several disciplines have influenced OB and its subareas by lending concepts to the discipline. Secondly, some changes in the intensions of the concepts that have been borrowed and adapted in OB have taken place; however, the analysis has shown that the majority of concepts that have been analysed have been adapted in OB without significant changes being made to their intensions. Thus, the most prominent feature of OB concepts in this respect is that conceptual changes have taken place in the extensions of the borrowed concepts, something which has resulted in new conceptual relations.

When it comes to the overall questions raised in the study, I have concluded that based on the results, OB can be described as an autonomous discipline. Secondly, terminological methods seem to offer a useful tool for clarifying relations and conceptual familiarities among interrelated disciplines. Consequently, I believe the methodology developed in the study can be a useful contribution in delineating related disciplines, something which is a challenge, for instance in the construction of terminologies, or the building of knowledge bases. In the attempt to establish knowledge bases, for instance, which are to contain terminology from several related disciplines, it has been a problem that the disciplines are highly interrelated and share many seemingly identical concepts.

A method for collecting corpus material, for instance, is to conduct searches on the internet for texts which contain some specified terms. By using the internet, terminologists are able to retrieve substantial amounts of text with relatively little effort. A challenge with this method is to decide to which discipline the concepts found in the selected texts belong. Another challenge is to assess which disciplines are actually represented in the retrieved texts. A knowledge base established according to terminological principles should take the concept of concept as a point of departure, which implies presenting the concept in its conceptual environment (i.e., with its conceptual relations explained). This will require that the various disciplines, such as OB, can be distinguished from related disciplines, such as marketing, strategy or even organisation theory.

To me, a promising point of departure seems to be to investigate the epistemological characteristics discussed in my doctoral thesis. The methodology will, obviously, not solve all problems of structuring the concepts of related disciplines since interrelatedness and multi-disciplinarity are such prominent characteristics of many disciplines. It may, nevertheless, increase the quality of the conceptual analyses.

## 5 REFERENCES


CERRELLA BAUER, S. (2004): Die Übersetzer-Terminologen als Wissensmanager. In: Hieronymus 3/2004, term bulletin of ASTTI (Association suisse des traducteurs, terminologues et interprètes), Bern, 3-16.

ANTIA, B.E. (2001). Metadiscourse in Terminology: Thesis, Antithesis, Synthesis. In B. Toft and H. Picht (eds.) Terminology Science & Research. Journal of the International Institute for Terminology Research 12 (1–2), 65–84.

BERNSTEIN, D.A., A. CLARKE–STEWART, E.J. ROY and C.D. WICKENS (1997). Psychology. 4th ed. Boston: Houghton Mifflin Company.

Blackwell Encyclopedic Dictionary of Organizational Behavior. Rev. ed. (1998). N. Nicholson (ed.) Massachusetts: Blackwell.

DEAUX, K., F.C. DANE and S. WRIGHTSMAN (1993). Social psychology in the ´90s. 6th ed. California: Brooks/Cole.

GARFIELD, E. (1979). Citation indexing: its theory and application in science, technology, and humanities. Information sciences series. New York: Wiley.

GIDDENS, A. (1997). Sociology. 3rd ed. Cambridge: Polity Press.

JAKOBSEN, O. (1980). Markedsføring som akademisk fag. Oslo: Universitetsforlaget.

JAKOBSEN, O. and K. GRØNHAUG (1993). Disciplinary influences on research in marketing: a citation analysis. Bodø: Bodø Graduate School of Business.

KAUFMANN, G. (2004). Evaluation from Prof. Kaufmann at NHH of the findings from the conceptual analysis presented in Chapter 6 and 7, personal communication.

KRISTIANSEN, M. (1997). A Terminological Analysis of Organisational Behaviour. Unpublished cand.merc. thesis. Bergen: Norwegian School of Economics and Business Administration.

KRISTIANSEN, M. (2004). The Multi-Disciplinary Nature of the Social Sciences. Investigating Disciplinary Autonomy in Organisational Behaviour by means of Terminological Analysis. Dr.art. thesis. University of Bergen.

KUHN, T. S. (1970). The Structure of Scientific Revolutions. Chicago: University of Chicago.

LAURÉN, Ch. and H. PICHT (1993). Vergleich der terminologischen Schulen. In Ch. Laurén and H. Picht (eds.) Ausgewählte Texte zur Terminologie. Wien: TermNet, 493–539.

LAURÉN, Ch., J. MYKING & H. PICHT (1997). Terminologi som vetenskapsgren. Lund: Studentlitteratur.

NORTHCRAFT, G.B. and M.A. NEALE (1994). Organizational Behavior – A Management Challenge. Orlando: 2nd ed. The Dryden Press.

PASSER, M.W. and R.E. SMITH (2001). Psychology: frontiers and applications. Boston: McGraw-Hill.

ROBBINS, S.P. (1996). Organizational Behavior: concepts, controversies, and applications. 7th ed. Englewood Cliffs, N.J.: Prentice-Hall.Shani, A.B. and J.B. Lau (1996). Behavior in Organizations – An Experimental Approach. USA: Irwin.

SAGER, J. (1990). A practical course in terminology processing. Amsterdam: John Benjamins Publishing Company.

SMITH, E.R. and D.M. MACKIE (2000). Social psychology. 2nd ed. Philadelphia: Psychology Press.

THAGARD, P. (1992). Conceptual revolutions. Princeton, N.J.: Princeton University Press.

WITTGENSTEIN, L. (1963). Philosophical investigations. Oxford: Blackwell.

**Anita Nuopponen**
**Assistant professor in Applied Linguistics**
**University of Vaasa**

# SUBJECT CLASSIFICATION AS A TERMINOLOGY RETRIEVAL TOOL IN AN ONLINE ENCYCLOPEDIA

There are several means to assist an information seeker in online information services e.g. in online glossaries, terminological databases, online encyclopedias, bibliographies, and link collections. A widely utilised method in free online glossaries from different subject fields is a simple search function where the information seeker types a word in a box. This kind of search, however, requires that the user knows one or more search terms to start with. Often only a single article with a term and a definition, or an equivalent in another language is shown to the user. This search method does not support browsing in cases in which the information seeker does not have an exact term in mind. In some glossaries, the word search leads the user to the wanted term entry in an alphabetical list of entries, thus showing all the term entries around it. In another type of typical online glossary, the material can be browsed by the alphabet. Sometimes the word search and the letter search are combined, see Figure 1.

## INTRODUCTION

### Remarks concerning ontological and epistemological issues

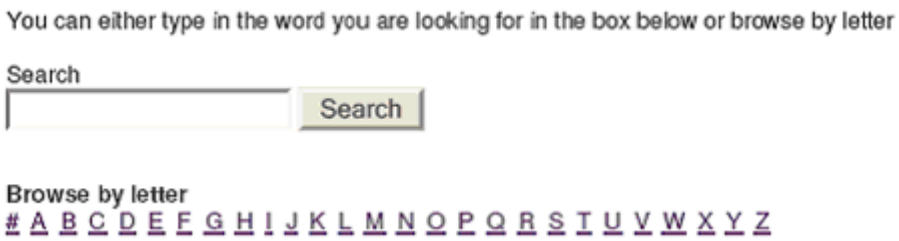


**Figure 1.** A typical interface of an online glossary1

Browsable subject (field) categories have been successfully used to assist information seekers in different kinds of online information resources such as link collections, e.g. Yahoo Directories2 . They are also used in online encyclopedias, e.g. Wikipedia (categories), Encarta (categories3 ), and Encyclopaedia Britannica Online (subjects4 ). The purpose of Wikipedia category classification is defined as follows:

> "Categories (along with other features like cross-references, lists, and infoboxes) help users find information, **even if they don't know that it exists or what it's called**."[5]

Browsable subject (field) classifications offer a solution when the information seeker does not have an exact search term in mind. This is also why I am interested in testing the function of the classification of Wikipedia. Traditional terminological databases utilise subject field classification, but not normally for browsing purposes. Instead, the classification assists in delimiting the field of the search term (see Figure 2), or to give information on the domain of the concept retrieved. Traditional terminological databases, such as Eurodicautom, are geared for translators and support searches where the user has a term to start the search with.

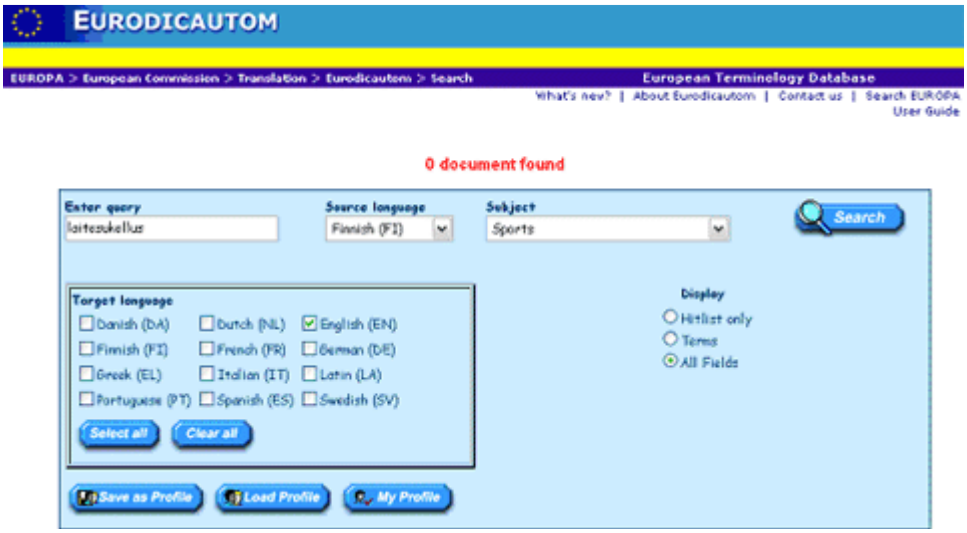


**Figure 2.** The interface of Eurodicautom terminology database.

The purpose of this paper is to discuss the use of browsable categories as a retrieval tool for terminological information. Instead of looking at existing or planned terminological vocabularies or data bases, this paper takes a look at an Internet based encyclopedia, Wikipedia. The question to be asked here is How to find desired terminological information by using the category classification provided by the encyclopedia? I set two information seeking tasks for myself in order to test the classification:

> 1) Find a definition of the concept Japanese tea ceremony.
> 2) Find English equivalents for the Finnish terms 'laitesukellus' (direct translation: "apparatus diving"), 'avovesisukellus' ("open water diving") and 'parisukellus' ("pair diving") in a case in which the information seeker is familiar with the concepts but is not sure about the English equivalents and cannot find them in normal dictionaries.

Before discussing this pilot study and the implications of the findings for a terminological hypertext system, the general ideas behind encyclopedias, Wikipedia and subject classification systems are scrutinized.

## 1 ENCYCLOPEDIA

Traditionally, encyclopedias are printed books or sets of books. They contain authoritative information about a variety of fields in the form of factual articles, normally subject to editorial approval. Articles vary in length, but are longer than the ones in glossaries and dictionaries. There are also vocabularies or glossaries, which are called 'encyclopedias' and encyclopedias, which are called 'terminologies', 'glossaries', or 'dictionaries'. Encyclopedias may be general in scope, such as the Encyclopedia Britannica or subject-oriented, such as The Encyclopedia of Philosophy. Articles in most printed general encyclopedias are organised in alphabetical order and the user must know what she/he is looking for. Indexes covering terms, which appear in articles and references in the text help in this. Articles in encyclopedias typically also contain references to authoritative books and articles on the subject.

The tasks of encyclopedias are, on the one hand, "to document and maintain authentic knowledge, ensuring and testifying its preservation over time and space", and on the other hand, "to provide adequate categorization and systematization of knowledge, providing easy access to knowledge for any interested person" (Lehner et al.: 3). For terminologists and other seekers of terminological information, encyclopedias have always been important reference sources. They systematize and define concepts and present the information in the form of a compact text. For an encyclopedia article, each subject or field is researched and presented as a whole. This makes it easy to extract terms, concepts, and definitions and to draw a graphical presentation of the concept system as well as to compile a glossary. Because of the systematic nature of encyclopedia articles, they are also used as sources for several ongoing projects in which terminology is extracted and knowledge or terminology bases are built automatically using natural language technologies (see e.g. Sui, Cui et al.: 2005).

Today, several printed encyclopedias have migrated into the Internet, e.g. The Encyclopædia Britannica Online6 and The Columbia Encyclopedia7. A further development are encyclopedias, which exist only on the www, e.g. Encyclopedia.com and Wikipedia. Many online encyclopedias are copies of the print version, or imitate one.

## 2 WIKIPEDIA

Wikipedia was selected as the object of study because it differs from other online encyclopedias in several ways, and because its popularity as a source of terminological information is growing all the time. Its content is dynamic: it is written collaboratively by volunteers, and anyone can add or change an article. Wikipedia's content is free whereas the commercial online encyclopedias may show only the beginning of the articles to those who have not registered (see e.g. Encyclopædia Britannica). Wikipedia was started in 2001 and is now operated by the non-profit Wikimedia Foundation. In November 2006, it had over 5 million pages, including more than 1,504,000 (compared to 961,000 in February 2006) in the English-language version, 88,000 (48,000) in Finnish, 195,600 (135,000) in Swedish, and over 7,500 (4,000) pages in Latin, just to mention a few of the numerous languages which have their own independent Wikipedia versions.8 The English version and its category classification will be used here.

Wiktionary, a sister project, is a similar collaborative project with the aim "to produce a free multilingual dictionary in every language, with definitions, etymologies, pronunciations, quotations, synonyms, antonyms and translations".9 Wiktionary also contains glossaries of different subject fields. However, they are not yet extensive enough and do not yet utilise the possibilities available on the web. Wikipedia itself is still more interesting for terminological purposes with its huge network of knowledge on any subject in various languages.

## 3 ORDERING AND CLASSIFICATION SYSTEMS

Different types of classification systems have been created for ordering and searching for books and information. Library and other classifications have found new life as knowledge organisation systems for networked knowledge. Originally, library classifications were meant for ordering "the fields of knowledge in a systematic way", bringing "related items together in the most helpful sequence", providing "orderly access to the shelves", and providing "an exact location for an item on the shelf." (Suman & Karmakar 2002.) These needs still exist, but on the World Wide Web, bookshelves are replaced by hypertext pages or databases. Library classifications have been designed to cover everything under the sun and above and must therefore remain on a high level of abstraction. As an example could be mentioned The Universal Decimal Classification (UDC), see Table 1.

Table 1. Outline of the UDC.

| | |
|---|---|
| 0 | GENERALITIES |
| 1 | PHILOSOPHY. PSYCHOLOGY |
| 2 | RELIGION. THEOLOGY |
| 3 | SOCIAL SCIENCES |
| 4 | VACANT |
| 5 | NATURAL SCIENCES |
| 6 | TECHNOLOGY |
| 7 | THE ARTS |
| 8 | LANGUAGE. LINGUISTICS. LITERATURE |
| 9 | GEOGRAPHY. BIOGRAPHY. HISTORY |

Like so many other information services, Wikipedia has its own subject field classification, Categories. In February and July 2006, the classification consisted of 9 main categories, which all have a large number of subcategories (see Table 2).

Table 2. Wikipedia's categories (12.07.2006)

| | |
|---|---|
| **Art and Culture** | Arts and crafts · Cultural movements · Entertainment · Films · Food and drink · Games · Languages · Literature · Mass media · Museums · Music · Mythology · Parties · Performing arts · Pets · Popular culture · Radio · Sports · Television · Traditions · Tourism · Toys |
| **Geography** | Africa · Antarctica · Asia · Australia · Europe · North America · Oceania · South America · Cities · Climate · Countries · Landforms · Maps · Parks · Subterranea · Towns · Villages |
| **History** | Africa · Asia · Australia · Eurasia · Europe · North America · Oceania · South America · By Period · By Region · By Country · By Topic · Colonialism · Historiography · Timelines |
| **Mathematics** | Algebra · Analysis · Arithmetic · Economics · Education · Equations · Geometry · Logic · Measurement · Numbers · Proofs · Theorems · Trigonometry · Statistics |
| **Natural sciences** | Applied sciences · Astronomy · Biology · Chemistry · Earth sciences · Ecology · Heuristics · Health sciences · History of science · Information science · Medicine · Scientific method · Physics · Protoscience · Scientists · Space · Systems theory |
| **Philosophy and Religion** | Lists · Branches · Movements · Schools and traditions · Theories · Arguments · Philosophers · Literature · History · By era · By region · Aesthetics · Epistemology · Ethics · Logic · Metaphysics · Buddhism · Christianity · Confucianism · Cults · Deism · Judaism · Islam · Religious Movements · Satanism · Taoism |
| **Social sciences** | Anthropology · Archaeology · Cultural studies · Demographics · Economics · International relations · Linguistics · Psychology · Media studies · Political science · Social scientists · Sociology · Sexology |
| **Society and People** | Biographies · Business · Communication · Education · Ethnic groups · Family · Finance · Gender · Government · Health · Home · Industries · Labor · Law · Mass media · Organizations · Politics · War |
| **Technology** | Agriculture · Architecture · Automation · Automobiles · Big Science · Biotechnology · Chemical processes · Computing · Electronics · Energy · Engineering · History of technology · Information technology · Internet · Manufacturing · Nanotechnology · Nuclear technology · Sound · Telecommunications · Technology forecasting · Tools · Transportation · Vehicles |

The classification in Table 2 will be used in this study, because when I made the analyses of the paths in January/February 2006 and in July 2006, the classification had not been changed. In November 2006, when checking some information, however, I found that the classification of categories had been extended to cover 2 new main categories. Also, the names of the main categories had been changed: Reference (new), Geography and places, Health and fitness (new), History and events, Mathematics and abstractions, Natural sciences and nature, People and self, Philosophy and thinking, Religion and belief systems, Social sciences and society, and Technology and applied sciences.10

The subcategories in the Wikipedia classification are linked to pages, which often contain a short introduction to the subject. On these introductory pages, there is a list of further subcategories as well as links to articles directly placed under the category in question. Every subcategory has these elements on its own page, too. Thus, some articles can be found after a couple of clicks, but for others, more clicking will be needed as we see below where the paths to the example information are described.

## 4 PATHS LEADING TO THE INFORMATION

In what follows, the paths leading to the desired information will be discussed in four stages: selecting the entry points (4.1), finding the right articles (4.2), tracing backwards to find alternative paths (4.3), and finally revisiting Wikipedia four months later (4.4).

## 4.1 Finding the entry point

Even though the Wikipedia category classification is comprehensive, it proved to be difficult to find the topmost categories for both tasks to start with. For finding the article "Japanese tea ceremony" the first guess, *Arts and Culture*, was not successful. *Ikebana* and *origami* were found directly under *Arts and Culture*'s subcategory *Arts and Craft*, but not Japanese tea ceremony which is very often associated with them as a hobby or an art form originating from Japan. The second guess was the path Geography: Asia: Japan, but it did not bring any results during the first two visits to the site in February or July.[11]

As to the second task, the target was an article or articles on the type of diving, in which breathing apparatus is utilized (fi *laitesukellus*) and a couple of its subtypes. The first guess was the main category Society and People, since the first supposition was that it would include Sports (see Figure 4). This proved to be a dead end, too.

In both cases, I had to return to the higher-level categories once more and read them through more carefully than I had initially done. The main category *Arts and Cultures* eventually proved to be a fruitful starting point for both tasks: for the tea ceremony its subcategory *Food and drink* and for diving *Sports*.

After I found the category Sports under *Art and Culture* – which I did not at first think of as the most obvious choice – the task of finding diving related information became easier: Sports had a subcategory *Water sports*, and *Diving* could be found under it.

## 4.2 Detours and ambiguities

Even though in both cases the entry points were found, there were still some problems before the desired information was discovered. The category classifications did not lead the information seeker directly to the actual articles but category articles were needed as intermediaries.

In order to find the path to tea ceremony, the subcategory *Food and drink* was tried. It had a subcategory *Ceremonial food and drink* which did not bring any results, but a parallel subcategory *Beverages* had a linked note on its introductory page saying: "*The main article for this category is* ***Beverages***". This category article was on a page of its own and it listed links to articles on different types of beverages including "Tea". In the article on tea, the Japanese tea ceremony was briefly mentioned , but without linking the term to the article "Japanese tea ceremony". A link to the article was finally found on the same page under the "See also" heading.

The first task was to find a definition for Japanese tea ceremony. In the introductory part of the article found, the Japanese tea ceremony is defined in the following way:

> "The **Japanese tea ceremony** (**chadō**, or **sadō**) is a traditional ritual influenced by Zen Buddhism in which powdered green tea, or matcha, is ceremonially prepared by a skilled practitioner and served to a small group of guests in a tranquil setting".

Some slight confusion is caused by the use of the terms as equivalents for the term 'Japanese tea ceremony' (see above) on the one hand, and for "the study or doctrine of the tea ceremony" on the other. This ambiguity is, however, something, which appears in other sources, too, and the hoped solution for it was not found here either even thought the right article was found.

Ambiguity or polysemy was also encountered in the case of the diving article, but on another level. In the main article "Diving", the term 'diving' did not, after all, refer to the concept, which I was looking for, not even to its superordinate concept. Actually, it was a co-ordinated concept, a fact clarified by a note on the page:

> "This article refers to the sport of jumping into water, often acrobatically. For swimming below the surface of the water, see underwater diving."[12]

When following this link, the "Underwater diving" article told me that

"**Underwater diving** refers to the practice of going underwater with or without breathing apparatus. [...] There are several types of underwater diving. [...] Scuba diving and surface supplied diving: swimming or walking underwater with breathing apparatus."[13]

The information provided in the article confirmed that the term 'scuba diving' was the equivalent of the Finnish term 'laitesukellus'. Following the link provided from the term 'scuba diving', an article was found which concentrated on this type of diving. In the article, however, no links or mention of the two further concepts I was looking for were found. Returning to the category Diving proved to be fruitful. There was a link to the article "Open-water diving", which confirmed that the Finnish term 'avovesisukellus' ("open water diving") can actually be translated with 'open-water diving' - it is not always certain that such a word for word equivalent is correct. As to the Finnish 'parisukellus' ("pair diving"), the task was slightly more difficult, since no direct translation of the term was to be found in the category Diving. It had, however links to the articles "Buddy check" and "Buddy system", both of which referred to the practices in diving in pairs. The nearest equivalent to the Finnish term proved to be 'buddy system', or 'buddy diving":

"When using the buddy system, the group dives together and co-operate with each other, so that they can help or rescue each other in the event of an emergency. […] With buddy diving, each of the divers is presumed to have a responsibility to the other. The "buddies" are expected to monitor each other […]" 14 In the article, the term 'buddy diving' is also used alternately with 'buddy system'. Thus all the tasks were performed and the desired information was found despite many dead ends.

## 4.3 Multiple paths

Once the relevant articles had been found, other possible pathways could be traced by following upwards the links provided at the bottom of every page. They lead to the superordinate categories. For both tasks, several alternative paths were found. Tracing upwards showed that there is a category called Tea, which has a subcategory Tea ceremony, the whole path being: Art and Culture – Food and drink – Beverages – Non-alcoholic beverages – Tea – Tea ceremony – Japanese tea ceremony (see Figure 3). It reflects the actual generic concept systems ("Beverages" and "Tea ceremony"), but for an information seeker it may be a bit too long. At least I did not even expect to find it and tried rather more associative paths at first. Another alternative and much shorter path was found when tracing upwards from the linked category Rituals via Culture to the uppermost category Society, which was not found in the beginning.

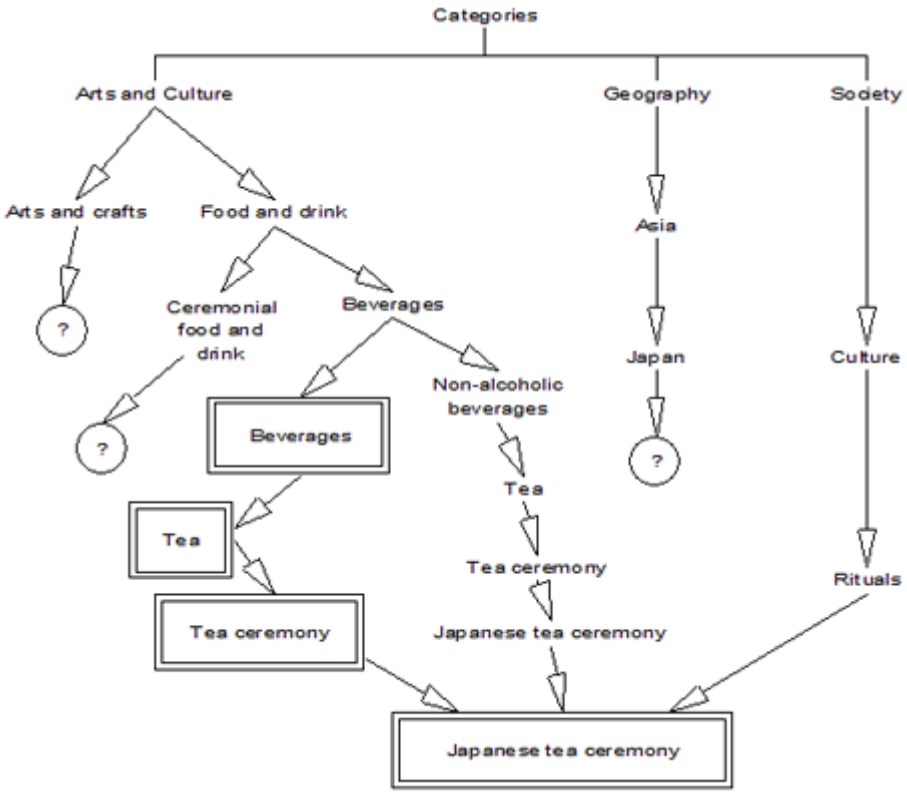


**Figure 3.** Paths to Japanese tea ceremony15

Tracking upwards showed that Diving could have been found through many different paths, the most obvious ones of which are included in Figure 4 together with the paths, which were found earlier.

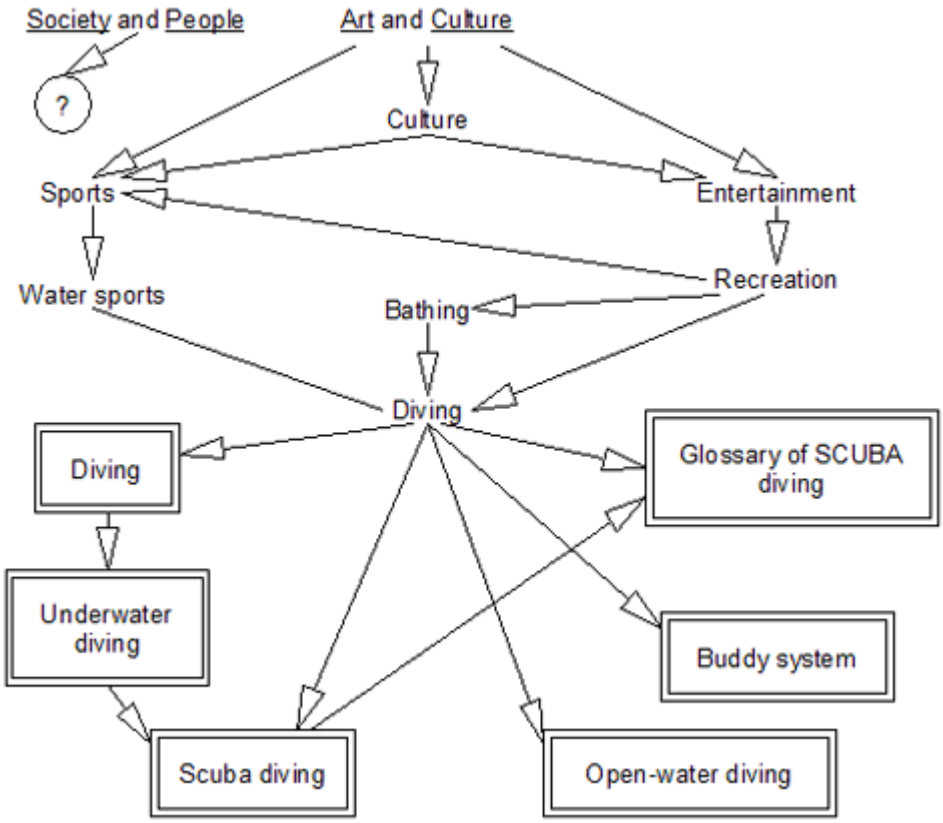

**Figure 4.** The paths to scuba diving

## 4.4 Opening new paths

As mentioned above, Wikipedia is dynamic in nature since new articles are added all the time and old ones updated and completed. Not only had the main categories been updated, but subcategories had also been changed somewhat after the visits in January/February and July. In my first visit, I was looking in vain for the article on tea ceremony via the path Geography: Asia: Japan. In November however, under the category Japan, there was a subcategory called Japanese Culture, and according to my earlier expectations, this path now leads to the subcategory Japanese tea ceremony. The classification of the higher categories in Wikipedia had undergone major reformulation, however, and the path had been lengthened (Geography and places: Countries: Countries by continent: Asian countries).

Some changes had also taken place in the article on scuba diving. Earlier it had no direct links to subtypes of scuba diving, but now on the third visit, it had an introductory section and direct links to some subtypes of scuba diving (e.g. recreational diving and technical diving) had been added. However, as before, the links to buddy system and open water diving were still missing.16

# 5 DISCUSSION

Collaborative online encyclopedias such as Wikipedia are revolutionising the genre of encyclopedia, and present an interesting object of study to terminologists, too. How to represent, organise, and find information in terminological glossaries and databases have been important topics for terminologists since the very beginning of the theory of terminology.

People often tend to think that all of knowledge can be organised in a single hierarchy, a comprehensive classification of all the phenomena in the universe, which makes up a summary of total human knowledge. At first sight, it may seem that Wikipedia categories follow strict hierarchic concept categories (see Table 2). Only the topmost categories are more or less mutually exclusive. Nevertheless, when we proceed to the categories at lower levels of abstraction, crossing paths will be found (see e.g. Figure 4). The same categories and articles may be found under several different categories. Instead of a single hierarchy, several overlapping ordering systems and concept systems interact". When browsing the categories, I even noticed cases in which a superordinate category became subordinate to its own subordinate category. However, Wikipedia instructions do not encourage this kind of looping. 17

Multiple entry points and overlapping hierarchies offer a smooth way to find the desired information. Different people have different points of view and start their search for the same item from different points of departure. In terminological analysis, we have also seen that the same concepts may belong to several different concept system types, e.g. generic, partitive, or functional concept systems. Instead of a hierarchy, the structure of a macro concept system rather forms a network of concepts (Nuopponen 1994). Despite the fact that the Wikipedia paths of categories were sometimes quite logical ones, including generic concept relations (e.g. Food and drink – Beverages – Non-alcoholic beverages – Tea – Tea ceremony – Japanese tea ceremony), they did not necessarily attract the information seeker as much as a supposedly shorter one with associative concept relations (e.g. here origination relation: Japan – Japanese culture – Japanese tea ceremony).

In addition to the hierarchical category classification, the English Wikipedia has an alphabetical listing of all the main and sub-categories, in which the categories Japanese tea ceremony and Diving could be found directly. It was also easy to find the relevant articles by means of the search function, e.g. by using the terms diving or diving apparatus, but the purpose was to navigate through the classification provided – after all, the classification exists for the purpose of assisting the information seeker. It must be added, however, that in both tasks I lost my way during the first efforts, and instead of finding the path along the categories to the target, I was able to navigate forwards by means of the articles and links in them. Thus the category classification system, together with the articles attached to them, formed a functioning retrieval tool: when categories did fail, articles helped me on and vice versa. In both cases, paths along categories were discovered afterwards when tracing the links to the upper categories from the articles. In a terminological hypertext system, all these four retrieval tools (categories, alphabetical lists, search, and links) would be needed to complement each other. In addition to these, Wikipedia has

other tools for information retrieval: overviews, featured content, lists of lists, glossaries, portals, and timelines. It also tries to apply other classification systems, e.g. the Library of Congress classification.18

When browsing Wikipedia's categories and looking for terminological information in the articles, an idea for further discussions and projects came up: There would certainly be a need for a collaborative terminological hypertext system similar to Wikipedia, where the lenghty encyclopedia articles would be replaced by entries with definitions and other terminological data compiled by means of terminological methods. The entries would be connected to each other via a subject category classification and links. Instead of building a hypertext system for terminology from scratch, would it be possible to publish terminological glossaries as part of Wikipedia's Wiktionary? Wikipedia and its sister projects as well as wiki applications offer versatile tools for collaboration and for publishing terminological glossaries online. Are these tools suitable for terminology work? Resources and copyrights are certainly an issue, but we could start with students' terminological projects and theses. These questions and ideas will be, however, left for future discussions and projects.

[1]Nowodiver's Glossary of scuba diving terms, http://www.nowodiver.net/en_ glossary.php
[2]http://dir.yahoo.com/
[3]http://encarta.msn.com/artcenter_0/Encyclopedia_Articles.html
[4]http://www.britannica.com/eb/subject
[5]Wikipedia; http://en.wikipedia.org/wiki/Wikipedia:Categorization (12.7.2006)
[6]http://www.britannica.com
[7]http://www.bartleby.com/65/
[8]http://meta.wikimedia.org/wiki/List_of_Wikipedias
[9]http://en.wiktionary.org/wiki/Main_Page
[10]http://en.wikipedia.org/wiki/Wikipedia:Categorical_index (28.11.2006)
[11]See Figure 4; dead ends are marked with a question mark.
[12]Wikipedia; http://en.wikipedia.org/wiki/Diving (7.2.2006)
[13]Wikipedia; http://en.wikipedia.org/wiki/Underwater_diving (7.2.2006)
[14]Wikipedia; http://en.wikipedia.org/wiki/Buddy_system (7.2.2006)
[15]Dead ends are marked with question marks and articles with double frames.
[16]Wikipedia; http://en.wikipedia.org/wiki/Scuba_diving (27.11.2006)
[17]Wikipedia; http://wikipedia/WikipediaCategorisation_FAQ.htm (12.7.2006),
http:// en.wikipedia.org/wiki/Wikipedia:Categorization (28.11.2006)
[18]Wikipedia; http://en.wikipedia.org/wiki/Wikipedia:Contents (11.12.2006)

REFERENCES

Lechner, Ulrike, Beat Schmid, Salome Schmid-Isler, Katarina Stanoevska-Slabeva (1999). Structuring and Systemizing Knowledge on the Internet - Realizing the Encyclopedia Concept as a Knowledge Medium. In Proc. Int. Conf. on Information Resource Management (IRMA'99). Available 28.11.2006: http://www. alexandria.unisg.ch/EXPORT/DL/10280.pdf

Nuopponen, Anita (1994). Begreppssystem för terminologisk analys. [Concept systems for terminological analysis] Acta Wasaensia. University of Vaasa, Vaasa.

Sui Zhifang, Gaoying Cui, Ding Wansong, Zhang Qinlong (2005). Domain Knowledge Engineering Based on Encyclopedias and the Web Text. The 5th Workshop on Asian Language Resources. Available 3.12.2006: http://www.cl.cs.titech.ac.jp/ALR/WS/5th/5thALR.pdf

Suman S., Debanshu Karmakar (2002). The Role of Library Classification in Organizing the Web. Available 28.11.2006: http://drtc.isibang.ac.in/~saiful/ greenstone/collect/Pdf/archives/HASH0193.dir/doc.pdf