# Contents

**1 Introduction** **1**

**2 German Verb Classes** **9**

2.1 Idea and Definition of Verb Classes . . . . . 9

2.1.1 Idea of Verb Classes . . . . . 10

2.1.2 Verb Classification . . . . . 10

2.1.3 Related Work on Verb Classes . . . . . 13

2.2 Class Properties . . . . . 25

2.3 Usage of Verb Classes . . . . . 104

2.4 Summary . . . . . 108

**3 Statistical Grammar Model** **109**

3.1 Context-Free Grammars and their Statistical Extensions . . . . . 110

3.1.1 Context-Free Grammars . . . . . 110

3.1.2 Probabilistic Context-Free Grammars . . . . . 112

3.1.3 Head-Lexicalised Probabilistic Context-Free Grammars . . . . . 118

3.1.4 Summary . . . . . 122

3.2 Grammar Development and Implementation . . . . . 122

3.2.1 Grammar Development for Lexical Verb Information . . . . . 123

3.2.2 The German Context-Free Grammar . . . . . 124

3.3 Grammar Training . . . . . 154

3.3.1 The Statistical Parser . . . 154
3.3.2 Training Strategy . . . 155
3.4 Grammar-Based Empirical Lexical Acquisition . . . 157
3.4.1 Subcategorisation Frames . . . 157
3.4.2 Selectional Preferences . . . 158
3.4.3 Related Work on H-L PCFGs . . . 158
3.5 Grammar Evaluation . . . 164
3.5.1 Subcategorisation Lexica for Verbs . . . 164
3.5.2 Evaluation of Subcategorisation Frames . . . 170
3.5.3 Lexicon Investigation . . . 172
3.5.4 Related Work . . . 176
3.6 Summary . . . 178

**4 Clustering Algorithms and Evaluations 179**

4.1 Clustering Theory . . . 179
4.1.1 Introduction . . . 180
4.1.2 Data Objects, Clustering Purpose and Object Features . . . 181
4.1.3 Data Similarity Measures . . . 182
4.1.4 Clustering Algorithms . . . 184
4.2 Clustering Evaluation . . . 190
4.2.1 Demands on Clustering Evaluation . . . 190
4.2.2 Description of Evaluation Measures . . . 193
4.2.3 Comparison of Evaluation Measures . . . 201
4.3 Summary . . . 205

**5 Clustering Experiments 207**

5.1 Clustering Data . . . 207

5.1.1 German Verbs and Verb Classes . . . 208

5.1.2 Feature Choice . . . 209

5.1.3 Data Illustration . . . 226

5.1.4 Summary . . . 233

5.2 Verb Class Experiments . . . 233

5.2.1 Clustering Methodology . . . 233

5.2.2 Baseline and Upper Bound . . . 234

5.2.3 Experiment Results . . . 234

5.2.4 Summary . . . 247

5.3 Experiment Interpretation . . . 248

5.3.1 Interpretation of Experiment Outcome . . . 248

5.3.2 Feature Manipulation and Class Coherence . . . 261

5.3.3 Summary . . . 263

5.4 Optimisation Criteria . . . 264

5.4.1 Feature Variation . . . 264

5.4.2 Feature Selection . . . 269

5.4.3 Optimising the Number of Clusters . . . 271

5.4.4 Verb Sense Disambiguation . . . 275

5.4.5 Summary . . . 276

5.5 Large-Scale Clustering Experiment . . . 276

5.6 Related Work . . . 282

5.6.1 Automatic Induction of Class-Relevant Features . . . 282

5.6.2 Automatic Induction of Classes . . . 286

**6 Conclusion** **291**

6.1 Contributions of this Thesis . . . . . 292

6.1.1 A Small-Scale German Verb Classification . . . . . 292

6.1.2 A Statistical Grammar Model for German . . . . . 293

6.1.3 A Clustering Methodology for NLP Semantic Verb Classes . . . . . 294

6.2 Directions for Future Research . . . . . 297

**A Subcategorisation Frame Types** **299**

**B Corpus-Based Analysis of Subcategorisation Frames** **301**

**C Large-Scale Set of German Verbs** **305**

**Zusammenfassung** **313**

**Summary** **323**

# List of Tables


2.1 Class-based estimated frequencies of direct object nouns . . . . . 107

3.1 Example CFG . . . . . 111
3.2 Example PCFG (1) . . . . . 113
3.3 Example PCFG (2) . . . . . 114
3.4 Example H-L PCFG (rules) . . . . . 119
3.5 Example H-L PCFG (lexicalised parameters) . . . . . 120
3.6 Terminal grammar categories . . . . . 125
3.7 Terminal features . . . . . 126
3.8 Examples of grammar terminals . . . . . 127
3.9 Subcategorisation frame types: VPA . . . . . 131
3.10 Subcategorisation frame types: VPP . . . . . 132
3.11 Subcategorisation frame types: VPI . . . . . 133
3.12 Subcategorisation frame types: VPK . . . . . 133
3.13 Generalised frame description . . . . . 134
3.14 Clause type examples . . . . . 135
3.15 Auxiliary combination with non-finite verb forms . . . . . 139
3.16 Subcategorisation frame distribution for *glauben* . . . . . 160
3.17 Refined np distribution for *reden* . . . . . 161
3.18 Nominal arguments for *verfolgen* in na . . . . . 162
3.19 Nominal arguments for *reden ber*$_{Akk}$ ‘to talk about’ . . . . . 163

3.20 Lexical subcategorisation for *befreien* . . . 165
3.21 Lexical subcategorisation for *zehren* . . . 165
3.22 Examples for purely syntactic lexical subcategorisation entries . . . 166
3.23 Examples for PP-refined lexical subcategorisation entries . . . 166
3.24 Frequencies of *Duden* verbs in training corpus . . . 171
3.25 Evaluation of subcategorisation frames . . . 171
3.26 Investigation of subcategorisation frames . . . 176

4.1 Example evaluation for class-based P/R . . . 196
4.2 Example evaluation for pair-wise P/R . . . 197
4.3 Example evaluation for adjusted pair-wise precision . . . 197
4.4 Example evaluation for mutual information . . . 198
4.5 Example evaluation for Rand index . . . 199
4.6 Example evaluation for adjusted Rand index . . . 200
4.7 Example evaluation for matching index . . . 201
4.8 Comparison of evaluation measures . . . 204

5.1 Empirical properties of gold standard verb classes . . . 208
5.2 Frame distributions for *glauben* . . . 212
5.3 Frame+PP distributions for *reden* and frame type np . . . 213
5.4 Nominal arguments for *verfolgen* in na . . . 214
5.5 Nominal arguments for *reden ber*$_{Akk}$ ‘to talk about’ . . . 215
5.6 Selectional preference definition for *essen* in na as based on GermaNet nodes . . . 218
5.7 Selectional preference definition with GermaNet top nodes (1) . . . 221
5.8 Selectional preference definition with GermaNet top nodes (2) . . . 222
5.9 Frame+Pref distributions of *verfolgen* and frame type na . . . 223
5.10 Combined Frame+Pref distributions of *essen* and frame type na . . . 224
5.11 Examples of most probable frame types (1) . . . 227

5.12 Examples of most probable frame types (2) . . . 228
5.13 Examples of closest verbs . . . 230
5.14 Examples of nearest neighbour verb pairs . . . 231
5.15 Examples distances between verbs in same or different classes . . . 232
5.16 k-Means experiment baseline and upper bound . . . 235
5.17 Comparing distributions (frame only, reduced verb set) . . . 241
5.18 Comparing distributions (frame+pp, reduced verb set) . . . 241
5.19 Comparing distributions (frame only, full verb set) . . . 242
5.20 Comparing distributions (frame+pp, full verb set) . . . 242
5.21 Comparing similarity measures (frame only, reduced verb set) . . . 243
5.22 Comparing similarity measures (frame+pp, reduced verb set) . . . 243
5.23 Comparing similarity measures (frame only, full verb set) . . . 244
5.24 Comparing similarity measures (frame+pp, full verb set) . . . 244
5.25 Comparing clustering initialisations (frame only, reduced verb set) . . . 245
5.26 Comparing clustering initialisations (frame+pp, reduced verb set) . . . 245
5.27 Comparing clustering initialisations (frame only, full verb set) . . . 246
5.28 Comparing clustering initialisations (frame+pp, full verb set) . . . 246
5.29 Comparing feature descriptions on reduced verb set . . . 247
5.30 Comparing feature descriptions on full verb set . . . 247
5.31 Comparing the amount of PP information (reduced verb set) . . . 265
5.32 Comparing the amount of PP information (full verb set) . . . 265
5.33 Comparing selectional preference slot definitions on full verb set . . . 267
5.34 Comparing selectional preference frame definitions on full verb set . . . 268
5.35 Comparing optimal feature sets . . . 270
5.36 Large-scale clustering on frames . . . 278
5.37 Large-scale clustering on frames and PPs . . . 278
5.38 Large-scale clustering on frames, PPs and preferences . . . 278

A.1 Subcategorisation frame types . . . . . 300

B.1 Corpus-based analysis of subcategorisation frames (1) . . . . . 302
B.2 Corpus-based analysis of subcategorisation frames (2) . . . . . 303
B.3 Corpus-based analysis of subcategorisation frames (3) . . . . . 304

C.1 Large-scale set of German verbs (1) . . . . . 305
C.2 Large-scale set of German verbs (2) . . . . . 306
C.3 Large-scale set of German verbs (3) . . . . . 307
C.4 Large-scale set of German verbs (4) . . . . . 308
C.5 Large-scale set of German verbs (5) . . . . . 309
C.6 Large-scale set of German verbs (6) . . . . . 310
C.7 Large-scale set of German verbs (7) . . . . . 311
C.8 Large-scale set of German verbs (8) . . . . . 312

# List of Figures


3.1 Syntactic analyses for *John loves Mary* and *John loves ice-cream* . . . . . 111

3.2 Syntactic analyses for *John ate that cake* . . . . . 113

3.3 Syntactic analyses for *John eats the cake with a spoon* . . . . . 116

3.4 Syntactic analyses for *John eats the cake with icing* . . . . . 117

3.5 Syntactic analysis for *John blames Mary for her anger* . . . . . 121

3.6 Syntactic analysis for *John loves Mary for her smile* . . . . . 121

3.7 Top-level clause construction . . . . . 129

3.8 Nominal syntactic grammar categories . . . . . 144

3.9 Proper names . . . . . 145

3.10 Noun phrases generating pronouns and cardinals . . . . . 146

3.11 Noun phrases introducing relative and interrogative clauses . . . . . 146

3.12 Prepositional phrase arguments . . . . . 148

3.13 Prepositional phrase arguments in relative and interrogative clauses . . . . . 149

3.14 Prepositional phrase adjuncts . . . . . 150

3.15 Attributive adjectival phrases . . . . . 151

3.16 Predicative adjectival phrases . . . . . 152

3.17 Adverbial phrases . . . . . 153

3.18 *Duden* lexical entry for *zehren* . . . . . 169

4.1 Algorithm for agglomerative hierarchical clustering . . . . . 186

4.2 Algorithm for k-Means clustering . . . . . 188

4.3 $APP$ evaluation on introducing errors . . . . . 203
4.4 $PairF$ evaluation on introducing errors . . . . . 203
4.5 $Rand_{adj}$ evaluation on introducing errors . . . . . 203

5.1 GermaNet hierarchy for noun *Kaffee* ‘coffee’ . . . . . 216
5.2 Propagating frequencies through GermaNet hierarchy . . . . . 217
5.3 Varying the number of clusters on reduced verb set (evaluation: $APP$) . . . . . 273
5.4 Varying the number of clusters on reduced verb set (evaluation: $PairF$) . . . . . 273
5.5 Varying the number of clusters on reduced verb set (evaluation: $Rand_{adj}$) . . . . 273
5.6 Varying the number of clusters on full verb set (evaluation: $APP$) . . . . . 274
5.7 Varying the number of clusters on full verb set (evaluation: $PairF$) . . . . . 274
5.8 Varying the number of clusters on full verb set (evaluation: $Rand_{adj}$) . . . . . 274