Table of contents

Acknowledgements **7**

Terms and definitions **15**

List of figures **17**

List of tables **20**

Introduction **23**

1 Context, objectives and structure of the investigation **23**

1.1 Background 23

1.1.1 Constant development of the lexicon across borders 23

1.1.2 The desire to control lexical development 24

1.1.3 Deliberate lexical interventions: Successes and challenges 25

1.2 Motivation and objectives 26

1.2.1 Lexical knowledge: Exploring lexical environments 28

1.2.2 Lexical opinion: Gathering naturally occurring data 31

1.3 Context 33

1.4 Overall methodology and research questions 35

1.5 Thesis structure 38

PART 1
THEORETICAL BACKGROUND 41

2 Deliberate lexical interventions and their effectiveness 43
2.1 Introduction 43
2.2 Why: Goals of an extralinguistic nature 46
2.3 What: Goals of a linguistic nature in the speech community 47
2.3.1 Filling a lexical gap … 48
2.3.2 … or modifying the existing lexicon 50
2.4 How: The route from aim to achievement for language managers 53
2.4.1 From in vitro to in vivo lexical items 53
2.4.2 The actualization issue 56
2.4.3 Speakers as the source of lexical change 58
2.5 Outcomes and effectiveness of deliberate lexical interventions 68
2.5.1 Possible and desired outcomes of deliberate lexical interventions 68
2.5.2 The effectiveness of deliberate lexical interventions 70

3 Intervening on speakers' lexicon under conditions of uncertainty 73
3.1 Introduction 73
3.2 Problem statement 73
3.2.1 Language as a complex adaptive system 73
3.2.2 Introduction of a new lexical item and marketing 77
3.2.3 Deliberate lexical intervention as a cocreation process 80
3.3 Evaluating deliberate lexical intervention effectiveness in the speech community 83
3.3.1 Previous studies on lexical knowledge 84
3.3.2 Previous studies on lexical opinion 96
3.3.3 Previous studies on lexical replication 104

3.4 Proposals for language managers 105
3.4.1 Exploring speakers' lexical environment 106
3.4.2 Exploring speakers' lexical opinions in context 111

PART 2
ESPERANTO FOR SCIENTIFIC RESEARCH 119

4 Using the Esperanto speech community to study lexical phenomena 121
4.1 Introduction 121
4.2 The emergence of a decentralized speech community 121
4.2.1 From a plan … 121
4.2.2 … to a fully functioning language … 124
4.2.3 … spoken mainly by a decentralized community 125
4.2.4 Defining the Esperanto speaker 126
4.2.5 Counting Esperanto speakers 130
4.2.6 The status of Esperanto in society and science 132
4.3 Lexical change in Esperanto 134
4.3.1 Democratically completing the plan 134
4.3.2 Lexical norms in Esperanto 146

5 Electronic networks discussing Esperanto lexical items 149
5.1 Introduction 149
5.2 Five online networks under the microscope 149
5.2.1 Astronomia Terminaro 154
5.2.2 Esperanto-tradukistoj 155
5.2.3 Lingva Konsultejo 155
5.2.4 Retposhta rondo por redaktantoj 156
5.2.5 Vivo-Vikio 158

5.3 Linguistic knowledge and language proficiency of network contributors 159
5.3.1 Objective of the survey 159
5.3.2 Methods 159
5.3.3 Results 165
5.3.4 Discussion 172

Part 3
Empirical investigation and proposal 177

6 Speakers' lexical environments: insights from focus-groups 179
6.1 Introduction 179
6.2 Methods 181
6.2.1 Focus group rationale 181
6.2.2 Discussion guideline 182
6.2.3 Choice and recruitment of participants 183
6.2.4 From samples to focus groups 188
6.2.5 Analytical framework 193
6.3 Results 195
6.3.1 Is lexical information needed? 195
6.3.2 External search for information: Diversity on the rise 197
6.3.3 Collaborative dictionaries: Not trustworthy but useful 203
6.3.4 When speakers turn their filters on 206
6.4 Summary 208

7 Methods: A proof of concept for detecting opinionated autonyms 210
7.1 Introduction 210
7.2 Theoretical foundations for the detection of opinions and autonyms based on natural language processing 211
7.2.1 Opinionated autonyms in corpora 211

7.2.2 Building and manually annotating a test set 216
7.2.3 Adapting the Metalinguistic Operation Processor..... 217
7.3 Corpus compilation 218
7.3.1 Guiding principles 218
7.3.2 Corpus collection 220
7.3.3 Corpus cleaning 222
7.3.4 Morphological tagging and pre-editing 223
7.3.5 Choosing units of analysis 224
7.4 Grammatical features of opinionated autonyms in Esperanto..... 225
7.4.1 Opinion indicators in written language 225
7.4.2 Autonymy indicators in written language..... 232
7.5 Building and evaluating indicators for opinionated autonyms in Esperanto 240
7.5.1 Developing sentiment and metalinguistic lexicons for Esperanto 240
7.5.2 Indicators for detecting opinions in Esperanto corpora.... 244
7.5.3 Indicators for detecting autonyms in Esperanto corpora.. 247
7.6 Extracting opinionated autonym candidates from the corpus..... 250
7.7 Units of analysis extracted from the corpus..... 252
7.8 Summary 254

8 Results: Lexical criteria in context.....256
8.1 Introduction..... 256
8.2 Analysing opinionated autonyms..... 256
8.3 Speakers' lexical criteria..... 260
8.3.1 Properties of the lexical item 262
8.3.2 Reference to language use..... 275
8.3.3 Reference to other language items..... 280
8.3.4 Reference to language instances or resources 283

8.3.5 Extralinguistic statements 287
8.4 Summary 288

9 Using naturally occurring data to absorb uncertainty in deliberate lexical interventions 290
9.1 Introduction 290
9.2 Assessing and understanding speakers' lexical knowledge 291
9.2.1 Data on lexical-source recall 291
9.2.2 Data on lexical-source opinion 293
9.3 Assessing speakers' lexical opinions 295
9.3.1 Data on lexical item polarization at the lexical item level 295
9.3.2 Data on lexical item polarization at the designational paradigm level 297
9.4 Understanding speakers' lexical opinions 299
9.5 A model for action based on naturally occurring data 301

CONCLUSION 305

10 Conclusion and future research 307
10.1 Objectives and methods 307
10.2 Summary of main findings and achievements 307
10.3 Future research 309

References 311

Endnotes 359

Appendices 367