
Contents

<i>List of figures</i>	vii
<i>List of tables</i>	viii
<i>List of boxes</i>	xi
<i>Preface</i>	xiii
<i>Acknowledgements</i>	xv
<i>List of abbreviations</i>	xix
<i>List of sources</i>	xxi
1 Parallel text corpora: a general overview	1
1.1 Different types of text corpora	2
1.2 Problems in using text corpora	8
1.3 Text corpora and the translator	13
1.4 The usefulness of corpora to researchers	15
2 Designing and compiling a parallel corpus	19
2.1 A few notes on software and hardware	20
2.2 Planning the corpus	22
2.3 Inputting the texts	26
2.4 Aligning the texts	29
2.5 Annotating the texts	34
2.6 Storing the corpus	37
2.7 Copyright issues	39
3 Using parallel corpora: basic search procedures	45
3.1 Conducting searches	45
3.2 Concordancing	48
3.3 Frequency lists	51
3.4 Collocations	54

4	Processing search results	60
	4.1 Comparing translation equivalents from a parallel concordance	60
	4.2 Processing frequency lists	65
	4.3 Studying collocations	71
	4.4 Using parallel corpora in monolingual research projects	83
5	Using parallel corpora: more advanced search procedures	99
	5.1 Software tools for quantitative analysis of corpora	100
	5.2 Checking the reliability of research data	102
	5.3 Measuring central tendencies in the data	105
	5.4 Quantitative analysis of concordances	121
	5.5 Collocations revisited	123
	5.6 Keywords	133
6	Applications of parallel corpora	147
	6.1 Parallel corpora as online dictionaries	147
	6.2 Parallel corpora in lexicography	149
	6.3 Terminology studies and multilingual corpora	155
	6.4 Morphology and syntax	163
	6.5 Pragmatics	173
	6.6 Translation studies	184
7	A survey of available parallel corpora	197
	7.1 Multilingual parallel corpora	198
	7.2 Bilingual parallel corpora	206
	7.3 Interpretation corpora	209
	<i>Final remarks</i>	214
	<i>Glossary</i>	216
	<i>Appendix 1: Corpus-based M.A. theses at the University of Tampere</i>	222
	<i>Appendix 2: Sample programs</i>	224
	<i>Index</i>	229