

CONTENTS

1. INTRODUCTION	1
2. DESCRIPTIVE TOOLS OF METALEXICOGRAPHY	4
2.1 The general metalexicographical situation	4
2.2 Lexicography and computers	6
2.3 Metalexicographical descriptive apparatuses with and without computers in mind	7
2.3.1 Apparatuses for the comparative description of a collection of machine-readable dictionaries: the BONNLEX lexicon system and the MDW draft for lexical microstructure	8
2.3.2 Descriptive apparatuses for monolingual dictionaries	12
2.3.2.1 Schaefer's proposal for a "Gestaltung der lexikalischen Mikrostruktur"	12
2.3.2.2 The Longman Dictionary of Contemporary English (LDOCE)	15
2.3.3 Descriptive apparatuses for translation dictionaries. The Canadian Dictionary: An Experiment in Automated Lexicography	18
2.3.4 Descriptive apparatuses for onomasiological dictionaries	20
2.3.4.1 The DANTERM record	20
2.3.4.2 Roget's Thesaurus, and thesauri or conceptual dictionaries of the vocabularies of individual authors	23
2.4 The need for a tertium comparationis	25
3. A TAXONOMY FOR LEXICOGRAPHICAL DATA	29
3.1 Taxonomy for lexicographical data	32
3.1.1 Project-specific information types	33
Language	
Approach	
Medium of expression	

3.1.2	General information types	35
	Etymological information	36
	Origin	
	Parallels	
	Graphic information	36
	Orthographical information	
	Graphic symbol	
	Grammatical information	37
	Part of speech	
	Information on inflection	
	Word formation	
	Syntax	
	Phonetic information	38
	Prosodic features	
	Segmental features	
	Pragmatical information	38
	Contextual information	
	Usage	
	External reference	
	Evaluative information	
	Administrative information	
	Semantic information	41
	Subject classification	
	Semantic relations	
	Content specification	
	Equivalence	
3.2	The linking of information	50
3.3	Overview of the main groups for general information types, categories and subcategories	51
4.	USING THE TAXONOMY	52
4.1	Academic dictionaries	54
4.1.1	Use of the taxonomy to classify lexicographical information for establishing a format for an academic dictionary	54
4.1.2	General points about Gammeldansk Ordbog	54
4.1.3	Establishing editorial principles for the GLDO	56

- 4.2 Translation dictionaries: describing Vinterberg and Bodelsen's Dansk-Engelsk Ordbog with the taxonomy 72
 - 4.2.1 Describing an existing dictionary 72
 - 4.2.2 Description using the project-specific information types of the taxonomy 75
 - 4.2.3 Lemma, sublemma and lexical units 75
 - 4.2.4 Incorporating the general information types of the taxonomy 77
 - 4.2.5 A brief account of the other information types in Dansk-Engelsk Ordbog 88
 - 4.2.6 What emerged from using the taxonomy 89
- 4.3 Term banks: examples of the description of DANTERM, the Danish Term Bank, with the help of the taxonomy 99
 - 4.3.1 DANTERM in general 99
 - 4.3.2 DANTERM compared with the project-specific information types of the taxonomy 103
 - 4.3.3 DANTERM compared with the general information categories of the taxonomy 105
 - 4.3.4 Summing-up 115
- 5. DICTIONARY FORMATS 118
 - 5.1 The description of the structure of dictionaries 118
 - 5.2 Formal languages 119
 - 5.2.1 Backus-Naur notation 121
 - 5.2.2 The ETTO Notation 122
 - 5.2.3 Bracket diagrams 123
 - 5.3 Format description and practical lexicography 125
 - 5.4 An example: Dansk-Engelsk Ordbog 126
 - 5.4.1 The ETTO Notation 127
 - 5.4.2 The bracket diagram 130
 - 5.4.3 Extended BNF 131
 - 5.5 Relationship with the taxonomy 133
 - 5.5.1 The entry 134
 - 5.5.2 Lexical units 134
 - 5.5.3 The format categories and the taxonomy 135

5.5.4	The taxonomy categories and the format	138
5.6	Mark-up systems	139
5.7	The Oxford English Dictionary - an example	144
5.8	Format description and computerized registration of the structure of a printed bilingual dictionary	146
5.8.1	Dansk-Fransk Ordbog described with the bracket diagram	147
5.8.2	Conclusions on description with the bracket diagram	151
5.8.3	Computerized analysis and the establishment of a database	152
5.8.4	The representation of the entry structure in the database	158
6.	EXPERIMENTS WITH FORMATTED ENTRY OF DICTIONARY DATA	161
6.1	System requirements and experimental material	161
6.1.1	The word-processing component	161
6.1.2	Data management	163
6.1.3	Introduction to experiments with formatted input of dictionary data	164
6.2	Gammeldansk Ordbog	166
6.3	The Supplement to Ordbog over det Danske Sprog	169
6.3.1	Short description of the ODS and the ODSS	169
6.3.2	A format for the ODSS	171
6.4	Compulexis and the GLDO/ODSS	175
6.4.1	Compulexis	176
6.4.2	Field code names for information categories	188
6.4.3	Use of the general facilities of the Compulexis system in the GLDO and ODSS	190
6.4.4	Use of the specialized facilities of the Compulexis system in the GLDO and ODSS	190
6.4.4.1	Validity checks	190
6.4.4.2	Automatic generation of field content and codes (data entry screen and data entry form printout)	193
6.4.4.3	Automatic generation of alterations in field content in on-screen and printout column forms on the basis of field content in the on-screen data entry form	196

6.4.4.4	Field content visible in data entry form but suppressed in column form	210
6.4.4.5	Automatic alphabetization of entries	217
6.4.4.6	Control of typography	219
6.4.4.7	Embedded fields	221
6.4.4.8	Additions to the Compulexis standard character set	226
6.5	Conclusions on the data entry experiments	227
6.5.1	The GLDO	227
6.5.2	The ODSS	228
6.5.3	The Compulexis system	230
7.	COMPLEX DICTIONARY DATA IN DATABASE SYSTEMS	231
7.1	Introduction	231
7.2	The structuring of data	231
7.2.1	The bracket diagram	231
7.2.2	Entity/relationship diagrams	234
7.3	Lexicographical data management requirements	239
7.3.1	Data structure	239
7.3.2	Data entry and changes in the data entered	241
7.3.3	Character set and fonts	242
7.3.4	Data transfer	242
7.3.5	Searches	243
7.3.6	Sorting	248
7.3.7	Presentation	249
7.3.8	Data protection	250
7.3.9	Statistical functions	251
7.3.10	Conclusion	251
	REFERENCES	253
	RESUME FRANÇAIS	262
	DEUTSCHE ZUSAMMENFASSUNG	273

