

162nd CSO Meeting, 14-15 June 2005
Proposal for a new COST Action

COST A31

**“STABILITY AND ADAPTATION OF CLASSIFICATION
SYSTEMS IN A CROSS-CULTURAL PERSPECTIVE”**

Proposer: Dr. Thekla Wiebusch,
CRLAO (CNRS),
54 Bd Raspail, 75006 Paris
Tel. +33-1-49542403
wiebusch@rz.uni-leipzig.de

COST National Coordinator: Dr. Michel Gorlicki
Ministère de la Recherche et des Nouvelles
Technologies Direction de la Technologie -
Bureau de la Coordination Européenne et des
Relations Internationales
1, rue Descartes 75231 Paris Cedex 05
FRANCE
michel.gorlicki@technologie.gouv.fr

TC-SSH Rapporteur: Mr. Jan-Arne Eilertsen
The Research Council of Norway
Stensberggata 26
0131 Oslo
NORWAY
Tel:+ 47 22 03 73 89
Fax:+ 47 22 03 74 09
jae@rcn.no

DRAFT
MEMORANDUM OF UNDERSTANDING

for the implementation of a European Concerted Research Action designated as

COST Action A31

**“STABILITY AND ADAPTATION OF CLASSIFICATION SYSTEMS IN A
CROSS-CULTURAL PERSPECTIVE”**

The Signatories to this Memorandum of Understanding, declaring their common intention to participate in the concerted Action referred to above and described in the Technical Annex to the Memorandum, have reached the following understanding:

1. The Action will be carried out in accordance with the provisions of document COST 400/01 "Rules and Procedures for Implementing COST Actions", the contents of which the Signatories are fully aware of.
 2. The main objective of this Action is to gain a deeper understanding of human categorisation as it is reflected in different types of classification systems.
 3. The economic dimension of the activities carried out under the Action has been estimated, on the basis of information available during the planning of the Action, at Euro 22 million in 2004 prices.
 4. The Memorandum of Understanding will take effect by being signed by at least five Signatories.
 5. The Memorandum of Understanding will remain in force for a period of 4 years, calculated from the date of first meeting of the Management Committee, unless the duration of the Action is modified according to the provisions of Chapter 6 of the document referred to in Point 1 above.
-

COST A31

**“Stability and adaptation of classification systems
in a cross-cultural perspective”**

Abstract

The main objective of this Action is to gain a deeper understanding of human categorisation as it is reflected in different types of classification systems. More specifically, increased knowledge should be achieved concerning the following aspects of classification systems: Types of classification systems and their main distinctive features, universal as well as culture/context dependent features of classification systems, domain specificity and internal domain structure, the comparative development of classification systems, the interrelation between stabilization and adaptative potential, and the relation between individual level and cultural level categories.

A. Background

Research in Human Categorization was first initiated in the United States (Rosch 1973, 1981, Lakoff 1987), but soon gained ground in Europe. Its progress was closely related to the rise of the cognitive sciences as an independent discipline. The seventies, eighties and nineties saw a wealth of publications in cognitive linguistics (Wierzbicka 1972-94), ethnobiology (Brown 1974-84, Atran 1987, Berlin 1976,1977,1992), and other disciplines on universals in human categorization . The central issue was the understanding of so-called prototype or radial categories allowing for graded membership, overlapping categories, chaining and other features of classification that were contrasted with a so-called “Aristotelian” or “objectivist” model of classification supposing hierarchical systems of univocal class membership based on distinctive features. The “prototype” approach was applied to a range of linguistic phenomena (Kleiber 1990, Declés 1998, Dubois 1993, 1995). In anthropology it led to an increased research in folk taxonomies of plants and animals (Berlin et al. 1968-1992, Brown 1974-1984 and many others), body part terminology (Burton and Kirk 1979) and artefacts (Kempton 1978, 1981); in linguistics it fostered the study of systems of nominal classification such as Bantu noun class systems (Senft 2000), numeral classifier systems (cf. Aikhenvald 2000), gender systems and verbal classifiers.

This development went hand in hand with a vast amount of descriptions of single instances of classification systems (Hundius & Kölver 1983), their acquisition (Sanches 1977, Erbaugh 1986) and cognitive implication (Craig, ed. 1986). There have also been comparative studies within single types of classification systems such as folk taxonomies (Brown 1984, Berlin 1992) or noun classification systems (Adams

1982). On the one hand, recent typological studies have given an overview of the subtypes of noun classification systems and their geographical distribution (Grinevald 1999, Aikhenvald 1997, Senft 2000). On the other hand, research projects as those of James Tai at Chia-yi, Taiwan, have investigated the semantic features and cognitive principles of Chinese dialects such as Minnan and Hakka in ever greater detail. Human categorization in determinatives of hieroglyphs have been studied by a project of the universities of Göttingen and Jerusalem (David 2000, Goldwasser 2002) and classifiers in sign languages/script (Emmorey, ed. 2003) received increasing attention.

Parallel to this development, computer linguistics have experienced a range of new developments. The detection of correlating features by using statistical tools is already well established, but the increasing memory capacities of computers as well as the development of efficient software, involving clustering and multidimensional scaling, allows the application to ever more complex systems. The recent work on the origin of language and the connection between linguistic, archeological and genetical data have led to the development of a range of new software tools for modeling language evolution (see Wang , Ke & Min, in press) and analysing phylogenesis, taking phonetic, cladistic or probabilistic approaches such as Paup, Phylip or MrBayes. Results from the humanities, especially from philosophy, cognition and semantics have been realized to be important in the development of software for knowledge retrieval and organization in the internet, multilanguage lexical tools, and data mining. These software applications (such as Semantic Web, WordNet, to mention only a few) make use of insights on ontologies (Johansson 1989), natural language processing and construction of categories in the human brain to develop feasible systems to exploit and integrate complex and contradictory data pools and information systems.

It seems necessary at this point to bring together the different approaches in a systematic way and build a network of the loosely – if at all – connected European research on human categorization to make a typological study of classification in languages and writing systems, seeing them as different expressions of human categorizations, employing different strategies according to the purposes of the systems and their specific constraints.

There is no other COST-Action touching upon classification systems or human categorization. The ESF-programme “Origin of Man, Language and Languages” also has an interest in the development of linguistic phenomena, but its focus is on the earliest development of language as a human capacity as such, on the one hand going far beyond classification systems, on the other hand not including script or systems of knowledge representation.

The focus of the “Mind, Language and Metaphor” Euroconferences (ESF) is on the role of imagination in human thought. The “Classification”-Action will cooperate with this programme as far as imagination is involved in classification systems. For contacts with experts of sign language and sign writing the Action can

built on structures established by the completed ESF-network INTERSIGN. (Sign Linguistics and Data Exchange).

As knowledge representation is becoming a key topic in processes of applied information sciences such as datamining or the development of multilingual lexical tools, a cooperation with the European Network of Excellence on Semantic Interoperability and Datamining in Biomedicine may be of mutual benefit.

The COST programme offers a unique and necessary opportunity to establish a network between a large number of research projects in very different disciplines. As a comparative approach to classification systems in this dimension is a novelty, only a period of cooperation between these diverse perspectives and methodologies will show which lines of research will be the most promising. The flexibility of COST allows to make the necessary adjustments in the course of the Action. Another important aspect is the possibility to integrate and support young researchers by granting short term missions and giving them opportunities to present their work in Working Group meetings to get feedback from experts of different disciplines. Promising young researchers can be encouraged to pursue this line of research during training schools and be integrated at a later point of the Action.

B. Objectives and Benefits

The main objective of this Action is to gain a deeper understanding of human categorisation as it is reflected in different types of classification systems. More specifically, increased knowledge should be achieved concerning the following aspects of classification systems:

- (1) Types of classification systems and their main distinctive features
- (2) Universal as well as culture/context dependent features of classification systems
- (3) Domain specificity and internal domain structure
- (4) The comparative development of classification systems
- (5) The interrelation between stabilization and adaptative potential
- (6) The relation between individual level and cultural level categories

Working towards this main objective, the Action aims to improve interdisciplinary communication on the topic of classification systems. The main focus will be on intensifying the dialogue between theory and description, making easy to grasp information about new models and theories available.

A second focus will be on creating an exchange between researchers working on different cultures and periods and establish a joint basis for communication in the form of clear, compatible definitions and the clarification of terminologies. This will bring a broader perspective to the researchers involved.

A third focus is to enhance the exploitation of new computer tools for data collection and analysis by making information on these tools available to the network and beyond, comparing their usefulness for certain purposes and increase the dialogue

between computer scientists and users from the humanities to cooperate in the development of software tools.

This interdisciplinary exchange between otherwise often isolated or loosely connected scientific communities from different parts of Europe will substantially improve the quality and effectiveness of researchers working on related topics.

Better knowledge about human categorization and classification systems is centrally relevant in the wider social context of the information age. The mastering of knowledge systems and understanding of knowledge representation is an essential skill for obtaining information as well as for exploiting information in different information systems. This can be seen in the Semantic Data Mining projects but also in everyday tasks such as arranging products in a supermarket, automatic translation or designing a Web Portal.

The understanding of stabilizing forces and adaptive mechanisms of classification is a precondition for effectively dealing with the consequences of globalization. Processes such as the integration of new EU countries and the harmonization and standardisation of basic systems of society, e.g. educational systems, require an unprecedented rapidity of adaptation of concepts and categories on an individual level as well as in official classification systems. The intensive economic interaction between Europe and Asia, for example, entails the integration of hardly compatible management and computer systems from different countries or enterprises. The obtained knowledge about factors facilitating or inhibiting changes in classification systems will contribute to raising awareness of cognitive factors affecting culture contact and global integration, and to design more adequate tools and policies.

C. Scientific programme

In order to reach the objectives outlined in Section (B), the following Scientific programme will be pursued during the Action:

Typology of classification systems: In this first part of the scientific programme, a typology of classification systems will be established. On the one hand, their functions and conditions of usage have to be carefully described according to features and definitions to be agreed on during the initial phase of the Action. These formal conditions have an important impact on strategies of categorisation employed and the evolution of the systems. On the other hand, the relevant semantic domains of different systems as well as the salient criteria for classification have to be compared cross-linguistically and put into relation with the formal features of the system. In order to compare these different systems and cultures, there are two necessary steps:

- (1) Collection of equivalent data from different systems. It is the intention for each system to give a broad overview of the main semantic domains and their structure, along with a finer analysis of a selection of semantic domains from different realms of knowledge. There will be at least one domain from nature/living beings (e.g. animals, plants, the body, weather); culture/artefacts (e.g. clothes, buildings, writings) and abstract notions or eventualities (e.g. movement, emotions, colours, events). The in-depth treatment of semantic domains from different contexts will

make it possible to analyse the obtained features of classification systems concerning domain specificity.

- (2) Developing a methodology for establishing equivalence between data from different languages. In order to obtain comparable data, not only the semantic domains will must be agreed, but also the amount and type of data to be furnished must be fixed. Especially in the case of historical sources, not all information will be available.

The basic tool for systematically comparing and analysing these fragmentary and partly incoherent data will be a joint relational database prepared for this purpose by the proposer of the Action. The database will contain general information on the structure and function of the classification systems as well as detailed data on certain semantic domains. The treatment of such different systems and exotic scripts necessitates a flexible structure of the database and specific solutions developed in collaboration between linguists and computer specialists with experience in this field. Thus, in the course of the Action, software tools will be developed or adapted according to the needs of the researchers.

In its long development, mankind has brought forward a range of classification systems, varying according to their purposes, the constraints they are subject to and the way they reflect the internal organisation of knowledge or perception of the world.

In this Action, the systems will be preliminarily subdivided into two categories. On the one hand, we have planned and systematized knowledge systems such as the semantic organisation of reference works with an explicitly classifying function and very often a certain hierarchical depth (relying primarily on class inclusion). On the other hand, we have more or less naturally grown systems of classification such as numeral classifiers in languages and semantic determinatives in script which are often implicit and lack hierarchical depth.

These classification systems not only have different functions, but they are used by different – though overlapping – groups of people. Numeral classifiers are a linguistics device used with nouns in counting, in demonstrative or possessive constructions. They exist in many languages, for example in Sinitic languages, Japanese or Thai. Numeral classifiers for verbs occur with certain verbs when actions or events are counted. Verbal classifiers mark features of nouns on verbs. Semantic determinatives in logographic writing systems give clues to the semantic domain of a word, mostly in combination with one or more phonetic components. They occur in Chinese and Japanese characters, in Egyptian hieroglyphs in Sumerian cuneiforms, and probably also in Mayan script. In sign languages the “classifiers” are gestures indicating the meaning of a word.

Folk taxonomies can be observed in the everyday usage of language. They may be used even in topically arranged reference works, such as the Chinese *leishu* “classified” books, but also in Plinius’ *Historia Naturalis* and some historical encyclopaedias. The latter may also display organisation based on expert knowledge or specific philosophical, religious or political beliefs with a continuum to scientific taxonomy. A basic difference between folk taxonomies in spoken language and the topical arrangement of reference works is the internal structure of the domains. The linear text of a reference work with its fixed order gives possibilities to indicate relative closeness of domain members, rankings of importance or size and informal subgroups. Such additional parameters are also available in Chinese character dictionaries arranged according to “radicals” mostly derived from semantic markers.

The focus of the comparative work will be on East Asia, mainly China, because this culture uniquely provides practically all classification systems in question with a time span of records from up to 3.400 years. China is thus most suited to find the system-dependent features of classification systems given a “constant” culture and language. Chinese systems will then be compared to systems from other cultures.

Standardisation and Adaptation: A second part of the Action deals with the description of stabilising and adaptative factors in classification systems and the way they shape their evolution in certain contexts. The methodology will be quantitative or qualitative, depending on the nature of the data available. This part will make use of tools for modelling of the evolution of complex systems. Such factors may be universal, specific for certain languages or cultures or depending on the nature and functions of the system. System-specific components of classification systems have up to now hardly been studied and will be one of the focus issues of this Action.

One of the main functions of classification is to order the perceived phenomena so that they can be systematically thought and communicated about. This means that standardisation and stability belong to the key features of a classification system and the individual as well as the groups using it will make efforts to maintain the conventions. Standardisation appears especially desirable for scientist and in data processing (such as artificial language, automatic translation or semantic data mining). But a completely rigid system is not useful, as it does not allow to adapt to new developments in the environment itself or the attitude towards it. Nor would it be feasible to predetermine the sub-classification of every possible event, emotion or object with all the possible variation of details.

There are various forces of stability and standardisation at work in each system, such as psychological or cognitive factors, i.e. the stability of a learned category or cognitive universals, and political or cultural factors, i.e. the reinforcement of orthographic conventions for semantic markers in public schools, the insistence of parents on using the “correct” classifier, the transmission of scientific taxonomies or conventions of library codes or organisation of encyclopaedias. The way and degree to which they exert their stabilising influence on the different types of classification systems have to be determined.

As a counterpart to the stabilising forces there exist numerous sources for variation and evolution in classification systems, which have recently become the object of research. These can be subdivided into three groups: (1) Reasons in the “object world”: new, previously unclassified objects or the existence of ambiguous (ly perceived) objects such as bats (mammal/bird/insect) or whales (mammal/fish) in the animal world. (2) Variation in the choice of classification criterion by the user (e.g. *tiao* “stick” (criterion: shape: long/flexible) or *wei* “tail” (criterion: meronymic -> taxonomic) for fish in Chinese classifiers). (3) Contact with new languages or cultures/subcultures using a different classification. A special case of this could also be the propagation of new standards of orthography or knowledge organisation by the government, where the striving for synchronic and global standardisation and coherence leads to diachronic and local adaptation.

In addition, the interaction of these antagonistic forces are determined by system specific features and constraints. One example is the degree of stylisation in writing systems allowing for the creation of new “determinatives” and restricting the number of distinguishable “pictograms” for related objects. Another example is the

long term influence of a reference work transmitted for centuries as opposed to the short term effects of oral transmission in case of folk taxonomies or classifier usage.

Theory and Empirical Evidence: The essential of different models of classification will be described and their advantages and drawbacks for the evaluation of data from different systems will be discussed. Without a priori preference for specific models, different approaches will be made mutually understandable by clarifying definitions and underlying assumptions as well as their implications for the work on classification systems. Existing software tools will be compared and analysed in view of theoretical and descriptive requirements.

Those researchers doing descriptive work will present their data coming from a vast range of classification systems to each other and to experts of different theories and models, which often are only working on one or few language(s) or classification system(s). They will explain the linguistic, cultural and historic context of their data and the nature of their sources and raise the attention to specific features of the systems. Together, the Action members will establish adequate definitions and short descriptions that can be understood outside the discipline. Researchers working on models of classification will use the different type of data to test and elaborate their models.

Human categorisation and classification systems are dealt with by many disciplines in the humanities, information sciences and the life sciences. The existing research is marked by the construction of elaborate theories in different fields on the one hand and a vast amount of empirical and descriptive work on the other hand. Classification is a major issue in the philosophical field of ontology, in formal and cognitive semantics, in anthropology, with major inputs coming from neurobiology and findings on the function of learning and memory in the brain. The last century, especially since the seventies has witnessed controversies on the rule- or analogy-based nature of classification, also expressed as critical distinctive features vs. overall similarity, radial categories vs. componential analysis. Most recently, important models come from the information sciences trying to design or work with knowledge systems from different contexts.

The interdisciplinary work of this Action should lead to a more comprehensive and integrated model of human categorization and classification systems on the one hand and to theoretically more well informed descriptive work and manuals for fieldwork, experiments and philological work on the other hand.

D. Organisation

The Action will be subdivided into several Working Groups according to the topics and objectives defined in (B) and (C).

- (WG 1) Theory and empirical evidence
(possibly with a subgroup on software tools)
- (WG 2) Knowledge systems and explicit classification (reference works)
- (WG 3) Implicit classification in language and script
- (WG 4) Standardization and Adaptation

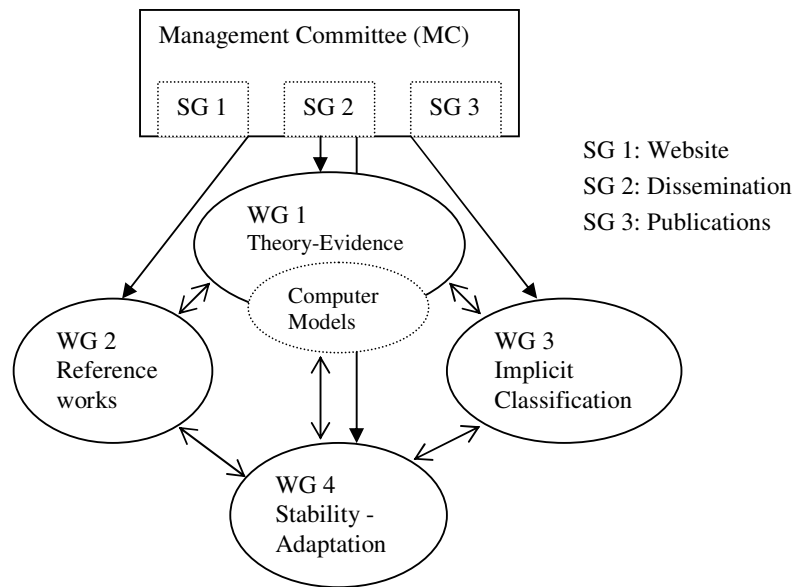
This subdivision may be subject to adjustments due to the focus of research interests of additional participants. The Working Groups will be organized during the initial

Management Committee meeting and may be subject to additional changes after the experience of the first working period.

The activities of the Action and the cooperation of the Working Groups will be coordinated by the Management Committee. The MC will set up small groups with special tasks such as maintaining the website, promoting the dissemination of results, organizing and editing publications.

The main occasions for presenting the results of the national research projects and the cooperation of the WGs will be the joint workshops and the final conference. WGs and MC will meet between workshops as indicated in (E). WGs will invite representatives from related international research programmes or attend their meetings.

Between the meetings, scientific cooperation and planning of activities will be maintained by mailing lists and telephone conferences. In the second half of the Action, some short term missions to partner institutions with special facilities or programmes will be arranged especially for younger participants. All participants will receive a regular electronic newsletter with information on activities and results of the Action and relevant developments in the field, to which all Working Groups will contribute.



E. Timetable

The Management Committee and the Working Groups will meet 1-2 times a year according to their needs. One of the meetings coincides with the workshop or conference.

The Action will be divided into 3 main phases, the preparatory Phase, the Working Phase and the Evaluation and Dissemination Phase. There will be an initial MC meeting, two workshops and one final conference to be organized during the course of the Action.

During the first phase, the Website will be established and the Initial Meeting prepared by the Management Committee. In this meeting, participants will report on their present research and the state of the art of their disciplines with respect to the Action. The meeting will also serve to form the Working Groups and agree on a more

detailed plan of activities for the Action. During the first phase, the MC will actively look for additional participants and work on initiating additional national or cooperative projects.

During the second phase, the Working Groups will cooperate on the defined research topics. They will – in particular – contribute to the Website of the Action as a forum of interdisciplinary exchange as outlined in C. The MC in collaboration with the Working Groups will organize the two workshops and the final conference. Each workshop and conference will have a session for each of the WGs with a focus topic, as well as a general session.

From the third year on, the MC will prepare the evaluation and joint publications of the results of WGs as well as other means of disseminating the results outside the scientific community involved. In this phase, short term missions, especially of young scholars will be organized.

START	WG 1		WG 1		WG 1		
Initial Meeting	WG 2	Workshop I	WG 2	Workshop II	WG 2	Final Conference	Final Report Conference Proceedings
	WG 3		WG 3		WG 3		
	WG 4		WG 4		WG 4		
MC	MC		MC		MC		
	Cooperation of Working Groups on defined domains Preparation of Workshops and Conference with topical focus				STSMs, Evaluation, Publications		

Year 1	Year 2	Year 3	Year 4
--------	--------	--------	--------

F. Economic dimension

The following COST countries have actively participated in the preparation of the Action or otherwise indicated their interest:

- Austria
- Denmark
- Croatia
- Finland
- France
- Germany
- Hungary
- Israel
- Italy
- Lithuania
- The Netherlands
- Norway
- Poland
- Romania
- UK

On the basis of national estimates provided by the representatives of these countries, the economic dimension of the activities to be carried out under the Action has been estimated, in 2004 prices, at roughly 22 Mio. Euro.

This estimate is valid under the assumption that all the countries mentioned above, but no other countries will participate in the Action. Any departure from this will change the total cost accordingly.

G. Dissemination plan (1/2-1)

The results of the Working Groups will be published on a regular basis in recognized journals and the proceedings of international conferences to make them available to the several scientific communities of the Action. Interdisciplinary studies will be especially encouraged.

The participants of the Action will present their results in international conferences of different disciplines. All of them, and especially the responsables of the Working Groups will make efforts to organize specific panels and subsections on topics concerning the Action in national and international conferences.

In order to make the results available to students and young scientists, the participants of the Action will offer classes at summer schools organized in disciplines included in this Action, e.g. ESSLi (European Summer School in Logic, Language and Information) of FoLLI.

Parts of Workshop 1 and Workshop 2, and the whole of the Final Conference will be open to non-participants. The main results of this Action, including the most important papers of the Conference will be published in a special monograph.

The Website entertained by the Action will constantly publish information about new results in the form of working papers or brief reports.

In order to make the results of this Action accessible to a more general public, participants will be encouraged to write articles for reference works, textbooks or journals addressed to a more general public. They will also engage in preparing working materials for modules on classification systems or human categorization and the relevant methods and theories to be used within university courses.

Finally, they will also cooperate with educational institutions and museums to organize special exhibitions, e.g. on the history of thought and changes in knowledge systems, comparative usage of categorization in script and signing. Results about stabilizing and adaptative factors in human categorization and classification systems in situations of culture and language contact will be made available to persons responsible for globalisation processes and international standardisation in international enterprises and relevant administrations.

COST A31

**“STABILITY AND ADAPTATION OF CLASSIFICATION
SYSTEMS IN A CROSS-CULTURAL PERSPECTIVE”**

ADDITIONAL INFORMATION

NOT PART OF THE MoU

Contents

1. The Three Dimensions of the Action and the Sequence of the Focus Issues
2. Preliminary Working Schedule
3. The Joint Database
 - Data collection, systematization and input
 - Structure of the database
4. Methodology and Sources for collecting data on classification systems
5. List of Experts
 - (A) Experts engaged in preparing the proposal and already agreeing to participate
 - (B) Potential participants not yet contacted
6. Selected Bibliography of the Participants

1. The Three Dimensions of the Action and the Sequence of the Focus Issues

As the treatment of the questions that form the steps towards the objective of the Action are dependent on the data available in the database and on the progress of the national research projects, the working programme of the Action will be organized as to allow an efficient treatment of topics of increasing complexity (see working schedule). The sequence of the research work is structured along three dimensions – *Typology of Classification Systems, Domain specificity and Stability and Adaptivity* with development of computational tools. In each dimension, one or two elements will form the focus of one year of the Action. The comparative work on relevant theories will support all these dimensions:

Typology: (1) Numeral Classifiers, (2) Semantic determinatives, (3) Reference Works

This sequence is motivated by the wealth of existing data and ongoing research of participants of the Action. A preliminary input of data on contemporary languages, drawn mostly from the literature, will give a fast overview and allow a test run of the database itself as well as software tools that are going to be developed in order to compare features of animal classification in different. In another approach, clustering and multidimensional scaling will be used to detect correlations between semantic and syntactic features in different subtypes of classifier systems and eventually discern subgroups of languages with similar classification strategies. This line of research will be carried out by integrating data from the AUTOTYP sub-project on "Attributionstypologie" (University of Leipzig). After evaluating this test-phase, the database will be revised, and data collected in fieldwork and from historical sources (see below) supplemented, to obtain more systematic results.

The subsequent work on semantic determinatives and reference works (see will result in analysis of specific features of categorization in each of these systems across cultures on the one hand, and in comparison of variation of strategies of categorization according to the functions and characteristics of different types of systems before the same cultural and linguistic background. The main work of developing software for these question lies in the heterogeneous nature of the data that have to made compatible. In the end of this sequence it will be possible to

establish a typology of the classification systems treated, with their defining features and parameters.

Domain Specificity: (A) Natural entities, (B) Artefacts, (C) Eventualities

The in-depth comparison of strategies of categorisation and development of classifier domains will be limited to three representative semantic domains within this Action. According to the findings of Tversky and Hemenway (1984) and Hirschfeld & Gelman (eds. 1994) on domain specificity, these domains are chosen to highlight diversity in classification of **natural entities** (e.g. animals or plants), **artefacts** and **abstract notions** or **eventualities** (i.e. actions, events and states). At least one domain will be selected from each of the three fields in the beginning of the Action.

Whereas for **animals** classification, there is a vast amount of literature, especially from ethnobiology (Brown 1984, Berlin 1992), and existing data on Chinese from the Wiebusch (forthcoming) as well as for hieroglyphs from a research project by Orly Goldwasser (University of Jerusalem), the latter two domains have been little studied from a cognitive point of view, so that new models and principles will have to be established in the course of the Action and data have to be specifically collected. **Artefacts**, but also in part eventualities, obtain a different status in categorization, as they are culture-specific and made by man himself, who also classifies them. Thus, there are no "cuts in nature" to be followed. Moreover, the prototypical elements of an artefact domain may very quickly according to fashion or technical innovation.

The domain of **eventualities** offers the opportunity to study a hitherto neglected subsystem of numeral classifiers used for verbs in Chinese and some other neighbouring languages, and its relation to the numeral classifiers used for nouns denoting eventualities. As a result of this sequence - in integration with the other dimensions of the Action - it will be possible to show how differences in the classified objects / concepts themselves - or in their cognitive status - influence the predominance of criteria of classification in single systems or universally.

Stability and Adaptation

The above two dimensions treat classification systems without specific regard to their historical evolution and interaction with other systems. But indeed, diachronic changes of all systems studied are remarkable, and among the parameters determining these changes are system internal as well as external factors. **Four major topics** are going to be treated in this dimension:

(a) the **diachronic evolution of single systems** and their **determining parameters**. For this purpose, quantitative and qualitative data on the core domains of each classification systems are going to be treated by a software, to be developed for this purpose, modelling its evolution and possible influence of changes in certain parameters on this system. Likewise, the likelihood of certain changes given additional features of the classifiers or systems will be calculated.

(b) The comparison of **diachronic development of different types of classification systems**:

Using the results of the first step, the Action will explore whether there are paths of change shared by two or more different types of systems, and which processes only apply in a certain type.

(c) **Effects of system contact:** Each type of system analysed within this Action has been exposed to contact with equivalent systems or been subject to borrowing. This concerns systems as a whole as well as single classifiers or their application to certain objects. This project will treat these borrowing effects in a systematic manner and apply recent computational tools such as cladistic approaches to phylogenesis, but also networks and splits to detect these effects. An especially interesting case will be the borrowing of Chinese classifiers to Korean and Japanese, with donor and receptor languages belonging to completely different linguistic types and historical data available in all three languages.

(d) **Comparison of synchronic data of Chinese "dialects" and neighbouring language families with diachronic data from Chinese:** The rich data on classifiers for Sinitic languages and their neighbours offer an unique opportunity to study **the interrelation of the development inside a language family and effects of borrowing on a classifier system**. Models and statistical tests will allow to make a very concrete comparison of synchronic and diachronic data and show the convergences and differences between these orthogonal plans. These four research topics will lead to a very comprehensive understanding of the evolution of different types of classification systems determined both by system-immanent parameters, system-external changes in the linguistic environment, cultural context and contact situations.

The preliminary working schedule is derived from the sequence of topics dealt with under the three dimensions as outlined above:

2. Preliminary Working Schedule

The main activities of the COST Action, i.e. the organization of workshops and conferences and the maintenance of the website have been outlined in Part I "Organization" and "Timetable". Although a final decision on the concrete working programme, the details of the focus issues and domains, will be taken by the MC and the WGs, the following represents the preliminary working schedule of the Action. With the integration of new theoretical models and software tools, a different approach might appear more feasible.

First meeting of MC

- Revision of work programme for Year 1,
- formation of subgroups responsible for Website, planning of meetings, dissemination of results and publications
- selection of additional participants
- preparation of Initial Meeting

Initial Meeting of participants

- formation and first session of Working Groups
- presentation of national research projects
- decision on exact parameters of database
- priority list of languages
- priority of semantic domains, minimum data included per language
- definite working programme for 1st year
- presentation of theoretical models to be discussed

Working Groups (Year 1-2):

- Focus topics:** **database**, clarification of **terminology** and **definitions**
Implementation of database (preliminary version) and data input
Manual for database use
Working paper/glossary with important **terminology and definitions** (taking into account different conventions across the disciplines, languages, classification systems and models involved)
Development and test-runs of **statistical tools**
- Focus system:** **numeral classifiers**
Exploration of **correlation** between **semantic** and **syntactic features** of numeral **classifiers**
- Focus domain:** **A: (e.g. animals)**
the **animal domain** in different Chinese classification systems
animal classification in **classifier** systems **across languages** (with preliminary data)
animal categorization in Chinese **script** and Egyptian hieroglyphs (preliminary)
- Focus theoretical model:** (to be determined)

Workshop I

- Evaluation of database, suggestions for revisions (if necessary)
- presentation of preliminary results
- Work programme for 2nd and 3rd year (MC meeting)
- Design of questionnaires, checklists and word lists for collecting and supplementing data on additional languages

Working Groups (Year 2-3):

- Focus topic:** **modelling diachronic evolution**
of classification systems; likely paths of change development and test of programs
- Focus system:** **semantic determinatives**
formal and semantic similarities and discrepancies in Chinese, Egyptian, Sumerian and Mayan writing systems
- Focus domain:** **B (e.g. utensils)**

human categorization of artefacts in different classification systems

Focus theoretical model (to be determined)

Workshop II

- presentation of results
- evaluation of computer tools
- Planning of publications, dissemination of results and STMs, (MC meeting)

Working Groups (Year 3-4):

Focus system: reference works, CL for verbs, eventually sign languages/ script

Focus domain: C (e.g. eventualities)
relation between CL for verbs and CL for nouns designating eventualities

Focus topics: effects and modelling of **culture and language contact**, borrowing
adaptation of classification systems in **globalization processes and international standardisation**
domain specificity of categorization
extension of diachronic models to **multilingual systems**

Synthesis: **Typology of classification systems**

Final Conference

- Presentation of results
- Public access to database
- Evaluation of Action and planning of dissemination of results and further research and collaboration (MC)

Working Groups and Subgroups of MC:

- Publication of dictionaries/reference works on classifiers
- Publication of results
- Final evaluation

3. The Joint Database

The basic tool for comparing the results of the national research projects is a joint database. As the data come from a variety of languages and sources and belong to different types of classification systems, the first step consists of establishing a set of features and a standard for data collection and data input that make data comparable across important dimensions and will allow further analysis by computational tools. This coordinating work will be done by the Working Groups and the MC. Collection and input of data will be carried out within national research projects. As not all participants are known yet, the following can only be a preliminary plan for the first three years.

- Data collection, systematization and input

As described above, there will be a minimum number of three semantic domains, one from nature, one from artefacts and one from abstract notions or eventualities (semantic domains A, B, and C). The MC may decide to include more domains, if this seems necessary or feasible. Individual participants or national research projects can choose to add semantic domains optionally. In order for the WGs to do analysis and obtain results on the focus topics fixed in the Work programme, the input of data has to be organized serving the sequence of topics defined above. The order of data input also takes into consideration the availability of data from previous projects and the literature and the time necessary for collecting supplementary data.

Thus, the general procedure of the work will be guided by the following principles:

1. Selected languages first – additional languages later
2. Input of existing data (from own projects or the literature) first – supplementing with specifically collected data later
3. Contemporary or very well documented stages first – historical or badly documented stages later
4. Overview of the system and formal features first – fine structure later
5. Always proceed from semantic domain A to C
6. Additional semantic domains or particular features optional

Year 1-2

Data systematization and input on:

- **Classifiers** in Contemporary Chinese, Japanese, Korean (Syntax and semantic domains A-C)
- Contemporary Chinese dialects (from literature and data from Chung-cheng University project, syntax and semantics A-C)
- Himalayan and other South-East-Asian languages (from literature, syntax and semantics A-C)
- Historical stages of Chinese (semantics A: animals)
- **Semantic determinatives** in Chinese script (from selected dictionaries, semantics A)
- topically arranged **reference works** (Chinese) (selected sources, semantics A)
- **sources** for classifiers and other data
- relevant **literature**

Year 2-3: Data collection, systematization and input on::

- **Classifiers** in selected contemporary languages (supplementing data of year 1 according to evaluation meeting)
- in historical periods of Japanese and Korean
- and of Chinese (semantics B)
- **Semantic determinatives:**
 - of Chinese script, formal features, semantics B, supplement to semantics A from additional sources
 - of Egyptian hieroglyphs (formal features, semantics A and B)

- of Sumerian cuneiforms (formal features, semantics A and B)
 - of other writing systems (e.g. Mayan script)
 - topically arranged **reference works** (Chinese)
- semantics B (from selected sources)

Year 3-4: Data collection and input on:

- **Classifiers** in Altaic languages (syntax and semantics A-C, from literature)
- in Chinese (historical, semantics C)
- - Chinese dialects, historical (syntax and semantics A-C)
- additional languages, e.g. Amerindian languages, according to expertise of participants
- **phonetic reconstruction** of classifiers in different languages (if possible)
- **numeral classifiers for verbs**
- Sinitic languages, some South-East-Asian languages, syntax and semantics A-C
- **Semantic determinatives:** in Chinese script, semantics C
- in Egyptian hieroglyphs, Sumerian cuneiforms, Mayan hieroglyphs, semantics C
- topically arranged **reference works** (Chinese)
- (formal features, semantics C (from selected sources), additional sources for semantics A and B
- reference works from other cultures
- **additional classification systems**, such as sign languages/script and folk taxonomies

This preliminary Plan will vary according to the participants of the Action and focus topics of the Working Groups. For the first year, the Action can rely on data already collected for some Chinese classification systems (see below) to create a test-version of the database.

A database of different types of classification systems is currently being designed by a project “” with the Centre de Recherches linguistiques sur l’Asie Orientale and the Laboratoire Dynamique du Langage (DDL) as main project partners. At present, this database is mainly conceived for East Asian languages, but is kept flexible enough to include data on other languages and writing systems. The structure of the database is outlined below:

- The Design of the Database

The **database** will be designed in consideration of the AUTOTYP-databases in Leipzig/ Berkeley. It will be in *filemaker pro*, allowing online data input and integration of tables in more generally available formats, such as Excel. The treatment of different scripts and rare Asian characters will be insured by using an extended version of Unicode, and a typeface with more than 70.000 characters available at the CRLAO. For the integration of Egyptian and Mayan hieroglyphs as well as Sumerian cuneiforms the experts for the digital treatment of these writing systems will cooperate in a Working Group.

The database will consist of **several modules:**

1. Domains of classifiers/determinatives/superordinate terms

2. **General and formal features** of the respective classification systems
3. **Phonological/lexical information** on classifiers, determinatives etc.
4. **Sources** and
5. **Literature** modules - shared by all classification systems and languages

Modules (1) – (3) are designed specifically according to the parameters of each type of classification system. The following is an example for numeral classifiers:

1. Data tables on the **domains of numeral classifiers**:

For the nouns in the domains chosen:

- | | | |
|--|---------------------|---|
| • language (stage), | • source | • possible classifiers |
| • Chinese/ Japan/
Korean characters | • type of source | • semantic relation between CL and noun |
| • meaning | • dating | • information on usage (style, discourse function) if available |
| | • quantitative data | |

These data have to be linked to the language internal structure of the relevant lexical domain and prepared for carrying out cross-linguistic searches for similar objects/ concepts taking into account physical, functional and conceptual differences of seemingly equal objects in different languages and cultures.

2. Data sheets on **general or syntactic features** of the classifier system in each language

for each language:

- | | | |
|---|--|--|
| • definition(s) of CL, | • number of different CL, | • relation between CL and measure words (MW), distinguishing criteria/tests |
| • obligatoriness, | • main semantic distinctions, | • information on related particles/ structures in the languages (e.g. number marking, subordinative particles, demonstratives) |
| • nouns used with/ without CL | • existence and usage of default CL | |
| • standard noun phrase(s) with CL, word order | • existence of auto-classifiers | |
| • usage with demonstratives, possessives | • existence of verb CL, syntactic properties | |

3. Data sheets with **phonological and lexical information** on CL in each language

- for each classifier:

- | | |
|--|---|
| • pronunciation | • cognate classifiers in other languages |
| • reconstruction of earlier stages of the language (if possible) | • meaning (if possible) |
| • source language(s) (if borrowed) | • general information about domain, semantic or syntactic peculiarities |

4. Database on **sources**:

The database on sources will contain information on the type of source (e.g. genre), its way of transmission, earliest edition, dating(s), the author, his native region, observations about the nature of the language (dialect, style).

5. Database on **literature**:

Analogous modules will be created for other types of classification systems.

4. Methodology and sources for collecting data on classification systems:

In order to decide on the concrete structure of the database, the parameters to be included and possible difficulties, it is important to understand the specific characteristics of each type of classification system, and to determine which information on the systems itself and on historical and cultural background will have to be included in the database.

Systems of classification and relevant features:

Three types of classification systems **are in the focus of the Action:** numeral classifiers, semantic determinatives in writing systems and topically arranged reference works. Other classification systems, such as folk taxonomies, radicals used for arranging character dictionaries and semantic elements in sign languages/script are also taken into account.

Among the three core classification systems, one is primarily part of the spoken language, two exist only in written form. Two of them are implicit systems, that are not systematized, whereas one reflects intentional and systematic organization of concepts. All systems are represented in the Chinese language. This allows a comparison of features across systems keeping the linguistic and, in part, the cultural background constant (for diachronic variation, see below). None of these classification systems is unique to Chinese, though. This allows to compare certain features along different languages within the same system.

Examples for classification systems in other languages/cultures

- | | |
|---|---|
| • Numeral classifiers | Japanese, Korean, Tai, Vietnamese, Burmese, Indonesian, various languages from the Pacific-Rim area |
| • Semantic determinatives in writing systems | Japanese (borrowed from Chinese), Egyptian hieroglyphes, Mayan script, Sumerian cuneiforms |
| • Semantic organization of reference works | e.g. Arabian, Latin, Japanese, French |
| • Superordinate – subordinate terms (folk taxonomies) | Practically all languages |
| • Radicals | Japanese (borrowed from Chinese) |
| • Semantic elements in sign languages/script | e.g. English, French, German, Chinese sign language and many others |

Numeral classifiers (CL) form a class of words which are used in some languages when a noun is used with a numeral. They exist in a large number of languages with different syntactic properties (Aikhenvald 2000). Here. Mandarin Chinese will serve as an example: The standard word order is:

(Dem) Num (Mod.) N - e.g. Zhe liang BEN shu – Dem two CL book – these two books.

There are several dozens of different classifiers, which are selected according to the meaning of the noun, thus rendering noun classes, which are neither historically stable nor synchronically clear cut, due to possible variation of classifier assignment. Mandarin Chinese, as many other languages, has a default CL for unmarked objects,

also used to replace more specific CL in certain discourse contexts. Other languages have echo- or auto classifiers repeating all or part of the noun. Typologically speaking, numeral classifiers form an areal phenomenon spread across East-South-Asia, as well as some adjacent regions. The word order of Num, CL and N varies (see Greenberg 1972), and different languages display a different number of CL. Apart from the typical count construction, CL also appear in other linguistic contexts, such as possessive constructions, anaphoric constructions, without numeral, assuming the function of indefinite or definite article depending on the language.

Already in the seventies, some universals regarding the domains and selection criteria of numeral classifiers have been proposed, which were attributed to cognitive processes of human categorization. Major **criteria** of categorization have been found to be **extension/shape, flexibility/ material** (secondary), **taxonomic** (natural type or function) and **part-whole-relations** (Allan 1977, Tai 1994). In many subtypes of classifier systems, these relations are implicit and show many features of prototype classification, such as graded membership, fuzzy borders and chaining (see Lakoff 1987). The importance of different criteria varies among languages and changes diachronically (Erbaugh 1986).

Several **non-alphabetic writing systems** of the world make use of **semantic determinatives** to indicate the semantic domain of a word that it is used in, e.g. **Chinese characters and Egyptian hieroglyphs**. More than 90% of the Chinese characters are composed of one phonetic and one semantic element. The phonetic element roughly indicates the pronunciation of the character in question (being a grapheme which alone is or was pronounced similarly) and the semantic element, itself unpronounced in this character, though having a pronunciation when used independently, gives a hint to the semantic domain – a link which may have bleached in the course of time. In Chinese, these two elements, graphically not separable, form the character which usually represents a monosyllabic morpheme, in ancient Chinese often equal to a word. The characters containing the same semantic element thus form a class of words. E.g. the semantic determinative *ma* “horse”, is a semantic determinative in characters for horses (in general, of different colour or age), camels, donkeys, galloping, riding, saddles and so on.

The semantic determinatives of Egyptian hieroglyphs, though showing some structural differences due to peculiarities of the hieroglyph script and the different type of language they represent, display astonishing similarities in principles of categorization.

The use of **topically arranged reference works** probably predates that of dictionaries arranged according more formal criteria such as pronunciation. Such works - in China as in the occidental tradition - have a **double function**: their immediate practical purpose **is information retrieval** - in the case of Chinese *leishu* important information are quotations from the classics or other literary works containing the word in question. But in addition, many of these works **represent the structure of the world** or proper organization of knowledge in the eyes of their authors. Accordingly, many Chinese *leishu* are organized in the order of “Heaven/

weather”, “Earth”, “Man and his affairs” and “Objects/living beings”, but the concrete chapter headings their exact order and content depends on priorities of the period of their creation, the philosophical school of its author and the immediate purpose.

Specific features of such reference works in comparison to other systems of classification are: their **systematic approach to classification**, attempt for **balanced structure, clear boundaries** between classes and explicit naming of classes. Moreover, the written format requires and allows for **linear arrangement** expressing **relative importance** or **perceived similarity** of class members. As these works are composed by single authors, there is **no continuous development** to be observed, but the chapter headings and their respective entries certainly in part reflect the underlying continuous development of the Chinese lexicon and folk taxonomies.

Folk taxonomies may be described as the "unscientific" conceptions of members of a culture of the classification of objects in his environment. These folk taxonomies are represented in the lexicon in the form of superordinate and subordinate terms, but also as implicit knowledge about relatedness or class inclusion among concepts. It belonged to the basic works on human categorization to show that there exists a "preferred" or **basic level of categorization** of objects with high salience with maximal feature clustering and other relevant phenomena, that such basic objects were often conceived as organized around a prototypical member and such similarity counts more than abstract biological evidence.

Classifiers in sign(ed) languages are similar to semantic determinatives in writing systems in that they can occur with any word class and are visually perceived. Sign languages offer some **additional dimensions**, though, as they allow to make use of three-dimensional space and movements and allow the simultaneous usage of more than one signing element at the time. Some of these options are lost in **sign script**.

Sources available for each classification system

The availability and nature of sources is crucial for the implementation of the Action. The historical depth of data, their nature and reliability varies for each system and each language:

The **numeral classifier systems** belong to the most well documented systems across a wide variety of languages: For many **contemporary languages**, the data can be obtained from grammars, from classifier dictionaries and from a wealth of articles and theses.¹ For the purpose of this Action, these data have to be carefully examined and systematicized, in many cases probably to be supplemented by own field data to be elicited, as they are of a varying quality, and have been obtained by different standards. Although these data can be used to get a first overview across languages, the Action will develop an own standard for data entered into the database during its first meeting.

For historical data on numeral classifiers research is limited to the transmitted texts in the relevant languages and information in early grammars. This imposes temporal limits on the availability of data, but also leads to lack of data on certain

¹ For an overview about the publications compare Aikhenvald (2001) and Wiebusch (forthcoming)

semantic domains not treated in the sources. This constraint necessitates a careful choice of the core semantic domains to be studied as well as the usage of a very large text corpus.

The largest time depth for the study of numeral classifiers is available in Chinese: Excavated inscriptions date back to the 14th century BC (oracle bone inscriptions), transmitted texts to the 8th-10th century BC (*Shijing* and *Shujing*). From this period on until Late Modern Chinese, a rich text corpus has been established (see Peyraube (1988), Yang (1993), Wiebusch 2000) of works best representing the vernacular of their time. Most of these texts are now easily searchable as electronic resources, especially in the corpus of the Institute of Linguistics of the Academia Sinica in Taiwan, and that of Modern Chinese (ca. 10.-19. century) of the Academy of Social Sciences in Beijing, with whom the Action is going to cooperate. This long time depth allows to study the evolution of the Chinese classifier system from its earliest stage on to its development into a fully obligatory system.

The historical texts available for other languages only reach back to 7th century (Old Khmer), the 8th century (Japanese), or the 15th century (texts in Korean phonetic script). Many historical Japanese texts pose additional problems as they are often modelled on Classical Chinese and the pronunciation of the characters is not clear. This problem is to be treated by the subgroup on historical development and borrowing.

Data on Chinese dialects or Sinitic languages can be obtained from early Missionary grammars – mostly describing Min, Hakka or Cantonese - some of which date back to the 16th or 17th century. Additionally, texts in the diachronic corpus for Chinese can be evaluated according to their regional provenience. The reconstruction of regional variants of Mandarin, as well as the history of Sinitic languages is part of ongoing international research, e.g. by Hilary Chappell at Melbourne, with whom the Action will cooperate.

Writing systems with semantic determinatives are not as widespread as numeral classifiers, but data are more easily available: In principle every text is a possible source, as these elements are present in nearly every character or written word. For the Chinese writing system, the first sources thus date back to the 14th century BC and the development can be followed until modern times, for Egyptian hieroglyphs the data range from the 3rd millennium BC over a period of ca. 3.000 years.

For an analysis of the diachronic development, only excavated documents or originals should be accepted as sources, as in copying, texts were often adapted to a more modern orthography, which could imply addition, omission or change of determinative. In recent years, a large number of collections and dictionaries of characters from Chinese archeological sites have been published, that can serve as research tools. For a later stage, **Chinese character dictionaries** (Bottéro 1993) – the first one transmitted from ca. 120 AD (*Shuowen jiezi*) - can serve as additional sources. For Egyptian hieroglyphs, there is the Ägyptisches Wörterbuch at Berlin as well as a large amount of digitalized data from two Göttingen/Jerusalem projects on

determinatives in hieroglyphs (Thyssen foundation and VW foundation). Currently there is a comparative project on Egyptian hieroglyphs and Sumerian cuneiforms planned by Orly Goldwasser (Jerusalem) and Gebhard Selz (Wien), as well as one on Chinese characters and Mayan hieroglyphs by Nikolai Grube (Bonn), the results of which can be used by the Action.

In China, there is a long tradition of **topically arranged reference works** called *leishu*, the earliest being preserved from the 7th century (see Kaderas 1998). The model for another tradition of topically arranged dictionaries, the *Erya*, is dated to Qin or Early Han dynasty (ca. 3rd-1st c. BC). Other works with topical arrangement, which may be viewed as precursors of these traditions can be dated back even earlier. As a comparison, reference works from the occidental tradition such as Plinius' "Historia naturalis" or certain florilegia of the medieval period will serve as sources.

Folk taxonomies of animals and plants, but also other semantic domains have been studied by ethnobiologists, so that there is a wealth of descriptions for languages of native peoples being the object of anthropology. There has also been some comparative research regarding the establishment of terms for cognitively significant levels of hierarchy, such as life forms (see Brown 1984). For contemporary data on additional languages, it is possible, for a first approximation, to consult monolingual dictionaries and ask informants. Here, the influence of scientific taxonomy taught at school has to be taken into consideration.

Recently, research on **sign(ed) languages** has led to a wealth of publications even on "classifiers" in such languages (Emmorey (ed. 2003). A dictionary of Taiwan Sign Language as well as a comparative grammar of Taiwanese and Chinese CL are currently prepared at the Chung Cheng University of Chiayi, Taiwan. Moreover, dictionaries on various sign scripts are also being compiled.

5. List of experts

(A) Experts engaged in preparing the proposal and already agreeing to participate

*Experts recommended for MC membership is marked by an *.*

AUSTRIA

Prof. Wolfgang U. Dressler e-mail: wolfgang.dressler@univie.ac.at
Dept. of Linguistics, University of Vienna Tel.: +43-1-4277-41710
Berggasse 11 Fax: +43-1-4277-9417
A-1090 Vienna

Prof. Dr. Erhard Oeser e-mail: erhard.oeser@univie.ac.at
Institut für Wissenschaftstheorie Tel: 01/4277-47621 od. -47602
Universität Wien Fax: +43 1 4277-9476
Sensengasse 8/10
A-1090 Wien

Univ. Doz. Dr. Gerhard Budin e-mail: Gerhard.Budin@univie.ac.at
Institut für Wissenschaftstheorie Tel.: (+43-1) 4277-47623
Universität Wien
Sensengasse 8/10
A-1090 Wien

Prof. Dr. habil. Gebhard J. Selz * e-mail: gebhard.selz@univie.ac.at
Lehrkanzel für Altsemitische Philologie und Tel.: 01-4277-43410
Orientalische Archäologie Fax: 01-4277-9434
Institut für Orientalistik
Universität Wien
Spitalgasse 2, Hof 4
A-1090 Wien

CROATIA

Mario Brdar (Associate Professor) e-mail: ritamario@dravanet.hu
Department of English Tel.: 031 / 211 - 400
Faculty of Philosophy Fax.: 031 / 212 – 514
University of Osijek
Lorenza Jägera 9
Osijek

DENMARK

Prof. William B. McGregor e-mail linwmg@hum.au.dk
Afdeling for Lingvistik Phone +45 8942 1111
Aarhus Universitet Direct phone +45 8942 6559
Nobelparken
Jens Chr. Skous Vej 7
8000 Århus C

Paul John Frandsen E-mail: pjf@hum.ku.dk
Egyptology Tel.: 353 28904
University of Copenhagen

FINLAND

Steve Legrand

Dept. of Computer Science and Information
Systems
University of Jyväskylä
PO Box 35
SF-40351 Jyväskylä
FIN-40014

Finland

email: stelegra@cc.jyu.fi

Tel.: +358 14 260 3010 (office)

+358 14 260 3024 (amanuensis)

Fax: +358 14 260 3011

FRANCE

Chappell, Hilary (Prof.)

Centre de recherches linguistiques sur l'Asie
Orientale (CRLAO) (CNRS/EHESS)
54 Bd Raspail
75006 Paris

e-mail: chappell@ehess.fr

Tel.: 0033-1-49542418

Prof. Colette Grinevald

Laboratoire Dynamique Du Langage
UMR CNRS 5596
Institut des Sciences de l'Homme
14 avenue Berthelot
69363 Lyon cedex 07
France

Colette.Grinevald@univ-lyon2.fr

Nishio, Sumikazu (PhD)

INALCO (Département Langue et civilisation du
Japon)
Centre Dauphine,
Place du Maréchal de Lattre de Tassigny
75116 Paris

e-mail: sumikazu.nishio@ish-lyon.cnrs.fr

Mahe Ben Hamed (PhD)

Laboratoire Dynamique Du Langage
UMR CNRS 5596
Institut des Sciences de l'Homme
14 avenue Berthelot
69363 LYON CEDEX 07

e-mail: Mahe.Ben-Hamed@ish-lyon.cnrs.fr

Christophe Coupé (PhD)

Laboratoire Dynamique Du Langage
UMR CNRS 5596
Institut des Sciences de l'Homme
14 avenue Berthelot
69363 LYON CEDEX 07

e-mail: ccoupe@ish-lyon.cnrs.fr

Tel.: +33- 4 - 72 72 64 63

Fax: +33 - 4 - 72 72 65 90

Prof. Georges Métaillé

Centre Alexandre-Koyré Histoire des sciences et
des techniques (CAK), EHESS
Muséum national d'histoire naturelle - Pavillon
Chevreul
57 rue Cuvier - 75231 Paris Cedex 05

Tel.: 01 43 36 70 69

Fax: 01 43 31 34 49

e-mail: metaille@mnhn.fr

Thekla Wiebusch *
Centre de recherches linguistiques sur l'Asie
Orientale (CRLAO) (CNRS/EHESS)
54 Bd Raspail
75006 Paris

e-mail: wiebusch@rz.uni-leipzig.de
Tel.: 0033-1-49542403

GERMANY

Prof. Dr. Balthasar Bickel
[Institut für Linguistik](#)
Universität Leipzig
Beethovenstraße 15
04107 Leipzig

E-Mail: bickel@uni-leipzig.de
Tel. +49-341-973 7604
+49-341-973 7610 (secretary)
Fax +49-341-973 7609

Prof. Walter Bisang *
Institut für Allgemeine und Vergleichende
Sprachwissenschaft
Johannes Gutenberg-Universität Mainz
FB 1420
55099 Mainz

Tel. Uni.: +49(0)6131/ 39-2-2778
Fax Uni.: +49(0)6131/ 39-2-3836
Tel. Priv.: +49(0)6131/ 34617
E-mail: wbisang@mail.uni-mainz.de

Prof. Dr. Nikolai Grube *
Institut für Altamerikanistik und Ethnologie
Universität Bonn
Römerstraße 164
53117 Bonn

Telefon: 0228 / 73-4412
Fax: 0228 / 73-4385
ngrube@mail.utexas.edu

Prof. Dr. Gordon Whittaker
Philosophische Fakultät
Seminar für Romanische Philologie
Kulturanthropologie und Altamerikanistik
Georg-August-Universität Göttingen
Humboldtallee 19
37073 Göttingen

Tel.: +49 551 39 4188
E-mail: gwhitta@gwdg.de

HUNGARY

Hrisztalina Hrisztova-Gotthardt
PhD-School of Applied Linguistics
Department of Philosophy
University of Pécs
Olga utca 1.X.32
7632 Pécs

E-Mail: xpucu@freeweb.hu
Telefon (privat): +36/20/4904033

Dr. Imre Szigeti *
Germanistisches Institut
Katholische Peter-Pazmany-Universität
2087 Piliscsaba
Egyetem u. 1.

Tel.: +36/26-375-375
Fax: +36/26-374-570
e-mail: imros@btk.ppke.hu

Rita Brdar-Szabó
Institute of German Studies
Faculty of Philosophy
ELTE
Budapest

e-mail: ritamario@dravanet.hu

ISRAEL

Prof. Orly Goldwasser *
Department of Egyptology
The Hebrew University of Jerusalem
Jerusalem 91905

e-mail: goldorly@hotmail.com

Arlette David (PhD)
Lecturer
Department of Egyptology
The Hebrew University of Jerusalem
Jerusalem 91905

Email: msarlett@mscc.huji.ac.il
Fax: +972-9-8919793

Racheli Shalomi-Hen (PhD)
The Department of Bible, Archaeology and
Ancient Near Eastern Studies
M.B. 653
Ben-Gurion University of the Negev
Be'er-Sheva 84105

phone: 972-08-6461092 (secretary)
or 08-6461031 (administrator)
Fax. 972-08-6472913

ITALY

Prof. Carla Bazzanella
Dipartimento di Filosofia
Università degli Studi de Torino
Via Sant'Ottavio 20 – 10124 – Torino

Tel. +39 11 670.37.36
Fax +39 11 812.45.43
Tel. 0116703728
Fax 0118124543
carla.bazzanella@unito.it

LITHUANIA

Prof. Inesa Seskauskiene
Department of English Philology
Vilnius University
Universiteto g.5
LT-01513 Vilnius

e-mail: inesa.seskauskiene@flf.vu.lt.

NETHERLANDS

Gunter Senft
Max Planck Institute for Psycholinguistics
Wundtlaan 1
PB 310
6500 AH Nijmegen

Phone: +31 24 352 1274
Fax: +31 24 3521213
E-mail: Gunter.Senft@mpi.nl

NORWAY

Prof. Dr. Christoph Harbsmeier
Department of East-European and Oriental
Studies
University of Oslo
P.O.Box 1030 Blindern
0315 Oslo

Phone: (+47) 22 85 67 06
Fax: (+47) 22 85 41 40
E-mail: christoph.harbsmeier@east.uio.no

POLAND

Prof. Barbara Lewandowska-Tomaszczyk *
President of the Polish Cognitive Linguistics
Association
Department of English Language
Kosciuszki 65,
90-514 Lodz

e-mail: blt@kryisia.uni.lodz.pl/ blt@uni.lodz.pl

Jacek Walinski (PhD)
Head of Computer Linguistics Unit
Department of English Language
University of Lodz
Kosciuszki 65,
90-514 Lodz

e-mail: walinski@uni.lodz.pl

Piotr Pezik, MA
Automatic Semantic Clustering and Information
Retrieval
Department of English Language
University of Lodz
Kosciuszki 65,
90-514 Lodz

tel: (+48) 42 665 52 20
fax: (+48) 42 665 52 21

ROMANIA

Dan Cristea
Faculty of Computer Science
"Alexandru Ioan Cuza" University of Iasi
16, Gen. Berthelot St.
Iasi - 6600,
Romania

Phone: +40.232.201542
Fax: +40.232.201490
Email: dcristea@infoiasi.ro
URL: <http://www.infoiasi.ro/~dcristea/>

UNITED KINGDOM

Prof. Roy Ellen *
Department of Anthropology
Eliot College
University of Kent at Canterbury
Canterbury,
Kent CT2 7NS

Email: R.F.Ellen@kent.ac.uk
Tel (44-) (0) 1227 764000 extn. 3421
Fax (44-) (0) 1227 827289

Dr. Kasia Szpakowska
Lecturer in Egyptology
Dept. of Classics and Ancient History University
of Wales Swansea
Swansea, SA2 8PP, Wales

Tel.: 01792 513104
e-mail: K.Szpakowska@Swansea.ac.uk

Angela McDonald PhD,
University of Liverpool,
Egyptologist (Animal Classification)
58 Weavers Way, Tillicoultry,
Clackmannanshire,
Scotland FK13 6BD

Tel. 01259 752930
email: secquin2003@yahoo.co.uk

TAIWAN, R.O.C.

Prof. James Tai
Graduate Institute of Linguistics
Center for Research in Cognitive Sciences
160, San-Hsing, Ming-Hsing, 621, Chiayi

e-mail: lngtai@ccunix.ccu.edu.tw
Tel.: (05)2720411 (ext. 16651)
Fax: (05) 2720491

6. Selected Bibliography of the Participants

- Ahmad, K. and Rogers, M. (In preparation): Artificial Intelligence and Terminology: The Processing of Knowledge.
- Ahmad, K. & Rogers, M. (2001): The Analysis of Text Corpora for the Creation of Advanced Terminology Databases. In S.-E. Wright & G. Budin (eds.), *The Handbook of Terminology Management*. Amsterdam and Philadelphia: John Benjamins. (In Press).
- Ahmad, K., Vrusias, B., Ledford, A. (2001): Choosing Feature Sets for Training and Testing Self-organising Maps: A Case Study. *Neural Comput & Applic* (2001), 10(1), p.56--66.
- Archer, D., Rayson, P., Piao, S. Tony McEnery, A.M. (2004): 'Comparing the UCREL Semantic Annotation Scheme with Lexicographical Taxonomies', in Proceedings of the EURALEX-2004 Conference, Lorient, France, pp. 817-827.
- Atran, Scott (1990): *Cognitive Foundations of Natural History: Towards an Anthropology of Science*. Cambridge: Cambridge University Press.
- Atran, Scott (1987). "Origin of the Species and Genus Concepts: An Anthropological Perspective", *Journal of the History of biology*, 20(2), pp. 195-279.
- Atran, Scott; Estin, Paul; Medin, John C. D. (1997). "Generic species and basic levels: Essence and appearance in folk biology", *Journal of Ethnobiology*, 17(1): 17-43.
- Atran, Scott (1999). Folkbiology, in Robert Wilson and Frank Keil, Ed. *The MIT Encyclopedia of the Cognitive Sciences*, pages 316-317. MIT Press.
- Atran, Scott (1998). Folkbiology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences* 21:547-609.
- Atran, Scott (2003). A Metamodule for Conceptual Integration: Language or Theory of Mind?. *Behavioral and Brain Sciences*.
- Atran, Scott (2004a). Adaptationism for Human Cognition: Strong, Spurious or Weak?. *Mind and Language*.
- Atran, Scott (2004b). In God we trust: the evolutionary landscape of religion. Oxford University Press.
- Baker, P., Bontcheva, K., Cunningham, H., Gaizauskas, R., Hamza, O., Hardie, A., Jayaram, B.D., Leisher, M., McEnery, A.M., Maynard, D., Tablan, V., Ursu, C. and Xiao, Z. (2004): 'Corpus linguistics and South Asian languages: Corpus creation and tool development', *Literary and Linguistic Computing*, Volume 19, Issue 4, pp 509-524.
- Bisang, Walter (1993). "Classifiers, quantifiers and class nouns in Hmong". *Studies in Language*. 17.1, 1-51.
- Bisang, Walter (1999). "Classifiers in East and Southeast Asian languages: Counting and beyond." In: Gvozdanovic, Jadranka. ed. *Numeral types and changes worldwide*, 113 - 185. Berlin: Mouton de Gruyter.
- Bisang, Walter (ed. 2001). *Aspects of typology and universals*. Berlin: Akademie-Verlag.
- Branco, A., McEnery, A.M. & Mitkov, R. (eds.) (forthcoming): *Anaphora Processing: Linguistic, Cognitive And Computational Modelling: Selected Papers From Daarc 2002*, John Benjamins, Amsterdam.
- Budin, G. (1991a). The Application of terminology-based knowledge data bases in the humanities and the social sciences and its impact on research methods. In: BEST, H.; MOCHMANN, E.; THALLER, M. *Computers in the Humanities and the Social Sciences*. München/London/New York/Paris: Saur, 1991, p. 337 - 342.
- Budin, G. (1991b). The structure and role of specialized information in scientific and technical terminologies. In: Bock H.-H.; IHM P. (Eds.) *Classification, Data Analysis, and Knowledge Organization Models and Methods with Applications* (1991), Springer Verlag, p. 216-220.
- Budin, G. (1993) Knowledge Organization and Modelling of Terminological Knowledge. In: SCHMITZ, K.-D. [ed.]. *TKE'93: Terminology and Knowledge Engineering. Proceedings Third International Congress on Terminology and Knowledge Engineering, 25 - 27 August 1993*, Cologne, FRG. Frankfurt/M: INDEKS-Verlag, 1993, p. 1 - 7,
- Budin, G. (1994a) Organisation und Evolution von Fachwissen und Fachsprachen am Beispiel der Rechtswissenschaft. In Wilske, D. (Hrsg.). *Proceedings of XIV. VAKKI Seminar Voyri, Finland 12/13. February 1994*. Vaasa: University of Vaasa, S. 9-20
- Budin, G. (1994b). Some hypotheses about concept representations. In Proceedings of the 9th LSP-Conference Bergen, Norway, 1994
- Budin, G. (1994c): Trends in communication structures in science and research - a systemic view. In Proceedings of the 9th LSP-Conference Bergen, Norway, 1994

- Budin, Gerhard (1996). Wissensorganisation und Terminologie - Komplexität und Dynamik wissenschaftlicher Informations- und Kommunikationsprozesse. Tübingen: Gunter Narr Verlag (Forum für Fachsprachenforschung Band 28)
- Coley, J. D., Medin, D. L., Proffitt, J. B., Lynch, E., & Atran, S. (1999). Inductive reasoning in folkbiological thought. In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 205-232). Cambridge, MA: Bradford.
- Coupé, Christophe (2003) "De l'origine du langage à l'origine des langues: Modélisations de l'émergence et de l'évolution des systèmes linguistiques", Thèse de doctorat en sciences cognitives de l'Université Lyon 2. (Pdf - 10 Mo)
- Coupé, Christophe (2004), "A unified computer model for internal and external constraints in language evolution", Proc. of the 5th International Conference on the Evolution of Language, Leipzig, 31 March - 3 April 2004.
- Cristea, Dan, Nancy Ide, Daniel Marcu, and Mihai-Valentin Tablan (2000): Discourse Structure and Co-Reference: An Empirical Study, *The 18th International Conference on Computational Linguistics COLING'2000*, Luxembourg, July 31-August 4, 2000.
- Cristea, Dan (2000): An Incremental Discourse Parser Architecture, In D. Christodoulakis (Ed.) *Proceedings of the Second International Conference - Natural Language Processing - NLP 2000*, Patras, Greece, June 2000. Lecture Notes in Artificial Intelligence 1835, Springer.
- Cristea, Dan, Ovidiu Craciun and Cristian Ursu (1998): GLOSS - A Visual Interactive Tools for Discourse Annotation. In *Proceedings of the Workshop on Annotation Tools, ESSLLI'98*, Saarbruecken, August 1998.
- Cristea, Dan, Nancy Ide and Laurent Romary (1998): Veins Theory. An Approach to Global Cohesion and Coherence. In *Proceedings of Coling/ACL*, Montreal, August 1998.
- Cristea, Dan, Nancy Ide and Laurent Romary (1998): Marking-up Multiple Views Of A Text: Discourse And Reference. In *Proceedings of the First International Conference on Language Resources and Evaluation*, Granada, May 1998.
- David, Arlette (2000a), *De l'infériorité à la perturbation: l'oiseau du "mal" et la catégorisation en Egypte ancienne*. GOF IV. Reihe Ägypten 38. Classification and Categorization in Ancient Egypt 1. Wiesbaden: Harrassowitz.
- David, Arlette (2000b), *m.k + forme nominale au coeur de la controverse*. *Lingua Aegyptia* 7: 113-123
- David, Arlette (2003a), Review of Borchers, D., Kammerzell, F. & Weninger, S., Hieroglyphen Alphabete Schriftreformen. *Bibliotheca Orientalia* 60: 569-579
- David, Arlette (2003b). *Syntactic and Lexico-Semantic Aspects of the Legal Register in Ramesside Royal Decrees*, PhD Thesis, supervisor: Prof. Orly Goldwasser.
- David, Arlette (2004), Composés attributifs exocentriques, hypallage et métaphore, *Lingua Aegyptia* 12, in press (7 pp.)
- Dimitrova, Ludmila, Tomaz Erjavec, Nancy Ide, Heiki-Jan Kaalep, Vladimr Petkevic, and Dan Tufis (1998). Multext-East: Parallel and Comparable Corpora and Lexicons for Six Central and Eastern European Languages. In COLING-ACL '98, pp. 315-319.
- Ellen, Roy F. (1979): "Introduction". In: R.F. Ellen und D. Reason (Hg.): *Classifications in Their Social Context*. London, Academic Press, S. 1-32.
- Ellen, R. F. (1986): Ethnobiology, Cognition, and the Structure of Prehension: Some General Theoretical Notes. *Journal of Ethnobiology* 6: 83-98.
- Ellen, Roy F. (1993) The cultural relations of classification : an analysis of Nuaulu animal categories from central Seram. Cambridge [u.a.] : Cambridge Univ. Press, 1993
- Ellen, Roy F. (1998): Palms and the prototypicality of trees : some questions concerning assumptions in the comparative study of categories and labels. In: *The social life of trees*. - Oxford [u.a.] : Berg
- Erjavec, Tomaz, Dan Tufis, Tamas Varadi (1999). Developing TEI-Conformant Lexical Databases for CEE Languages. COMPLEX'99, pp. 205-209.
- Goldwasser, Orly (2002). Prophets, lovers, and giraffes: Wor(l)d classification in Ancient Egypt. *Göttinger Orientforschungen* 38. Classification and categorisation in Ancient Egypt,.3. Wiesbaden: Harrassowitz. (with an appendix by Matthias Müller)
- Goldwasser, Orly (2003). The building blocks of a conceptual category: The ancient Egyptian evidence. In L. Morra & C. Bazzanella (Eds), *Philosophers and hieroglyphs*. Rosenberg & Sellier: Torino.
- Goldwasser, Orly (forthcoming), Where is metaphor: Conceptual metaphor and alternative classification in the hieroglyphic script. *Metaphor & Symbol*.
- Goldwasser, Orly (forthcoming). A comparison between classifier languages and classifier script: The case of Ancient Egyptian, In G. Goldenberg & A. Shisha-Halevi (Eds.), *Ancient Egyptian, Neo-Semitic, methods in linguistics: Polotsky memorial volume*. Jerusalem : Magnes Press.

- Grube, Nikolai (2000). Monumentos esculpidos e inscripciones jerglíficas en el triángulo Yaxhá-Nakum-Naranjo. IN: El sitio maya de Topoxté: Investigaciones en una isla del lago Yaxhá, Petén, Guatemala, hg. von Wolfgang W. Wurster, S. 249-268. Mainz Philipp von Zabern.
- Grube, Nikolai (2000a). On Classic Maya Inscriptions. In: *Current Anthropology*, 41, Nr.5, S.837-838.
- Grube, Nikolai (2000b). Los gobernantes mayas. In: *Arqueología Mexicana*, Nr.50, S.72-77.
- Grube, Nikolai, David Stuart (2000). A New Inscription from Nim Li Punit, Belize. (= Research Reports on Ancient Maya Writing, 45). Center for Maya Research, Washington, D.C.
- Grube, Nikolai, Barbara McLeod, and Phil Wanyerka (1999), A Commentary on the Hieroglyphic Inscriptions of Nim Li Punit, Belize. (= Research Reports on Ancient Maya Writing, 41). Center for Maya Research, Washington, D.C.
- Haltmayer, Stephan [Hrsg.] , Oeser, Erhard und Durali, Teoman (2000) Homo pragmatico-theoreticus Zusatz z. Titel Philosophie - Interdisziplinarität und Evolution - Information ; Erhard Oeser zum 60. Geburtstag. Frankfurt am Main ; Wien [u.a.] Lang .
- Harbsmeier, Christoph (ed.): Thesaurus Linguae Sericae (TLS) □□□□□ An Historical and Comparative Encyclopaedia of Chinese Conceptual Schemes (A Construction Site on the Web). <http://www.sino.uni-heidelberg.de/database/tls/>.
- Harbsmeier, Christoph (1998): Language and Logic in Traditional China. Vol. 7.3 of Science and Civilisation in China (Cambridge: Cambridge University Press.
- Harbsmeier, Christoph ,with Lu Jianming and Ma Zhen. Modern Chinese Analytic Syntax. Vols. 1-2 (Oslo: Dept of East European and Oriental Studies, 1991-92)
- Harbsmeier, Christoph Comparative Studies in Latin and Classical Chinese Grammar and Rhetoric (Dept. of East European and Oriental Studies, 1993).
- Harbsmeier, Christoph (ed.) with Jiang Shaoyu (1999). Synonyma Serica Comparata. Dictionary of Classical Chinese Synonyms and Antonyms. Print-out of main results concerning 702 synonym groups.
- Hardie, Andrew, Piotr Pezik, and Eric Levin (2004). Overview of Part of Speech Tagging Systems.
- Nancy Ide, Dan Cristea, (2000): A Hierarchical Account of Referential Accessibility, accepted at *The 38th Annual Meeting of the Association for Computational Linguistics, ACL'2000*, Hong Kong, China.
- Legrand, Steve and JGR Pulido (2004): A Hybrid Approach to Word Sense Disambiguation: Neural Clustering with Class Labeling. Knowledge Discovery and Ontologies (KDO-2004) workshop in 15th European Conference on Machine Learning (ECML) and 8th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD) Pisa, Italy, September 24, 2004
- Legrand, Steve, P. Tyrvainen, and H. Saarikoski (2003), Bridging the Word disambiguation Gap with the help of OWL and Semantic Web Ontologies. Ontologies and Information Extraction International Workshop held as part of the EUROLAN 2003, Bucarest 28th July – 8th August 2003, pp. 29-35
- Legrand, Steve and P. Tyrväinen (2002): Connecting Lexical Entries to Distributed Real-World Knowledge, High Desert Linguistic Society V Conference, 1-2 November 2002, Albuquerque, New Mexico.
- Lewandowska-Tomaszczyk, B. (ed. 1993), Topics in English Lexicology, Lodz: Wydawnictwo Uniwersytetu Lodzkiego (Folia Linguistica 30).
- Lewandowska-Tomaszczyk, B. (1995 a) "Worldview and Verbal Senses", in Braj B.Kachru and Henry Kahane (eds.) Cultures, Ideologies, and the dictionary. Studies in honor of Ladislav Zgusta. Tübingen: Niemeyer [Lexicographica. Series Maior, Band 64], 223-235
- Lewandowska-Tomaszczyk, B. (1996) "Uniwersalizm a relatywizm w nowych teoriach kognitywnych", MateriaLy konferencji SWIATY ZA SLOWAMI, Lublin: UMCS.
- Lewandowska-Tomaszczyk, B. (1997) "Systemy reprezentacji jezyka w lingwistyce kognitywnej a ich realnosc psychologiczna" w Grucza, F. I M. Dakowska (eds.) Podejscia kognitywne w lingwistyce, translatoryce i glottodydaktyce. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego. 33-41.
- Lewandowska-Tomaszczyk, B. (1997) "Uncovering lexical meaning in language corpora". In: Lewandowska- Tomaszczyk, B. and P. J. Melia (eds.) Practical Applications in Language Corpora PALC'97 [Abstracts] [Proceedings]. Lodz: Lodz University Press.
- Lewandowska-Tomaszczyk, B. (1998 ed.) Lexical Semantics, Cognition and Philosophy, Lodz: Lodz University Press.
- Lewandowska-Tomaszczyk, B. (1998a) "Semantic representation in Cognitive Grammar" in Lewandowska-Tomaszczyk, B. (ed.) Lexical Semantics, Cognition and Philosophy. Lodz: Lodz University Press.

- Lewandowska-Tomaszczyk, B. (1998b) "Semantic analysis of the concepts of nationality and citizenship in Polish and English". In G. Wotjak (ed.) *Lexicology To-day*. Hamburg: Peter Lang.
- Lewandowska-Tomaszczyk, B. and P. J. Melia (eds. 1997) *Practical Applications in Language Corpora PALC'97 [Abstracts] [Proceedings]*. Lodz: Lodz University Press.
- Lewandowska-Tomaszczyk, B. and P. J. Melia (eds. 1997) *Practical Applications in Language Corpora PALC'97 [Proceedings]*. Lodz: Lodz University Press.
- Lewandowska-Tomaszczyk, Barbara (1990 a) Coleman, D.W., Lewandowska-Tomaszczyk, B.(eds), *Using Statistics for Language Studies*, Lodz: Wydawnictwo niwersytetu Lodzkiego, (Folia Linguistica 24).
- Lewandowska-Tomaszczyk, Barbara (1990 b) Tomaszczyk, J., Lewandowska-Tomaszczyk, B. (eds), *Meaning and Lexicography*, Amsterdam: John Benjamins.
- MacGregor, William B. and Hilary Chappell (eds. 1995): *The grammar of inalienability: a typological perspective on body parts and the part-whole relation*. Berlin: Mouton & deGruyter.
- MacGregor, William B. (1997): *Semiotic grammar*. Oxford: Clarendon Press.
- MacGregor, William B. (2002): *Verb classification in Australian languages*. Berlin & New York: Mouton de Gruyter.
- Marsico, E., C. Coupé and F. Pellegrino (2002), "Evaluating the influence of language contact on lexical changes", *Proc. of the 3rd Conference on the Evolution of Language*, Ecole Nationale Supérieure des Telecommunications, Paris, April 2000, p.154-155.
- Marsico, E., F. Pellegrino and C. Coupé (2002a), "A semiotic approach to the origin of language", *Proc. of the 4th International Conference on the Evolution of Language*, Harvard, 27-30 March 2002, p.73.
- McEnery, A.M. & Wilson, A. (2001): *Corpus Linguistics* (2nd edition), Edinburgh University Press, Edinburgh.
- McEnery, A.M. & Xiao, Z. (2002): 'Domains, text types, aspect marking and English-Chinese translation', *Languages in Contrast*, Volume 2, Issue 2, pp 51-69.
- McEnery, A.M., Tono, Y & Xiao, Z (in press): *Corpus Based Language Studies*, Routledge, London
- McEnery, A.M. (forthcoming): *Corpora in Linguistics*, CUP, Cambridge.
- Medin, D.L & Atran. S. (Submitted). *The Native Mind: Biological Categorization, Reasoning and Decision Making in Development Across Cultures*. *Psychological Review*.
- Medin, D.L, Ross, N., Atran, S., Cox, D., Wakaua, H.J., Coley, J.D., Proffitt, J.B. & Blok, S. (Submitted). *The Role of Culture in the Folkbiology of Freshwater Fish*. *Cognitive Psychology*.
- Métaillié, Georges, 1988. "Des mots et des plantes dans le *Bencao gangmu* de Li Shizhen". *Extrême-Orient Extrême-Occident*, 10, pp. 27-43.
- Métaillié, Georges, 1989. "Histoire naturelle et humanisme en Chine et en Europe au XVIe siècle: Li Shizhen et Jacques Dalechamps", *Revue d'Histoire des Sciences*, 42(4): 353-374.
- Métaillié, Georges, 1992. "Des mots, des animaux, des plantes", *Extrême-Orient Extrême-Occident*, 14: 'Regards obliques sur l'argumentation en Chine', pp. 169-183.
- Métaillié, Georges, 1995. "Note à propos de citations implicites dans les textes chinois", *Extrême-Orient Extrême Occident*, 17: 131-139.
- Métaillié, Georges, 1997. "Ethnobotany in China", pp. 312-315, in: Selin, Helaine (ed.). *Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Métaillié, Georges, Hsu, Elisabeth, forthcoming. "Li Shizhen: an innovator in natural history?", in: Hsu Elisabeth (ed.). *Chinese medicine: Innovation, convention and controversy*. Cambridge: Cambridge University Press.
- Nishio, Sumikazu, 2000, « Classificateurs Numériques en Japonais : Constructions et Catégories », thèse de doctorat, Université Lumière Lyon 2, Lyon.
- Nishio, Sumikazu, 2001 « *Classificateurs numériques en japonais : constructions et catégories (Résumé de thèse)* », *Cahier de Linguistique Asie Orientale*, Vol.30, n°1, Paris.
- Oeser, Erhard (1994). *Terminology and philosophy of science*. Vienna, International Terminological Conference Terminology Science and Terminology Planning in Commemoration of E. Drezen (1892 - 1992), S. 24 - 34 :
- Oeser, Erhard (1996). *System, Klassifikation, Evolution . Historische Analyse und Rekonstruktion der wissenschaftstheoretischen Grundlagen der Biologie*. Wien, Braumüller .
- Oeser, Erhard (1997a) *Die Repräsentation von Objekten in Terminologie und Wissenstechnik Beiträge zur Terminologie und Wissenstechnik*. Vienna, S. 58 – 76.
- Oeser, Erhard (1997b). *Terminologische Wissenstechnik . Beiträge zur Terminologie und Wissenstechnik*. Vienna, S. 183 – 188.

- Oeser, Erhard (1988): *Gehirn, Bewußtsein und Erkenntnis* / Erhard Oeser u. Franz Seitelberger . - Darmstadt : Wiss. Buchges.
- Oeser, Erhard (1997). *Computerunterstützte Wissenschaftstheorie am Beispiel einer Multimedia-Datenbank über historische Erdbeben-theorien (HEAT)*. Beiträge zur Terminologie und Wissenstechnik. Vienna, S. 189 – 199.
- Pezik, Piotr (2001). *Ontological and Statistical Keyness in Knowledge Processing*
- Pezik, Piotr, Rafal Uzar, and Eric Levin, 2004. *Using Relational Databases for Corpus Linguistics*
- Poitou, Jacques (2004a): Prototypentheorie und lexikalische Semantik. = *LYLIA* 3 (2004).
- Poitou, Jacques (2004b): Prototypentheorie und Flexionsmorphologie. *Linguistik online* 2004.
- Poitou, Jacques (2003): Polysémie et catégorisation en chaîne. in : Rémi-Giraud, Sylvianne & Panier, Louis, *La polysémie ou l'empire des sens*. Lexique, discours, représentations. Lyon : PUL. Pages 29-38.
- Poitou, Jacques (2002a): Les "normes catégorielles" : catégories sémantiques et/ou listes de mots. Structuration des catégories sémantiques entre le cognitif et le linguistique. (with Danièle Dubois). *Intellectica* 2: 217-249.
- Poitou, Jacques (2002b): Normes catégorielles pour 22 catégories sémantiques du français et 10 catégories sémantiques de l'allemand. (en collaboration avec Danièle Dubois). *Cahiers du LCPE* 5: 35-118.
- Poitou, Jacques (2000a): Conceptions prototypiques et articulation du changement diachronique. Le changement linguistique. Evolution, variation, hétérogénéité. *Actes du colloque de Neuchâtel, 2-4 octobre 2000*. = *TRANEL* 34-35 : 101-115.
- Poitou, Jacques (2000b): Prototypes, saillance et typicalité. *Nouvelles terminologies* 21 (2000) : 17-26.
- Poitou, Jacques (2000c): Was ist Komposition ? - Zur Unterscheidung zwischen Komposition und Derivation. *Cahiers d'études germaniques* 39: 239-250.
- Poitou, Jacques (1999): Catégories sémantiques et cognitives. Une étude expérimentale en sémantique lexicale. (with Danièle Dubois). *Cahiers de Lexicologie* 74, 1 : 5-27.
- Rayson, P., Archer, D., Piao, S. and McEnery, A.M. (2004): 'The UCREL semantic analysis system' in *Proceedings of the Beyond Named Entity Recognition Semantic Labeling for NLP Tasks Workshop*, Lisbon, Portugal, pp 7-12.
- Riggs, Fred; Budin, Gerhard, Matti Mälkiä (in press). *Descriptive Terminology Work..* In: Wright, Sue Ellen; Budin, Gerhard. *Handbook of Terminology Management*. Amsterdam/ Philadelphia: John Benjamins.
- Selz, Gebhard J. (2000): *Schriifterfindung als Ausformung eines reflexiven Zeichensystems*. In: *Wiener Zeitschrift für die Kunde des Morgenlandes* 90 (2000), pp. 169-200.
- Senft, G. 1990. "Apropos "The Whole and its Parts": Classificatory Particles in Kilivila Language", in: Walter A. Koch, ed. *Das Ganze und seine Teile - The Whole and its Parts*. Bochum Publications in Evolutionary Cultural Semiotics, BPX 19. pp. 142-176. Bochum: Brockmeyer.
- Senft, G. 1991. "Network Models to Describe the Kilivila Classifier System". *Oceanic Linguistics* 30: 131-155.
- Senft, G. 1993. "A Grammaticalization Hypothesis on the Origin of Kilivila Classificatory Particles". *Sprachtypologie und Universalienforschung* (vormals ZPSK) 46: 100-112.
- Senft, Gunter 1996. *Classificatory Particles in Kilivila*. Oxford Studies in Anthropological Linguistics. New York: Oxford University Press).
- Senft, Gunter 1998. "Body and Mind in the Trobriand Islands". *Ethos* 26: 73-104. Papers in edited volumes:
- Senft, Gunter (2000 ed): *Systems of nominal classification*. (Language, culture and cognition; 4) Cambridge [u.a.] : Cambridge Univ. Pr..
- Senft, Gunter (2000a): What do we really know about nominal classification systems? In: Senft, Gunter (ed): *Systems of nominal classification*. (Language, culture and cognition; 4) Cambridge [u.a.] : Cambridge Univ. Pr., pp. 11-49.
- Senft, Gunter (in press). "Nominal Classification", in: Dirk Geeraerts and Hubert Cuyckens, eds.: *Handbook of Cognitive Linguistics*. Oxford: Oxford University Press.
- Szapkowska, K (1999). "A Sign of the Times." *LingAeg* 6: 163-6.
- Szapkowska, K (2000). "The Perception of Dreams and Nightmares in Ancient Egypt: Old Kingdom to Third Intermediate Period." Ph.D. Dissertation, University of California, Los Angeles.
- Szapkowska, K. (2003) *Behind Closed Eyes: Dreams and Nightmares in Ancient Egypt*. Swansea: The Classical Press of Wales.
- Tai, James (1994): *Chinese Classifier Systems and Human Categorization*. In: M. Chen and O. Tzeng (éds.): 479-494.

- Tai, James (1992): "Variation in Classifier Systems Across Chinese Dialects: Towards a Cognition-Based Semantic Approach". *Zhongguo jingnei yuyan ji yuanyue* 1: 587-608.
- Tai, James and Chao Fang-yi (1994): "A Semantic Study of the Classifier *zhang*". *Journal of the Chinese Language Teachers Association* 29: 67-78.
- Tai, James and Wang Lianqing (1990): "A Semantic Study of the Classifier *tiáo*". *Journal of the Chinese Language Teachers Association* 25.1: 35-56.
- Thouvenot, Marc (1997). L'écriture nahuatl. In *L'aventure des écritures*, 72-81. Paris: Bibliothèque nationale de France.
- Thouvenot, Marc (1999). *Tlachia-Pohua* : un outil pour réaliser des dictionnaires pictographiques mésoaméricains. Paris: Editions Sup-Infor.
- Thouvenot, Marc (2001). L'écriture aztèque : une écriture à découvrir. In *Histoire de L'écriture*, ed. A.M. Christin, 179-187. Paris: Flammarion.
- Thouvenot, Marc (2002a). *TEMOA* (en collaboration avec Paul Fisher). Paris: Editions Sup-Infor.
- Thouvenot, Marc (2002b). *CODEX XOLOTL, Etude d'une des Composantes de son Ecriture : Les Glyphes, Dictionnaire des Eléments Constitutifs des Glyphes*. Paris: Editions Sup-Infor.
- Tufis, Dan, H. Hamburger, R. Hashim, and J. Pan (1994) "Generation of Natural Language in an Immersive Language Learning System", in M.Browse (ed.) *Basics of Man-Machine Communication for the Design of Educational Systems*, Springer Verlag, 1994
- Tufis, Dan, and M. Zock (1997) – "Maximal Cohesion, Minimal Commitment and Lexical Pattern Matching", LIMS Research Report, December, 1997, 56p.
- Uzar, Raf, and Jacek Walinski (2001), A Comparability Toolkit, in (eds. B. Lewandowska-Tomaszczyk, Patrick James Melia) *PALC 1999 Proceedings*, Peter lang Publ.
- Uzar, Rafal, and Jacek Walinski (2001), Analysing the Fluency of Translators, in (ed. Wolfgang Teubert) *International Journal of Corpus Linguistics - Text Corpora and Multilingual Lexicography*, John Benjamins.
- Walinski, Jacek (2003), Lexical Acquisition of Terminology From Comparable Corpora, in (ed. B. Lewandowska-Tomaszczyk) *PALC 2001 Proceedings*, Peter Lang Publ., 2003.
- Walinski, Jacek (forthcoming), Acquisition of Terminology at the Phrase Level from Comparable Corpora, in (ed. B. Lewandowska-Tomaszczyk) *PALC 2003 Proceedings*, Peter Lang Publishing, forthcoming.
- Wiebusch, Thekla (1993) "The Role of Nominal Classifiers in Chinese and their Historical Evolution: a Case of Cyclic Change". *Faits de langues* 2, pp. 47-57 (En collaboration avec Alain Peyraube).
- Wiebusch, Thekla (1995) "Quantification and Qualification - Two Competing Functions of Numeral Classifiers in the Light of the Radical System of Chinese Script". *Journal of Chinese Linguistics* 23.2, juin, pp. 1-42.
- Wiebusch, Thekla (1998) "What Radicals Reveal about the Quantifying and Qualifying Function of Numeral Classifiers". B. T'sou, ed. *Studia Linguistica Serica*. Hong Kong, City University Press of Hong Kong. 1998, pp. 69-88.
- Wiebusch, Thekla (1999) "Classificateurs et clés graphiques en chinois". *Faits de langues* 14 (*La catégorisation dans les langues*), Paris, Ophrys. 1999, pp. 133-142.
- Wiebusch, Thekla (to appear). *Chinesische Numeralklassifikatoren für Tiere, eine diachrone Studie*. Wiesbaden, Harrassowitz, 2004.
- Whittaker, Gordon. (1991). *Calendar and script in protohistorical China and Mesoamerica : a comparative study of day names and their signs*. Bonn : Holos, 1991.
- Wilson, Andrew, and Piotr Pezik (2004). *Corpus Annotation Schemes*.
- Wright, Sue Ellen; Budin, Gerhard [Hrsg.] (in press). *Handbook of Terminology Management*. Amsterdam/Philadelphia: John Benjamins, 2 volumes.
- Xiao, Z., McEnery, A.M., Baker, J.P. & Hardie, A. (2004): 'Developing Asian language corpora: standards and practice', in *Proceedings of the 4th Workshop on Asian Language Resources*. March 25, 2004. Sanya, China. pp. 1-8.
- Xiao, Z. and McEnery, A.M. (forthcoming): 'Collocation, semantic prosody and near synonymy: a cross-linguistic perspective', *Applied Linguistics*.

Bibliographical References mentioned in Part I

- Adams, Karen L. (1982): *Systems of Numeral Classification in the Mon-Khmer, Nicobarese and Aslian Subfamilies of Austroasiatic*. Ph.D. Dissertation. The University of Michigan. University Microfilms International 8304433. Published as *Pacific Linguistics Series B - 101*, Canberra 1989.

- Aikhenvald, Alexandra Y. (2000): *Classifiers. A typology of noun categorization devices.* (Oxford studies in typology and linguistic theory) Oxford [u.a.] : Oxford Univ. Press. Berlin
- Berlin, Brent (1968): *Tzeltal Numeral Classifiers, a Study in Ethnographic Semantics.* Janua Linguarum, Series Practica 70. Den Haag, Mouton,
- ____ (1976): "The Concept of Rank in Ethnobiological Classification: Some Evidence from Aguarana Folk Botany". *American Ethnologist* 3.3: 381-399.
- ____ (1977): "Speculations on the Growth of Ethnobotanical Nomenclature". In: Blount u. Sanches (Hg. 1977): 62-102.
- ____ (1992): *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies.* Princeton, NJ: Princeton University Press.
- Berlin, Brent, Dennis E. Breedlove und Peter H. Raven (1969): "Folk Taxonomies and Biological Classification". In: Tyler (Hg. 1969): 60-65. Abdruck von *Science* Vol. 154: 273-275 (Okt. 1966).
- ____ (1973): "General Principles of Classification and Nomenclature in Folk Biology". *American Anthropologist* 75: 214-42.
- ____ (1968): "Covert Categories and Folk Taxonomies". *American Anthropologist* 70: 290-299.
- Brown, Cecil H. (1974): "Unique Beginners and Covert Categories in Folk-Biological Taxonomies". *American Anthropologist* 76: 325-27.
- ____ (1976): "General Principles of Biological and Non-Biological Folk Classification". *American Ethnologist* 3.3: 400-424.
- ____ (1977): "Folk Botanical Life Forms: Their Universality and Growth. *American Anthropologist* 79: 317-42.
- ____ (1979): "Folk Zoological Life Forms: Their Universality and Growth. *American Anthropologist* 81: 791-817.
- ____ (1984): *Language and Living Things: Uniformities in Folk Classification and Naming.* New Brunswick, N.J., Rutgers UP.
- Burton, Michael L. u. Lorraine Kirk (1979): "Ethn classification of Body Parts: A Three-Culture-Study". *Anthropological Linguistics* 21.8: 379-399.
- Desclés, Jean-Pierre (1998), "*Primitives, types, schèmes sémantico-cognitifs et polysémie verbale*", ed. F. Cordier, J. François, Jean-Emmanuel Tyvaert ; *Sémantique du verbe : Etat de la question*, 49 -62.
- ____ (1998a): Les représentations cognitives du langage sont-elles universelles? In *Essais sur la langage, logiques et sens commun.* Fribourg, Suisse, Editions Universitaires, p. 53-81.
- Dubois, D. (1993): *Lexique et catégories naturelles, Cahiers de Praxématique*, 21, Lexique et représentation des connaissances, Université Paul Valéry, Montpellier III, Praxiling.
- ____ (1995): *Catégories sémantiques "naturelles" et recherches cognitives: enjeux pluridisciplinaires. Linguistique et modèles cognitifs.* Acta Romanica Basiliensa 3, Université de Bâle, p. 76-104.
- Emmorey, Karen (2003 ed.): *Perspectives on classifier constructions in sign languages.* Mahwah, NJ: Erlbaum.
- Erbaugh, M. (1986): "Taking Stock: The Development of Chinese Noun Classifiers Historically and in Young Children. In: Craig (éd.): *Noun Classes and Categorization.* Amsterdam; Philadelphia: John Benjamins, pp. 399-406.
- Grinevald, Colette (1999): *Typologie des systèmes de classification nominale.* FDL 14:101-122.
- Hundius, Harald und Ulrike Kölver (1983): "Syntax and Semantics of Numeral Classifiers in Thai". *Studies in Language* 7.2: 165-214.
- Johansson, Ingvar (1989): *Ontological Investigations. An Inquiry into the Categories of Nature, Man, and Society.* Routledge, New York, London. 2nd ed.: Frankfurt: Ontos-Verlag 2004.
- Kempton, Willett (1978): "Category Grading and Taxonomic Relations: A Mug is a Sort of a Cup". *American Ethnologist* Vol. 6, No 1: 44-65.
- ____ (1981): *The Folk Classification of Ceramics: A Study of Cognitive Prototypes.* New York, Academic Press.
- Kleiber, G. (1990): *Sémantique du prototype,* Paris: PUF.
- Lakoff, Georg (1986): *Classifiers as a reflection of mind.* In Colette Craig (éd.): *Noun classification and categorization,* p. 13-51.
- Lakoff, Georg (1987): *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind.* Chicago, University of Chicago Press. Rosch 1973,
- Rosch, Eleanor (1981): *Prototype Classification and Logical Classification: The Two Systems.* In: E. Scholnick (éd.): *New Trends in Cognitive Representation.: Challenges to Poaget's Theory.* Hillsdale NJ: Lawrence Erlbaum Associates, S. 73-86.
- Sanches, Mary (1977): "Language Acquisition and Language Change: Japanese Numeral Classifiers". In: Blount und Sanches (Hg. 1977): 51-64.

- Wang, William S-Y. , Ke J.Y. & Minett, J.W. (2004): Computational Studies of Language Evolution. In: Huang, C.R. & Lenders, W. (eds.): Computational Linguistics and Beyond, Nankang: Academia Sinica: Institute of Linguistics, pp. 65-106.
- Wierzbicka, Anna (1972): Semantic Primitives. Reihe: Linguistische Forschungen 22. Übersetzt von Anna Wierzbicka und John Besemeres. Frankfurt, Athenäum Verlag.
- ____ (1985): Lexicography and Conceptual Analysis. Ann Arbor MI, Karoma Publishers.
- ____ (1989): Prototypes. *Linguistics* 27: 731-67.
- ____ (1992): Semantics, Culture, and Cognition: Universal Human Concepts in Culture-Specific Configurations. New York, Oxford University Press.
- ____ und Cliff Goddard (Hg. 1994): Semantic and Lexical Universals: Theory and Empirical Findings. Studies in Language Companion Series Bd. 25. Amsterdam/ Philadelphia, J. Benjamins.

Bibliographical References mentioned in Part II:

- Allan, Keith (1977): Classifiers. In: Language. Journal of the Linguistic Society of America 53 (1977) - pp. 285-311
- Bottéro, Francois (1996): Semantisme et classification dans l'écriture chinoise. Les systèmes des caractères par clés du *Shuowen jiezi* jusqu'au *Kangxi zidian*. Paris, Collège de France.
- Erbaugh, M. (1986): "Taking Stock: The Development of Chinese Noun Classifiers Historically and in Young Children. In: Craig (éd.): Noun Classes and Categorization. Amsterdam; Philadelphia: John Benjamins, pp. 399-406.
- Hirschfeld, L.A. and S. A. Gelman eds. (1994): Mapping the Mind. Domain Specificity in Cognition and Culture. New York: Cambridge University Press.
- Kaderas, Christoph (1998): Die *leishu* der imperialen Bibliothek des Kaisers Qianlong (reg. 1736): Untersuchungen zur chinesischen enzyklopädie. Asien- und Afrikastudien der Humboldt-Universität zu Berlin Bd. 4. Wiesbaden, Harrassowitz.
- Lakoff, Georg (1987): Women, Fire, and Dangerous Things. What Categories Reveal about the Mind. Chicago, University of Chicago Press.
- Peyraube, Alain (1988): Syntaxe diachronique du chinois, évolution des constructions datives du XIVe siècle av. J.-C. au XVIIIe siècle. Paris, Collège de France.
- ____ (1998): On the History of Classifiers in Archaic and Medieval Chinese. In: Benjamin K. Tsou (éd.): *Studia Linguistica Serica*. Hong Kong, City University of HK, pp. 39-68.
- Tai, James (1994): Chinese Classifier Systems and Human Categorization. In: M. Chen and O. Tzeng (éds.): 479-494.
- Tversky, Barbara et Kathleen Hemenway (1984): "Objects, Parts and Categories". *Journal of Experimental Psychology: General* 113: 169-193.
- Yang-Drocourt Zhitang (1993a): Evolution syntaxique des classificateurs chinois (du XIVe siècle av. J.-C. au XVIIIe siècle). Thèse pour l'obtention du doctorat de nouveau régime. Paris.