Contrastive Corpus Linguistics

Patterns in Lexicogrammar and Discourse

Edited by
Anna Cermakova, Hilde Hasselgård, Markéta Malá
and Denisa Šebestová

BLOOMSBURY ACADEMIC
LONDON • NEW YORK • OXFORD • NEW DELHI • SYDNEY

Contents

List of figures ix
List of tables xi
List of contributors xiv

Introduction *Hilde Hasselgård, Anna Cermakova, Markéta Malá and Denisa Šebestová* 1
1 The Present Status and Recent Trends in Corpus-Based Contrastive Linguistics *Karin Aijmer* 11

Part I Lexicogrammar in Contrast

2 Seeing Through Languages and Registers: A Closer Look at the Cognates SEE and SE *Signe Oksefjell Ebeling* 29
3 Periphrastic Genitive Constructions in English and Norwegian *Hilde Hasselgård* 62
4 Double Object Constructions in English and Norwegian: Verbs of SENDING, BRINGING, LENDING and SELLING *Thomas Egan* 87
5 Prepositional Patterns in English and Czech Newspaper Discourse *Denisa Šebestová and Markéta Malá* 108
6 A Cross-Linguistic Study of Journalistic Phraseology *Jiajin Xu, Guying Zhou, Xinlu Liu, Yuanyuan Wei, Ruchen Yu and Suhua Zhang* 132
7 Corpus-Based Contrast in Audiovisual Customization: A Pilot Study on *Can/Could* and Subject Pronouns in Spanish Dubbing *Camino Gutiérrez-Lanza and Rosa Rabadán* 157

Part II Discourse in Contrast

8 The Social Functions and Linguistic Patterns of *Please* and Its Norwegian Correspondences *Stine Hulleberg Johansen and Kristin Rygg* 183
9 Discourse Connectives in English and French: A Contrastive Study on Political Discourse *Diana Lewis* 204

10 Reporting Verbs in English, Czech and Finnish *Anna Cermakova and Lenka Fárová* 236
11 From Dashes to Dashes? – A Contrastive Corpus Study of Dashes in English, German and Swedish *Jenny Ström Herold and Magnus Levin* 260

Index 285

Figures

3.1 Dispersion of the Norwegian *til*-genitive and the English *of*-genitive 69
3.2 Possessive relations in periphrastic genitives 70
3.3 Animacy of possessor (N2) 72
3.4 Congruent and noncongruent translation correspondences 75
3.5 Distribution of congruent translations (*til* = *of*) across corpus texts 76
3.6 Percentages of animate and inanimate possessors in congruent and noncongruent translations between Norwegian and English 78
4.1 Percentages for the ditransitive and prepositional dative with the eight verbs in the original texts in the ENPC 92
4.2 Percentages of syntactically congruent and divergent translations of the eight verbs 92
4.3 Extent of semantic overlap between the verbs of sending, selling, lending and bringing in double object constructions 101
6.1 Density plots of variability and predictability across five languages 141
6.2 Mosaic plot of functional characteristics across five languages 144
7.1 *Can* (STs) > *poder* (TT1s and TT2s) 166
7.2 *Poder* < *can*: TT1s, TT2s and STs (Es) 169
7.3 *Could* (STs) > *poder* (TT1s and TT2s) 170
7.4 *Poder* < *could*: TT1s, TT2s and STs (Es) 172
7.5 Subject pronouns in Spanish: TT1s, TT2s and STs (Es) 173
9.1 Token frequency of connectives that occurred at least once in the comparable corpus 215
9.2 Main functions of *alors*, *then* and *puis* in the interviews data 219
9.3 Distribution of *then*, *alors* and *puis* in the data (excluding *alors que*) 227
9.4 Comparison of distribution of *alors* and *then* in a conversational comparable corpus and in the political interviews comparable corpus 228
10.1 *With*-PP: Congruency of verb translations (Czech and Finnish) 249
10.2 *Without*-PP: Congruency of verb translations (Czech and Finnish) 250
11.1 Structure of the Linnaeus University English-German-Swedish corpus (LEGS) 263
11.2 Dashes in English, German and Swedish originals in LEGS per 10,000 words 267

11.3 Frequencies of dashes and brackets in originals per 10,000 words 268
11.4 Functional categories of dashes in the LEGS material 268
11.5 Proportions of primary functions of dash-introduced text 270
11.6 Primary functions of dash-introduced versus bracket-introduced text 270
11.7 Sub-categories of content-oriented and interpersonal functions of dash-introduced text per 10,000 words 271
11.8 Positions of clauses and phrases introduced by dashes 273
11.9 Frequencies per 10,000 words of dashes in originals and translations 275
11.10 Proportions of retained dashes in translations 276
11.11 Real numbers for added dashes 276
11.12 Replacements from dashes: proportions of 'non-dash' correspondences in translations 279
11.13 Replacements into dashes: proportions of 'non-dash' correspondences in originals 279

Tables

2.1 Token counts in the ENPC original texts: dialogue vs. narrative 37
2.2 Distribution of the verb forms in raw numbers after disambiguation 39
2.3 Occurrences of *see* and *se* in the sub-corpora 41
2.4 Complementation patterns of *see* and *se* in the different sub-corpora 43
2.5 Semantic classification of *see* and *se* (according to SFG process types) 47
2.6 Overview of Premier League and *Eliteserie* clubs represented in the ENMaRC 60
2.7 Overview of authors and texts in the ENPC used in this study 61
3.1 Expressions of possessive relations in English and Norwegian 63
3.2 Size and structure of the English-Norwegian Parallel Corpus 66
3.3 Possessive meaning relations 68
3.4 Frequencies of periphrastic genitives in English and Norwegian fiction and nonfiction 69
3.5 Most frequent meaning relations with animate and inanimate possessors 73
3.6 Meaning relations and congruence 77
3.7 Noncongruent translation correspondences of the periphrastic genitive 78
3.8 Prepositions corresponding to *of/til* in noncongruent translations of the periphrastic genitive 80
4.1 Number of double object examples in the ENPC original language texts 91
4.2 Verbs used in congruent and divergent translations of *send* and *sende* 93
4.3 Verbs used in congruent and divergent translations of *sell* and *selge* 96
4.4 Verbs used in congruent and divergent translations of *lend* and *låne* 97
4.5 Verbs used in congruent and divergent translations of *bring* and *bringe* 99
4.6 Mutual correspondence of pairs of verbs in double object constructions 101
5.1 Corpora and tools used in the present study 110
5.2 Frequencies of the prepositional n-grams analysed 112
5.3 Position of the prepositions IN and V in the n-grams 113

5.4 Structural classification of prepositional n-grams 114
5.5 Complex prepositions containing *v* sorted by frequency 119
5.6 Complex prepositions with *in* ordered by frequency 122
6.1 Top 5 p-frames in the five languages 138
6.2 Examples of discourse-pragmatic functions of p-frames in the five languages 139
6.3 Summary of the 100 top 3–5-word p-frames of the five languages 140
6.4 Median variability and median absolute deviation of the 100 top p-frames across the five languages 141
6.5 Median predictability and median absolute deviation of the 100 top p-frames across the five languages 142
6.6 Mean predictability of the 100 top p-frames across five languages 144
7.1 *Poder* < *Can*: Verification of dubbese (TRACEci-CORPES XXI) 168
7.2 *Poder* < *Could*: Verification of dubbese (TRACEci-CORPES XXI) 171
7.3 Subject pronouns: Verification of dubbese (TRACEci-CORPES XXI) 174
7.4 *Yo*: Verification of dubbese by function (TRACEci-CORPES XXI) 174
7.5 *Usted*: Verification of dubbese by function (TRACEci-CORPES XXI) 175
7.6 *Él-ella(s)*: Verification of dubbese by function (TRACEci-CORPES XXI) 175
8.1 Summary of types of situations and interaction types in the ENPC 191
8.2 Patterns of *please* in English original and translated texts in the ENPC 192
8.3 Correspondences of *please* in ENPC 193
8.4 Typical translations from *please* to Ø in the ENPC 198
9.1 Comparable corpus 206
9.2 Normalized frequencies of coherence-marking expressions occurring more than five times per 100,000 words (*que* = *que/qu'*) 216
9.3 Usage of *then*, *alors* and *puis* in the interview data 226
10.1 English fiction corpus 240
10.2 Translation corpus 240
10.3 Distribution of *said* in the English fiction corpus 242
10.4 Distribution of the lexicogrammatical patterns modifying reporting *said* 243
10.5 Frequency distribution of *with*- and *without*-PP 247
10.6 Translation solutions of *with-PP* in verb congruent translations (Czech) 251
10.7 Translation solutions of *without-PP* in verb congruent translations (Czech) 251

10.8 Translation solutions of *with-PP* in verb congruent translations (Finnish) 252
10.9 Translation solutions of *without-PP* in verb congruent translations (Finnish) 253
10.10 Translation solutions of *with-PP* in verb-divergent translations (Czech) 255
11.1 Word counts for the LEGS sub-corpora 263