Preface

The basis for this volume is *Dictionaries. An International Encyclopedia of Lexicography* (HSK 5.1–5.3), the three volume publication edited by Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand and Ladislav Zgusta, published between 1989 and 1991, with 335 articles and a total of 3355 pages. For almost a quarter of a century this monumental work has internationally been the central reference source for dictionary research. Its influence has contributed significantly to the dynamic and innovative research activities that have characterized the discipline of lexicography and metalexicography.

It is quite natural that the three volumes will no longer fully reflect the current state of research in all its different subsections. During the last decades numerous new dictionaries and studies have been published. Even more significant have been changes in the methodological basis of dictionary work, especially the epochal innovations in the field of electronic and computer-assisted lexicography (cf. Article 1 and 2). An updating of HSK 5.1–5.3 has been perceived as an important desideratum for a considerable time. A complete revision has not been envisaged because many parts of the contents remain relevant even today. The publishing house and the editors have rather decided to prepare a supplementary volume, to be known as HSK 5.4. The supplementary volume links with the objectives formulated in the preface of the first volume (HSK 5.1, p. V–XV). The premises and subjects are complemented by new articles, following from the practice-internal and theoretical developments of the last 15 years. Special attention has been focused on electronic and computer-assisted lexicography. The following aims have been pursued in particular:

- Expansion and extension of the presentation of lexicography — especially the printed and electronic linguistic lexicography — of all language circles of the earth;
- Continuation of the presentation of the status and function of lexicographic reference works up to the present time within the cultural systems of societies, with special reference to the new digital forms of lexicography;
- Continuation of the already presented history of lexicography by means of selected examples from bigger cultures up to 2006;
- Expansion, deepening and modification of presented excerpts of a general theory of lexicography in its four sections, i.e. research into the use of dictionaries, research of dictionary criticism, historical and systematic dictionary research, with the aim that metalexicography can document its academic status as a scientific discipline with an independent formulation of a theory that has been established in the last decade;
- Complementation of the description of all phases of the lexicographic process, especially by means of the presentation of new methods based on electronic corpora, by means of the discussion of the new possibilities of computer-assistance as well as the consideration of the linking possibilities of different types of lexicographic processes;
- The presentation of lexicographic training and lexicographic institutions;
- The presentation of new metalexicographic methods, and as focus of the volume:
  - An extensive and detailed presentation of computational lexicography with regard to its historical development as well as its most important research fields, research perspectives, methods and tools of the last two decades.

On this basis the following thematic classification that includes 20 major chapters and in total 110 articles, has been envisaged for the supplementary volume.

Chapter I *Dictionaries in modern society: Current status and perspectives* (Articles 1–2) describes in its first part present and future prospects, problems and development possibilities of printed dictionaries in their different typological forms; the collaboration in lexicographic processes of printed dictionaries with corpus linguistic approaches is emphasized. In the second
part an analogous presentation is given of computational lexicography and its influence on linguistic lexicography and other fields of dictionary-related research.

In the last decade textal structures of printed dictionaries have often been in the focus of lexicographic research. The eight articles of Chapter II New developments in lexicographic theory I: Textual structures (Articles 3–10) attempt to present this research as exhaustively as possible in the following manner: A review article briefly indicating all the types of textal structures is followed by articles on the following selected types of structures: macrostructures, access structures, microstructures and mediostructures as well as the textal architectures. The next article (Article 9) treats addressing and addressing structures and in the tenth and final article of this chapter a presentation is given of the different types of dictionary articles.

In Chapter III New developments in lexicographic theory II: Dictionary types and functions new dictionary types and new developments in selected dictionary types are presented in nine articles (nr. 11–19). Consideration is given to the lexicography of language for general purposes as well as to languages for special purposes. Textual structures in electronic dictionaries, mixed dictionary genres, language contact dictionaries and recent developments in English, French as well as German learners’ dictionaries are treated. There is also a review article of special purpose dictionaries of lexicography and those in progress. Article 19 concludes the chapter with a discussion of dictionary functions.

Chapter IV New developments in lexicographic theory III: Selected dictionary subjects consists of six articles (nr. 20–25). From the numerous themes, directed especially at aspects of the dictionary subject matter, treated in recent metalexicographic research, relevant and innovative aspects have been selected in such a way that consideration is given to a relatively wide spectrum of problems. Culture-bound lexical items in dictionaries, sensitive items and the treatment of meaning in prototype theory are treated in separate articles. Furthermore the concept of semiotaxis is presented as well as new developments in the lexicographic treatment of collocations and lexicographic examples.

The last chapter dealing with new developments in the theory of lexicography is Chapter V New developments in lexicographic theory IV: Research in dictionary production and use, consisting of six articles (nr. 26–31). The chapter starts with a presentation of the structure of lexicographic processes and their phases, contains a presentation of dictionary management and review articles on empirical usage research and on the methods developed and employed in recent research on dictionary use. Furthermore there are articles on the concept of simultaneous feedback and research on usage in the domain of electronic dictionaries.

Chapter VI deals with organizational questions and is entitled New developments in lexicographic and metalexicographic organisation. The five articles (nr. 32–36) present a review article on the training of lexicographers and the training possibilities, the establishment of lexicographic units, the spectrum of existing lexicographic associations, principles of the evaluation of dictionaries and dictionary projects and, in the last chapter, the aids that play a role in metalexicographic research.

Chapters VII to XII are dedicated to the lexicography of individual languages. The objective of each of these descriptions is to document the development of lexicographic research since the start of the nineties of the 20th century. The point of departure of these supplementary articles is the treatment of the lexicography of individual languages presented in HSK 5.1–5.3. With this background the latest developments with regard to methodology and form, from both a synchronic and a diachronic perspective, are discussed and commented on. The following focal areas have been identified for the supplementary volume 5.4: Chapter VII New developments in the lexicography of individual languages since 1990 I: The ancient languages of the Near East and the classical languages (Articles 37–38) treats, by means of examples, the lexicography of the ancient eastern languages as well as Greek and Latin. Chapter VIII New developments in the lexicography of individual languages since 1990 II: The Romance languages (Articles 39–44) presents the lexicography of the Romance languages (with a focus on the Iberoromance languages, French, Italian and Romanian). Chapter IX New developments in the lexicography of individual languages since 1990 III: The Germanic languages (Articles 45–51) is dedicated to the complete field of Germanic languages (Scandinavian languages, English and American English, German, Dutch and Afrikaans). Chapter X New developments in the lexicography of individual languages since 1990 IV: The Slavic languages (Articles 52–54) treats the lexicography of selected Slavic languages, divided into the domains of Eastern Slavic languages, Southern Slavic languages
and Western Slavic languages. Chapter XI Lexicography of Arabic and selected Asian languages (Articles 55–58) focuses on the lexicography of Arabic, Chinese, Japanese and Korean. Chapter XII Lexicography of selected African languages contains six articles (nr. 59–64). These articles focus on the lexicography of certain language families (the Nguni and Sotho languages), individual languages (Shona and Fang) and the languages of certain geographical regions (Central Africa and Western Africa). The articles in this chapter complement the limited presentation of the lexicography of the African languages presented in HSK 5.3 and recognize the rapid development of lexicography on the African continent.

After the treatment of the lexicography of selected single languages follow ten chapters on electronic and computational lexicography. The section starts in chapter XIII. The history of computational lexicography with an overview of computational lexicographic work of the 1950s and 1960s (nr. 65) and, in a second article (nr. 66), an overview of the time before 1970/1980.

The subsequent two chapters deal with typologies of electronic dictionaries, one each for interactive ones and for dictionaries for Natural Language Processing (NLP). Chapter XIV Typology of electronic dictionaries I: Electronic dictionaries for human use consists of 8 articles (nr. 67–74). The first two cover design criteria for interactive electronic dictionaries, the “added value” that the user may get from the computational medium, as well as monolingual and bilingual electronic dictionaries for human use. The subsequent articles are devoted to electronic versions of printed dictionaries, internet dictionaries, dictionary portals, electronic learner’s dictionaries and context-sensitive dictionaries, i.e. those that guide the reader of an online text to the relevant (e.g. idiomatic) reading of a word. Finally, the overview is completed by articles on large-scale documentary electronic dictionaries, electronic encyclopedias and products that combine a dictionary look-up with corpus search.

Chapter XV Typology of electronic dictionaries II: Electronic dictionaries for machine use (Articles 75–82) is still typological in nature, but focuses on Natural Language Processing (NLP) applications. Individual articles describe lexical needs, representation techniques, descriptive approaches and major realizations for selected subdomains of NLP: speech recognition and speech synthesis, text processing and spelling or grammar checkers, information retrieval, information extraction and data mining, question answering and natural language understanding systems, text generation systems, machine translation and computer-assisted language learning. The eighth article discusses generic large-scale electronic dictionaries intended for multiple applications.

Chapter XVI focuses on lexical representation formalisms: Models for the representation of dictionaries: The form aspect (83–89). In seven articles, it deals with the major representation formalisms for electronic dictionary resources: XML (and SGML), databases, feature structure-based formalisms, formalisms of knowledge representation as used in Artificial Intelligence, and hypertext. One article discusses dedicated formalisms for lexical representation and one describes objectives and topics of international standardization in the field of representation formalisms for electronic dictionaries.

Chapter XVII Models for the representation of linguistic data in electronic dictionaries: The content aspect is subdivided into 6 articles (90–95). The first five articles concern selected “levels of linguistic description”, needs, types of data, coding and representation schemes. These articles may serve as an exemplary description of lexical description work for Natural Language Processing; they cover word formation morphology, syntax, and lexical semantics, as well as dictionaries which explicitly make an attempt at describing and representing the interaction between syntax and semantics. The last article of this chapter describes typologies of metadata for electronic dictionaries.

So far, all chapters of this part of the handbook concern the electronic dictionary as a product: its types, its contents, and its form. Chapters XVIII and XIX are devoted, however, to dictionary making and to its computational aspects. As most dictionary making is nowadays supported to some degree by computational tools, obviously, these chapters concern both, electronic and paper dictionaries, for humans and machines. One chapter is devoted to corpus design for lexicography, the other to tools for data acquisition and to support for dictionary writing.

Chapter XVIII Computer-based dictionary making I: Acquisition of lexical data from corpora – corpus design (Articles 96–99) groups four methodological articles, two on corpus design criteria for lexicography (monolingual and bilingual), one on existing large corpora for lexicography (such as many of the national corpus projects), and a last one on possibilities and limitations of
the use of the World Wide Web for lexicography. Chapter XIX is on tools for lexicographers: Computer-based dictionary making II: Acquisition of lexical data from corpora and machine readable dictionaries – tools and procedures (Articles 100–107). Six of its eight articles cover the major aspects of corpus-based data provision for dictionary making: corpus pre-processing, the corpus-based design of a macrostructure, concordancing, the extraction from corpora of data on syntactic and collocational properties of words, and tools for the use of parallel and comparable corpora. The two other articles concern specialized tools and procedures for lexicography: the reversal of bilingual dictionaries A → B to make them usable as a starting point for a dictionary of the language pair B → A and finally lexico-graphic workbenches that include several, among others, corpus-based tools.

Computational work in terminology is the topic of chapter XX Computational terminology (Articles 108–110). As terminology and terminography are extensively dealt with in the HSK handbook on specialized language (HSK 14), only truly computational aspects of terminography are dealt with, here: the representation of terminological data in terminological databases and in formats for the exchange of terminological data, and approaches to and challenges in the extraction of terminological data from texts.

The volume is concluded by comprehensive indexes.

The first three volumes of the encyclopedia contained articles in English, German and French. However, all articles in the supplementary volume are exclusively written in English. It was clear to the editors that this decision would not be welcomed unanimously. The reasons for this decision are obvious. English is constantly expanding as international language and as the lingua franca of the sciences. This is an enormous advantage for the communication needs of the modern world with its manifold interconnections. Similar developments are also elsewhere known in history. During the early modern era Latin dominated as language of the learned world. Throughout Europe, French played a prominent role during the 18th and 19th centuries. Compared to both these languages English has decisive advantages that also result from profound changes in social parameters. It is a vibrant language that develops steadily, its use is not limited to a single social class and in comparison it is easy to learn. The dominance of English is the price of successful communication on international level and nobody will seriously deny its advantages.

An endeavor like the present one poses challenges with regard to the contents but also on an organizational level. Besides many enriching scientific and personal experiences there have also been many adversities over the years that had to be overcome collectively. Given that the preparation of this volume took several years, some of the articles may reflect an earlier or later state of the art (2008–2012); in the rapidly evolving domain of computational lexicography and electronic dictionaries, there may have been individual realizations since then which are not covered in detail; mostly these will however just prolong the methodological lines described in the respective articles. That the project has successfully been seen to its end can to a large extent be contributed to the fact that the editors, even in difficult phases, never questioned their commitment to the project or their respectful and amicable contact with each other. In addition the collaboration with the publishing house, especially with Barbara Karlson, was always and in each phase constructive and reliable. Without the support of many collaborators and colleagues the realization of this volume would not have been possible. We express our gratitude to Maria Hegner (Saarbrücken), Jean Hurinville (Paris, Art. 58), Laurel Loch (Stafford, USA), Stephanie Schweickard (Saarbrücken) and David Tollerton (Bangor, UK) for their assistance in the preparation and scrutiny of the English versions of articles, furthermore to Lissy Säuberlich (Heidelberg) for entering the data of Articles 3–10 and to Christina Doreen Görtz (Sömmerda) for the preparation of the illustrations of these articles. For their valued assistance in the coordination of editorial tasks we thank Maria Hegner, Lisa Sumski and Simone Traber (Saarbrücken), Julia Weidenkaff and Michael Krebes (Stuttgart), Antonia Kruse (Hildesheim) and Liana Roos (Stellenbosch). Last but not least a special word of thanks to all the authors of this volume who, almost without exception, participated in this project also during its difficult phases in a constructive way and with cooperative understanding.

Autumn 2013

Rufus Gouws, Stellenbosch
Ulrich Heid, Hildesheim
Wolfgang Schweickard, Saarbrücken
Herbert Ernst Wiegand, Heidelberg