



Uncertainty in deliberate lexical interventions

Exploring Esperanto speakers' opinions
through corpora

Mélanie Maradan

Mélanie Maradan

Uncertainty in deliberate lexical interventions

Mélanie Maradan

Uncertainty in deliberate lexical interventions

Exploring Esperanto speakers' opinions through corpora

Umschlagabbildung unter Verwendung der bearbeiteten Fotografie: Esperanto, Marcela McGreal
(CC-BY-Lizenz, <https://creativecommons.org/licenses/by/4.0/>)

peer reviewed content

Publiziert mit Unterstützung des Schweizerischen Nationalfonds
zur Förderung der wissenschaftlichen Forschung.



CC-BY-NC-ND

ISBN 978-3-7329-9237-9

ISSN 1862-6149

DOI 10.26530/20.500.12657/43200

© Frank & Timme GmbH Verlag für wissenschaftliche Literatur
Berlin 2021.

Herstellung durch Frank & Timme GmbH,
Wittelsbacherstraße 27a, 10707 Berlin.

Printed in Germany.

Gedruckt auf säurefreiem, alterungsbeständigem Papier.

www.frank-timme.de

Zugleich Cotutelle-Dissertation Université de Genève und
Universität Hildesheim 2019, Hil2

Welche von den Vorschlägen sich einbürgern,
darüber entscheidet letzten Endes stets
die Gesamtheit der Sprachbenutzer.

Eugen Wüster

Acknowledgements

First and foremost, I would like to thank the late Detlev Blanke, who sparked my curiosity and encouraged me to pursue a PhD related to Esperanto. Unfortunately, he left us too soon to see the result of my work, but I will never forget how much he helped me find relevant research material and how he systematically challenged me to think critically.

Both of my thesis advisors guided me through my years of research. I would have never started my PhD adventure without Klaus Schubert, who also guided me through the process of writing this work. Aurélie Picton supported me and provided constructive criticism, and I would not have survived the preparation of this PhD project without her.

I would also like to express special appreciation to those who are generally forgotten: from the invisible workers cleaning my office before I even arrived at work to the helpful secretaries sparing me much administrative hassle. I believe all work is noble and that my research would not have been possible, especially without the help of librarians who always obtained a copy of the papers and books I wanted, even in libraries on other continents and in languages they did not necessarily speak.

Dozens of people taught and helped me every step of the way, from professors who, despite their heavy workload, made time to read my PhD project in detail and offered me their honest opinion and advice, to the colleagues and other PhD candidates who put up with my ups and downs, mindset, sarcasm, and complaints, and who provided emotional support in situations I could not have handled on my own. Also, thank you to all the entertaining colleagues with whom I lunched and laughed.

Many thanks to Carlos Rodríguez Penagos, who kindly provided all programmatic components of the Metalinguistic operation processor (MOP) he developed (2004b).

Another thank you goes to the late Toon Witkam for making his ESPSOE Esperanto software packages (morphological tagging) freely available on the Web.

I would like to thank the Esperanto speakers who agreed to take part in my research, as well as all individuals who provided valuable information for my research. Special thanks to Bart Demeyere for sharing data of the ViVo interface.

I would also like to thank the Swiss National Science Foundation for the grant that allowed me to travel back and forth between the University of Hildesheim and the University of Geneva, as well as to spend two semesters at the former university. The joint degree could not have been possible without this financial support.

Thank you to my English reviewers for grammar editing and proofreading for spelling and punctuation errors, and to the German translator and my German thesis advisor for writing my German abstract. All errors remain mine.

Thank you to professors Sabine Fiedler, François Grin, Alexander Künzli, Nathalie Mälzer, and Jean Quirion for accepting to serve as examination board members.

Last but not least, I would like to thank my family, friends, acquaintances, and work partners for being patient and understanding when I made this research project the utmost priority.

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

Terms and definitions

Terms for which a definition is explicitly provided are presented in **bold type** in the running text. This list contains the terms thus defined and refers to the page on which they can be found.

- a posteriori planned language 123
- a priori planned language 123
- accuracy 215
- active lexical item recognition 86
- active lexical source recognition 86
- alternative lexical creation 51
- alternative resources 184
- autonym 232, 234
- autonyms 116
- classification 213
- coding frame 257
- coding scheme 257
- common language 112
- content analysis 256
- contribution 224
- corpus 218
- deliberate lexical intervention 25, 45
- deliberate multilingual secondary lexical creation 49
- denaskulo 126
- denominative variation 50
- designational paradigm 50
- DLI 43
- effectiveness 71
- Effectiveness of a deliberate lexical intervention 72
- electronic network of practice 34, 149
- environmental factor 55
- Esperantist 128, 130
- Esperanto speaker 130
- Esperanto studies 134
- Esperantology 133
- ethnic language 122
- external search for information 198
- F1-measure 215
- F-measure 215
- focus group 181
- foisonnement 50
- folk lexicography 143
- folk linguist 172
- folk speakers 27
- formal variation 50
- hard-to-reach population 130
- implantation coefficient 25, 104
- in vitro lexical item 54
- lexical implantation 26
- lexical item 53
- lexical item recall 86
- lexical knowledge 27, 86
- lexical opinion 27, 31, 96
- lexical recall 29, 85
- lexical recognition 29, 85
- lexical replication 27
- lexical selection 51
- lexical source 63, 85

lexical source awareness 29
lexical source recall 86
lexical standardization 69
lexicon 45
loan creation 136
metalanguage 112
metalinguistic awareness 112
metalinguistic lexical item 241
metalinguistic sign 232
metalinguistic statements 33
metalinguistic statements with an
 opinionated autonym 212
minority language 132
monitor corpus 219
natural language 54
naturally occurring data 32
neologism 136
neologismo 136
object language 112
opinion 225
opinionated MSa 212
ordinary sign 232
original lexical creation 49
passive lexical item recognition 86
passive lexical source recognition 86
planned language 121
planned language project 124
planned language *proper* 124
planned semi-language 124
precision 215
recall 215
representativeness 220
semantic false friend 265
sign in use 232
speaker 54
speech community 23
strategy 77
structural uncertainty 74
subjective statement 259
synchronous focus group 182
target lexical items 25
terminometrics 25
total implantation 68

List of figures

Figure 1.	Overview of the thesis methodology.....	37
Figure 2.	Example of the actualization process toward an outcome lexical situation desired by language managers. After a certain amount of time, a five-item designational paradigm (DP) has been reduced to a single-item designational paradigm by lexical selection.....	53
Figure 3.	Simulation of language change using the language change model in NetLogo (Wilensky, 2016), based on Nettle’s model using social impact theory (1999). The white variant propagates through the speech community (from 15 to 100 speakers in about 180 iterations).	59
Figure 4.	Three-step implantation process of a lexical item (LI) from a speaker’s perspective, based on the ideas expressed in (Martin, 1994, p. 33).....	65
Figure 5.	Possible outcomes of a deliberate lexical intervention aimed at implanting the target lexical item “mercatique” instead of “marketing” at the level of a single speaker (modification of the existing lexicon: lexical selection).....	66
Figure 6.	Three-step implantation process of a lexical item from a language manager’s perspective, adapted from the matrix in (Crawford & Bryce, 2003, p. 369).	67
Figure 7.	Typology of lexical change adapted from Munske (2015, p. 26).....	137
Figure 8.	Languages in which the 41 respondents from electronic networks of practice considerer the can have a smooth conversation (Esperanto is not included here).....	168
Figure 9.	Number of respondents from electronic networks of practice who indicated to have read works on Esperanto grammar or terminology partly or entirely.....	170
Figure 10.	Respondents’ relative position on the two dimensions of linguistic knowledge and language proficiency, according to my tentative metric for representing the two dimensions of linguistic knowledge and language proficiency.....	173

Figure 11. Domains for which focus group participants (there were 31 participants, see Section 6.2.42) reported having a diploma or work experience. Data were gathered through the prescreening survey.	189
Figure 12. Distribution, represented as a Venn diagram, of focus group participants according to the following criteria: professional language training and/or experience, known contributors to electronic networks of practice, current country of residence (ISO Alpha-2 country codes), and year of birth. The names are pseudonyms.....	191
Figure 13. AJAX Chat (Esperanto localization) in a Web browser with sample conversation involving one moderator (Melanie, light font) and two participants (Ludovikito and Klara, dark font).....	192
Figure 14. Classification of contexts with opinions and autonyms using a two-step binary classification for identifying opinionated autonym candidates (OAC).....	214
Figure 15. Representation of a language’s lexicon, containing three sets of lexical items: lexical items of the object-language, neutral lexical items, and metalinguistic lexical items. Metalanguage is formed by metalinguistic lexical items and by lexical items from any one of the three lexical sets standing in autonymical condition. Adapted from Rey-Debove (1978).....	236
Figure 16. Venn representation of contexts containing both opinion and autonymy (intersection, Set C.).....	251
Figure 17. Contexts as classified in absolute numbers (table) and percentages (graph). The contexts selected for the analysis of this chapter are those containing both autonymy and opinion (opinionated autonym candidate contexts or OAC, white on the graph).....	253
Figure 18. Process for directed content analysis (inspired from Hsieh & Shannon, 2005; Mayring, 2015, pp. 62–63; Zhang & Wildemuth, 2009).....	258
Figure 19. Using naturally occurring data (NOD) to assess lexical-source recall and better disseminate the lexical source containing target lexical item if the assessment is not satisfactory.....	293

Figure 20. Using naturally occurring data (NOD) to understand lexical-source opinion and improve the perception of lexical sources containing target lexical item in the case of negative perception.	295
Figure 21. Sample opinionated autonymical lexical items grouped in three clusters (positive, negative, and mixed opinions).	296
Figure 22. Using naturally occurring data (NOD) to assess lexical item polarization at the level of the lexical item and, if polarization is negative, move toward understanding speakers' lexical opinion through lexical criteria (9.4).	297
Figure 23. Comparison of statements made about four lexical item candidates competing as equivalents of "luminosity" (the amount of energy emitted by a star). "Lumeco" appears to be the best candidate in this set of results.....	298
Figure 24. Summary of statements concerning the autonym "pedologio," which is four negatively polarized opinionated statements and a neutral statement.	299
Figure 25. Using naturally occurring data (NOD) to decide whether a given target lexical item needs revision.	301
Figure 26. Model for action for language managers based on naturally occurring data (NOD). A "yes" means that no further activities should be undertaken.	303

List of tables

Table 1.	Two major types of deliberate lexical interventions (1. Filling a lexical gap, 2. Modifying the existing lexicon) according to the initial language situation.	52
Table 2.	Summary of methods used to investigate six aspects of lexical knowledge.	94
Table 3.	Translated and adapted from Remysen (2009).	100
Table 4.	Summary of methods used to investigate lexical opinions.	102
Table 5.	General typology of Esperantism adapted from Gobbo (1997, pp. 53–55: the sign “{}” stands here for combinations that do not exist).	129
Table 6.	Type of collaboration in the five electronic networks of practice. Classification based on Abel and Meyer (2013; roles of contributors as human users); Doan, Ramakrishnan, and Halevy (2011); and Klosa and Tiberius (2016; tasks of users in the lexical process).	151
Table 7.	Type of internet dictionary for electronic networks of practice linked to a dictionary, based on (Engelberg & Storrer, 2016).	153
Table 8.	The eight indicators used in the survey for assessing the two dimensions <i>linguistic knowledge</i> and <i>language proficiency</i>	160
Table 9.	Profile of ReVo contributors selected to pilot test the survey.	164
Table 10.	Types of notification chosen by the members of the Yahoo group tradukado. The figures are for October 2015 and were received by e-mail from the administrator.	165
Table 11.	Respondents’ self-assessment about their language and linguistic knowledge.	171
Table 12.	Domains for which respondents reported having formal training (school or university) or informal training (private seminars, self-study).	171
Table 13.	Summary of items respectively considered for the samples, excluded from the samples and included into the final samples.	187

Table 14.	Number of individuals who completed the prescreening survey vs. number of individuals in the four subsamples.	188
Table 15.	Sample contexts from the corpus illustrating types of metalinguistic statements in relation to opinion and autonomy and delimiting the notion of metalinguistic statement with an opinionated autonomy in Subset C.....	213
Table 16.	Contingency table for evaluating the performance of binary classification tasks.	215
Table 17.	Results of manual classification for the test set.	216
Table 18.	Figures for the five subcorpora <i>Astronomia Terminaro</i> , <i>Esperanto-tradukistoj</i> , <i>Lingva konsultejo</i> , <i>Retpoŝta rondo ...</i> and <i>ViVo-Vikio</i>	221
Table 19.	Examples of opinion-bearing lexical items for various parts of speech (POS) respectively formed from two opinion-bearing morphemes: the root <i>taŭg-</i> and the suffix <i>-ind</i>	229
Table 20.	Examples of lexical items (both opinion-bearing and nonopinion-bearing) from the corpus based on the polysemous root “signif.”	231
Table 21.	Ordinary signs and metalinguistic signs.	233
Table 22.	Degrees of explicitness of a metalinguistic statement with an opinionated autonomy according to the presence or absence indicators (opinion, metalanguage, and autonomy) in a metalinguistic statement with an opinionated autonomy. The figures result from a manual classification of the training set.	238
Table 23.	Results (precision and recall) for opinion indicators on the test set, in isolation (by type of indicator) and in combination (union of indicators).	246
Table 24.	Results (precision and recall) for autonomy indicators on the test set, in isolation (by type of indicator) and in combination (union of indicators).	250
Table 25.	Expected number of opinionated autonomy candidates and expected number of relevant contexts for each indicator combinations above assuming the test set is representative of the entire corpus.	252

Table 26. Opinionated autonym candidates (OAC) extracted from the corpus.	253
Table 27. Examples of lexical items that received positive comments for their conciseness.	269
Table 28. Examples of lexical items that were judged to be too long by speakers.	270
Table 29. Examples of lexical items whose grammar was deemed to be inappropriate by speakers.	271
Table 30. Examples of lexical items whose grammar was deemed to be appropriate by speakers.	271
Table 31. Examples of written sources in which speakers find lexical items.	278
Table 32. Examples of language resources in which speakers find lexical items.	286
Table 33. Examples of statements about lexical items linked to different types of linguistic variation.	288
Table 34. Summary of statements about the opinionated lexical item “pedologio.”	300

Introduction

1 Context, objectives and structure of the investigation

1.1 Background

1.1.1 Constant development of the lexicon across borders

Every day, new thoughts are expressed, and new objects, technologies, and processes are developed. Speakers constantly require new lexical items: *Smartphone*, *green technologies*, *e-signature*, and so on. How are these new realities being named? The lexicon of a language is not a fixed set carved in stone. It is not static (Tournier, 2007, p. 32) but finds itself in constant motion. For language to play its social role and meet communication needs, it must allow for the creation of new lexical items (Montané March, 2012, pp. 12–17). As language is a social interactional phenomenon (Croft, 2000, p. 89), it is intrinsically linked to the actual reality of the **speech community**¹ by which it is spoken, and it constantly adapts to speakers' needs (Domanskiy, 2016, p. 53), reflecting changes in society (Munske, 2015, p. 20).

Currently, the use of new lexical items is expanding. Michel and colleagues (2011), for instance, have found that the size of the English lexicon has been increasing at a rate of about 8,500 new items per year during the last 50 years. Petersen and his colleagues (2012) reported that there is “a dramatic increase in the relative use ... of newborn words over the last 20–30 years, likely corresponding to new technical terms, which are necessary for the communication of core modern technology and ideas” (p. 3).

Objects, ideas, and technologies are exchanged across countries and cultures, and this impacts the lexicon: New lexical items may be borrowed or coined in a receiver language under the influence of a primary language (see, for example,

1 Here, I use ‘speech community’ in a broad sense, i.e. as “any human aggregate characterized by regular and frequent interaction by means of a shared body of verbal signs and set off from similar aggregates by significant differences in language usage.” (Gumperz, 2009, p. 66) Depending on the context, the speech community can thus encompass anything from all the speakers of a language down to a very small group of people (see the discussion in Rampton, 2010).

Sablayrolles, Jacquet-Pfau, & Humbley, 2011). This phenomenon of imported neology is nothing new: Religion, military conquests, and colonization have repeatedly been key factors of language contact during various historic periods (Siemund, Gogolin, Schulz, & Davydova, 2013, p. 15). Today, the heterogeneity of languages has become the norm, considering that there currently coexist about 4,500 languages in some 200 countries across the globe (Edwards, 2012, p. 26). Multilingualism has become a fact of life (Antia, 2000, p. xxi), and a majority of the world's population has at least some multilingual competence (Bhatia & Ritchie, 2013). Even the minority of monolinguals is exposed to other languages because multilingual realities appear in people's neighborhoods, and one not need physically move. Language is presently all around: on store windows, commercials, traffic signs, and so on (Gorter, 2006, p. 1)

It is not the phenomenon itself that is so impressive but rather the speed at which it is currently occurring. At other periods in history, contacts between languages took place between a limited amount of speakers and most often over a prolonged period of time, which facilitated the "digestion" of new lexical items in the receiving language (Muñoz Martín & Valdivieso Blanco, 2006, p. 471). The rhythm and intensity of exchanges between languages today are not comparable. The increasing speed at which discoveries are currently being made and their swift propagation to other groups of people and parts of the world influence the frequency at which languages must update their lexicon (Sablayrolles, 2000, p. 381). Needs for specialized lexical items in companies and institutions have been growing (Slodzian, 2000, p. 29). Scientists and engineers name new concepts on a daily basis, and this means that journalists, professors, translators, and so on need to find equivalent lexical items in their respective languages (Estopà, Coromina, & Mestres, 2010, p. 16). Generally, the importance of imported neology is expected to increase in the future (Humbley, 2006, p. 199).

1.1.2 The desire to control lexical development

Neological creation can be spontaneous (i.e., the result of common language creativity mechanisms); it can, however, also be the result of a deliberate intervention (Cabré, 2002, p. 32; Díaz Hormingo, 2012, p. 108).

Human beings have always had the tendency to assess the processes taking place around them and have tried to manipulate them, and language change is no exception to this general principle (Landrø, 2008, p. 88). Regarding language, some human beings—here I call them **language managers**—have tried to

influence the (new) lexical items speakers use. In the present thesis, I call this type of intervention from their part a **deliberate lexical intervention** and define it as an intervention in the lexicon of a group of speakers made with the final objective of bringing about language use (full lexical implantation) for specific lexical items. Deliberate lexical interventions are triggered by a variety of extralinguistic motives, including but not limited to language standardization (Nahir, 2002, p. 273), purification (Landrø, 2008, pp. 30–33), and modernization (Ricento, 2000, p. 200).

1.1.3 Deliberate lexical interventions: Successes and challenges

Although it can be said that interventions in language date back to antiquity (Nekvapil, 2011, p. 872) and that the discipline of terminology, concerned with the conscious planning of new lexical items, began developing almost a century ago,² language managers designing deliberate lexical interventions have yet to address a range of fundamental and complex issues to achieve satisfying lexical implantation results. Deliberate lexical interventions are a goal-driven activity: They have an objective, which is pursued through a set of actions (practices) expected to lead to given results. The linguistic objective of a deliberate lexical intervention is to bring about usage from the speech community for specific lexical items chosen by language managers, which I here call **target lexical items**.³ The instruments and the activities used for making deliberate lexical interventions have differed greatly from case to case and have not always been based on scientific research (Bhreathnach, 2011). Language managers have often tried to implant target lexical items based on hypotheses concerning linguistic, sociolinguistic, and procedural factors (Quirion, 2012, p. 136). Oftentimes, the results of deliberate lexical interventions were far from being fortunate (see, for example, Mortureux, 1987, p. 250). As a matter of fact, the results of deliberate lexical interventions have not even been evaluated.

Over the last 2 decades, a new field of research has emerged that empirically shows the extent to which speakers are using the target lexical items of language managers. In this recent framework, called **terminometrics** (Quirion, 2003a, 2005, 2010), each lexical item can be semiautomatically associated with an **implantation coefficient**, which is a metric that describes how well a target lexical

2 The founding work being Wüster (1931).

3 Thoiron et al. (1993) speak of 'terme-cible'.

item performs in language use in comparison to its lexical competitors.⁴ The outcome of deliberate lexical intervention activities can thus be relatively well measured, although indirectly.⁵ Studies undertaken using this terminometrical framework have shown in which cases and to which extent target lexical items of language managers are used by the **target speech community** (Quirion, 2010; Vila i Moreno & Vila i Moreno, 2007).

Terminometrics can provide a reliable way to check whether deliberate lexical intervention activities generally lead to satisfactory outcomes. It is mainly designed to *assess the final results* of deliberate lexical interventions. However, terminometrics does not *explain the process* of **lexical implantation**—how lexical items enter language use—nor does it allow us to identify which actions in deliberate lexical intervention activities lead to unsuccessful end results. It does not allow for the assessment of other important aspects taking place during the lexical implantation process either.

Conducting deliberate lexical interventions is a complex endeavor comprising a large set of activities (see, for example, the components in the model developed by Bhreathnach, 2011). If, at the end of the process, language managers obtain low implantation coefficients for their target lexical items (which has repeatedly been the case), they should change something in their practices, but it does not tell them exactly what corrective actions they should take. Changing practices to improve the results of deliberate lexical interventions thus remains a guessing game to a certain extent.

1.2 Motivation and objectives

In terms of lexical implantation, reasoning forces us to be concerned with the transition from speech to language (Gaudin, 2007, p. 32). Ultimately, as it is always the speakers' decision as to whether to use the new lexical items that the authority (language managers) would like to implant (Loubier, 1994, p. 20; Wüster, 1931, p. 124), the speaker can be considered the interface between speech and language. The speaker is the barrier the target lexical items of language managers must overcome in their move from creation and selection to usage.

4 Performance is evaluated in terms of frequency of usage.

5 I am saying 'indirectly' because the implantation coefficient refers to a state of language use and does not differentiate between the impact of deliberate lexical intervention activities and that of environmental factors (see my detailed explanation in Section 3.3.3).

A large amount of previous researchers have focused on theories of language planning and language professionals and norm authorities: terminology policies, language planning agencies, professional writers, linguists, terminologists, translators, and so on. However, what is the true role of the language community, of “ordinary” people—I call them **folk speakers** here⁶—in lexical implantation? In this thesis, I argue for better monitoring of speakers’ roles in lexical implantation. Much like marketing managers who profit from understanding their customers’ journey from awareness to acquisition, I believe language managers should gain a thorough comprehension of the conversion funnel speakers go through, from becoming aware of the existence of a lexical item (here **lexical knowledge**) to forming an opinion toward them (here **lexical opinion**) to its actual use (here **lexical replication**). Lexical replication has already found a standard monitoring protocol (see Section 3.3.3) language managers could use, but lexical knowledge and lexical opinion less so. Thus, in the present work, I question how information from folk speakers could help language managers throughout the implantation process.

Engaging folk speakers in applied linguistic research is not a novel idea. Linguistics has long started to include nonscientific information in its approaches and, in particular, to use the construct of a native speaker to better understand language. With Harris, the native speaker only sat on the fence; with Chomsky, he was already primed for a career as an arbitrator (Antos, 1996, p. 256), and along with the development of folk linguistics starting in the 1960s, the common individual started to become an integral part of applied linguistics (Wilton & Stegu, 2011). In a similar way, and following Quirion’s largely yet unimplemented idea of involving speakers directly in lexical management activities (2012),⁷ I believe that engaging with speakers directly or indirectly as informants, as well as understanding how they come to know, evaluate, and choose the lexical items they use, will help language managers improve the results of deliberate lexical interventions.

How does the lexical implantation process take place among speakers (i.e., how do speakers go from ignoring a lexical item, to becoming aware of its existence, to actually using a lexical item?). One thing is certain: Speakers do not simply adopt the lexical items (variants) that are most common around them; otherwise, innovations would never spread within a speech community. This is

6 Building on concepts from folk linguistics (see e.g. Niedzielski & Preston, 1999).

7 There have been first timid steps in this direction, for instance wikiLF in the French language (FranceTerme, n.d.).

what Nettle calls the threshold problem (see Nettle, 1999, pp. 98–99). This implies that speakers apply some kind of filter and selection mechanism toward the new lexical items they are exposed to. These selection mechanisms have been studied by researchers: language managers have tried to develop indicators and methods for assessing and understanding how their target lexical items fare at each step of the lexical implantation process so as to increase the efficiency of their deliberate lexical intervention activities. Through this investigation, I propose to consider the lexical implantation process as a three-step process:⁸ (a) lexical knowledge, (b) lexical opinion, and (c) lexical replication.

A fairly reliable metric (the implantation coefficient) has been developed for assessing the final stage of the lexical implantation process (see Section 3.3.3), that of lexical replication (c), but the stages of lexical knowledge (a) and lexical opinion (b) are still poorly understood. The study conducted by Gresa Barbero (2016) is a good example of lexical knowledge: Members of the target speech community were unaware of the target lexical items that language managers wanted to implant.⁹ Thus, a paramount need must be satisfied: systematically exploring speakers' **lexical environments** on a case-by-case basis to understand how target lexical items could reach the target speech community. As far as lexical opinion is concerned, several researchers have approached the issue (see Section 3.3.2), but results are case dependent. This should not come as a surprise, as some lexical criteria stand in direct contradiction. For instance, there is always a tension between clarity and brevity in the (conscious or unconscious) choice of a lexical item (see, for example, Limaye & Pompian, 1991). Thus, it seems necessary to explore, also on a case-by-case basis, what lexical criteria are important to the members of the target speech community. In Section 3.2.1, I will further elaborate on why I believe lexical change cannot be planned a priori and should therefore be monitored. I provide an overview of previous works showing the issues I have identified for the two steps of lexical knowledge and lexical opinion, as well as the objectives I have set to propose solutions toward addressing the issues.

1.2.1 Lexical knowledge: Exploring lexical environments

I start with the need to examine speakers' lexical environments on a case-by-case basis to assess whether the target lexical item is reaching members of the target

8 This is explained in detail in Section 2.4.

9 This will be detailed in Sections 3.3.1 and 3.3.2.

speech community. Up until now, researchers have tried to measure six lexical knowledge dimensions:¹⁰ **lexical item recall**, **active lexical item recognition**, **passive lexical item recognition**, **lexical-source recall**, **active lexical source recognition**, and **passive lexical source recognition**. A target lexical item will not be used by speakers if it is unknown to them. Thus, language managers have been interested, for instance, in assessing the ability of speakers to recognize a target lexical item. To this end, they have used surveys or interviews, in which they directly asked speakers whether they knew the target lexical item (Allony-Fainberg, 1974). They further attempted to measure how well speakers connect a target lexical item with its corresponding concept. Semistructured interviews or participant observation was used: In these settings, researchers tried to make speakers spontaneously utter the target lexical item (Gaudin & Guespin, 1993; Vila i Moreno & Vila i Moreno, 2007) or gave speakers a definition or a picture of a concept and noted whether speakers expressed the target lexical item in response to this stimulus (Gouadec, Crespel, & Colombel, 1993; Nogué Pich & Vila i Moreno, 2007a; Thoiron, Iwaz, & Zaouche, 1993). Researchers have also attempted to evaluate whether speakers have come into contact with the sources (official documents, dictionaries, online terminology databases, etc.) containing the target lexical items. To evaluate this dimension, the researchers generally used semistructured interviews (Gresa Barbero, 2016; Vila i Moreno & Vila i Moreno, 2007).

Existing studies on lexical knowledge have shown that results are often not acceptable for language managers (e.g., speakers are unaware that the lexical source containing target lexical items even exists).¹¹ According to Allony-Fainberg (1974), for instance, the knowledge of a specific target lexical item can be as low as 24.7% for a specific group of speakers (pp. 501–502), meaning that only about one speaker out of four claims to know the lexical item considered. In Vila i Moreno and Vila i Moreno (2007), in 40% of cases, speakers do not know the lexical item (p. 79), only 35.7% know that the source containing the item exists (p. 76). In such cases, this strongly suggests that a large majority of speakers have probably *not* been exposed at all to the target lexical items that language managers would like to implant in language usage *or* to the lexical sources containing them. If a speaker must come into contact with the target lexical items that language managers would like to implant, the chance that this speaker will use this lexical item is approximately zero. For speakers to come into contact with a new

10 This is detailed in Section 3.3.1.

11 So are the metrics that were used, but this is out of scope in this introductory chapter. See for instance Thoiron and his colleagues for a discussion of methodologies (1993).

lexical item, they must be exposed to a source disclosing this linguistic innovation. Language managers have tried to disseminate their new lexical items, for instance, through institutional sources (dictionaries, newsletters, and official documents),¹² but if language does evolve and lexical awareness for target lexical items remains low, this must mean that speakers receive their new lexical material from sources other than those of language managers. If assessments are not satisfactory, what matters is investigating why this is the case.

Therefore, one objective of my thesis is to learn more about the lexical sources speakers use or, in other words, to explore what I here call speakers' lexical environment.

Objective 1: Explore speakers' lexical environments

In this thesis, I specifically question *where* speakers look for new lexical items.

In the past, speakers may have passively received new lexical items mainly from teachers and institutional dictionaries—the correct way to speak was defined by upper layers of society¹³—but in the current era of the participative Web and with the increasing digitization of society, speakers have access to a large range of potential lexical sources, and they seem to actively draw lexical material from them. According to Bonnin (2014), speakers unprecedentedly use new sources such as online forums or dictionaries, wikis, blogs, and machine translation systems as their reference, normative materials (p. 358). Investigating exactly which lexical sources speakers use could have fallen into the scope of lexicographical research, but to date, lexicography has mostly focused on specific dictionaries, usually institutional or otherwise prestigious ones (see Nesi, 2012a),¹⁴ and has largely

12 A few studies (e.g. Ballarin, 2009) questioned how language managers disseminate new lexical items, but literature on the topic remains scarce (see e.g. Martin, 1992, p. 34; Nic Pháidín, Ó Cleircín, & Bhreathnach, 2010, p. 955).

13 At the time when France was a monarchic society, for instance, the lexical norm was that of the words used by the French royal court, and only the king and a handful of renowned writers were in a position to coin new lexical items (Guilbert, 1975, pp. 50–51).

14 There are a few exceptions, e.g. Meyer & Gurevych (2012).

ignored the extralexicographical situations in which speakers find themselves.¹⁵ According to Tarp (2009), “no known user research has produced real information on the objective user needs, i.e. the needs that may occur in the extralexicographical situation preceding the dictionary consultation” (pp. 292–293). Only very recently (e.g., Kunkel, 2015) have scholars started to question *where* speakers find language-related information, especially on the Internet, and *what they do* with the pieces of information they collect.

1.2.2 Lexical opinion: Gathering naturally occurring data

As Thoiron et al. (1993) stated, lexical knowledge is a *precondition* for lexical implantation, but it by no means guarantees successful lexical implantation. Some lexical items are known by speakers but are not used (Thoiron et al., 1993, p. 69). Seen from a sociotheoretical perspective, the impact of a lexical item might depend on bias in favor or against the lexical item itself (functional selection), the status of the lexical item (social selection), the distance of the lexical item, and the number of sources using the lexical item (Nettle, 1999).¹⁶ I am writing *might* because this is an unverified theoretical assumption. The criteria that trigger or impede the acceptance of new lexical items among speakers remain largely unverified (Quirion, 2012, p. 137). In the present investigation, I group the criteria and factors impacting lexical implantation under the umbrella term **lexical opinion**.¹⁷

In empiry, researchers have examined functional selection in particular. Researchers of socioterminological studies, for instance, have concentrated on speakers’ opinion about specific lexical items or some of their properties (Gouadec et al., 1993; Nogué Pich & Vila i Moreno, 2007a, 2007b). A small number of researchers have tentatively gathered folk speakers’ metalinguistic statements about lexical items through interviews (Leblanc & Bilodeau, 2009). More integrative approaches include the empirical verification of possible lexical

15 Translation scholars (e.g. Künzli, 2001) have perhaps been the only ones to have investigated the question of the use of lexical sources, but only from the perspective of translators, i.e. language professionals, and only during the translation process.

16 Nettle’s paper applies to language variants in general. Here, I am transposing his statements to the lexicon.

17 This concept is further developed in Section 3.3.2.

implantation factors in corpus (Montané March, 2012)¹⁸ or the establishment of a relation between language attitudes and lexical usage (Triano-López, 2007).¹⁹

What is most striking about previous research about lexical impact is the repeated use of **intervention-generated data**.^{20, 21} Data have often been collected out of context and in the presence of the researcher(s). Such a methodological approach raises a whole range of validity issues (researcher effects, reactivity, observer effect,²² etc.; Speer, 2002, p. 511). Silverman (1998) mentioned, however, that “the particular strength of qualitative research ... is its ability to focus on actual practice *in situ*” (p. 3). There have been two main approaches to the collection of qualitative data: the collection of intervention-generated data and the collection of **naturally occurring data** (Ritchie & Lewis, 2003, p. 34). Also, if language managers must act quickly, planning and conducting interviews and surveys do not seem to be the best approach.

I would like to suggest that it is time to start studying lexical opinions using naturally occurring data to avoid the observers’ effect and to gather data quickly from the speech community. By naturally occurring data, I mean here that the recording of the data is “situated as far as possible in the ordinary unfolding of people’s lives, as opposed to being prearranged, set up in laboratories, or otherwise experimentally designed” (Hutchby & Wooffitt, 2008, p. 12). My investigation wishes to explore speakers’ lexical opinions *in context* and in the absence of both researchers and language managers.

Objective 2: Explore speakers’ lexical opinions in context using naturally occurring data

-
- 18 Montané March tested only six implantation factors: degree of dissemination, context of use (institutional versus academic or other contexts), length of lexical item, proximity of the lexical item with its donor language, transfer of an existing form (from general language or another domain), and lexical competition.
 - 19 In my opinion, the validity of this piece of research should be taken with a grain of salt since sound research approaches for language attitudes have been said to be missing (Arendt, 2010, p. 10; Casper, 2002, p. 95).
 - 20 Montané March (2012) is one of the few exceptions.
 - 21 By intervention-generated, I mean that data come into existence due to the researcher (which is the case e.g. for interviews), as opposed to naturally-occurring data, which exist whether or not the researcher studies them.
 - 22 See Cukor-Avila (2000) and Labov (1972, p. 109) for a discussion of the observer’s paradox.

As not only the researcher's absence but also speed is the main guiding principle for my approach, I propose devising a proof of concept based on **natural language processing** methods to capture naturally occurring data.

Collecting naturally occurring data may have been difficult in the past. Spheres of natural language were difficult for the researcher to access, and the technical means to record large collections of speech data were not yet available. Nowadays, I believe that a shift toward a methodology based on naturally occurring data is possible for three main reasons.

Firstly, speakers talk about their language, and they do so not only in private or corporate settings but also on the Internet (see, for example, Reyes & Bonnin, 2016), where a researcher can observe them while being invisible. There are **electronic networks of practice** on the Web where speakers create and/or discuss lexical items in collaborative settings. Gathering what speakers say about language—**metalinguistic statements**²³—in such settings seems to be of the utmost interest in the context of lexical implantation because speakers comment, for instance, on how lexical items are valued or ranked (Rodríguez Penagos, 2004, pp. 34–35).

My second reason for a shift toward naturally occurring data is the fact that a sound theoretical framework has been developed for the concept of metalanguage (i.e., what speakers say about language; Preston, 2004; Rey-Debove, 1978, 1985). Over the last few years, one has been able to successfully observe metalinguistic statements in extensive real language data (Achard-Bayle & Lecolle, 2009; Picton, 2014).

Lastly, Rodríguez Penagos developed a computer application to extract explicit metalinguistic information from large corpora (2004b), enabling a partly automated identification of metalanguage in large sets of data. The application was not designed for the study of lexical implantation,²⁴ but I will show that it can be adapted toward that purpose.

1.3 Context

As language is a social phenomenon, factors affecting lexical implantation might be highly context—but also language—dependent. In the present work, I limit its

23 I will further elaborate on metalinguistic statements in Sections 3.4.2 and 7.2.1.

24 Rather for “a variety of academic and technological tasks [...] from updating computational lexicons to driving graphical representation of conceptual change in Science.” (Rodríguez Penagos, 2004, p. VII–163)

scope to a single speech community. It does not pretend to be exhaustive but seeks to provide the foundations for a novel approach. I selected the speech community that appeared to be most adequate for developing a proof of concept for working with naturally occurring data. The choice fell on the Esperanto speech community²⁵ and was guided by two reasoned arguments.

Firstly, in the language community studied, speakers should have a tendency to explicitly form and express their opinion on language matters. This is to ensure that I as a researcher can collect **explicit metalinguistic statements**. Esperanto is well suited because its speakers generally cultivate their language and are highly critical about it. This is probably due to sociological motives: The members of the language community know that the overall social, psychological, and axiological structure of the community would be endangered if language abilities were not fostered (Rašić, 1994, p. 165). Most Esperanto speakers have a clear opinion on language matters and openly express it to their fellow speakers.

Secondly, this may seem self-evident, but there should be enough material for me as a researcher to examine. The Esperanto speech community brings a significant advantage in this regard: A large amount of lexicon-related metalinguistic statements are available. Many language resources have been produced through wide-scale collaboration (Schweder, 1999) and, with the ascent of the Web era, electronic networks of practice discussing lexicon-related topics have emerged. This can be explained by the fact that Esperanto as a language is not supported by any linguistic or economic power and that the speech community disposes of very few professional structures. Generally, speakers can only rely on themselves (Sakaguchi, 1998, p. 296) and often undertake activities on a voluntary basis.

Through this investigation, I take advantage of the **electronic networks of practice** of the Esperanto speech community. These networks are self-organized groups of speakers on the Web that help each other and share perspectives on specific issues. I examine two types of networks active in the lexical domain: three networks that seek to compile a collaborative dictionary on the one hand and two networks that try to solve ad hoc language-related problems on the other. Both general and specialized languages are of concern because I seek to capture a large spectrum of cases.

In the three networks seeking to compile a collaborative dictionary, members search for lexical items in existing sources or coined *ab ovo*, and they discuss their adequacy between network members. In the two networks solving punctual problems, among other things, members request information or opinions about lexical

.....
25 The speech community is presented in more details in Chapter 4.

items.²⁶ In the course of their discussions, members of these two types of networks give away two kinds of information that are of particular interest for the study of lexical implantation. Firstly, they reveal the sources where they find lexical material,²⁷ disclosing information about their lexical environment and, at times, their penchant for certain sources. Secondly, they also indicate which lexical items they personally use or would use and often explicitly explain their lexical choices²⁸ to their peers, providing information about their lexical opinion.

1.4 Overall methodology and research questions

To achieve Objectives 1 and 2, I have adopted a multimethod qualitative approach. In the present state of knowledge, as rather little is known about the lexical implantation process, I chose a qualitative approach because qualitative methods are best suited for studies of an explorative and descriptive nature (Boeije, 2010, p. 32). Here, I outline the methods that were employed during the research.

To start exploring speakers' lexical environment (Objective 1), I conducted focus group interviews with 31 participants (Esperanto speakers). In the focus group study, I examined the following two research questions:

Research Question 1:

What do speakers do when they need a (new) lexical item?

Research Question 2:

How do speakers perceive nonprofessional dictionaries?

.....
26 E.g. "How would you say [...] in Esperanto?"; "Is there an established expression for [...]?", etc.

27 E.g. "I found it in Kondratjev's dictionary"; "I found this form on the web"; "I found the word [...] in Túlio Flores' dictionary", etc.

28 E.g. "it appears 1,640,000 on Google"; "it is short and more or less pronounceable"; "it is well established"; "[...] and [...]" are both laconic Anglicisms, in fact they are slangs that are not immediately understandable out of context", etc.

I developed the first question to examine whether speakers report referring to some kind of lexical source when they need a new lexical item and, if such is the case, which source (e.g., institutional dictionary, website, collaborative dictionary, etc.). I used the second question to assess whether speakers consider using non-professional dictionaries and under which circumstances, following the idea of exploring the **potential of collaborative dictionaries** engaging nonprofessionals that was put forth by Quirion (2012). I analyzed the collected data using prevailing content analysis methodologies.

I chose focus groups as an initial method toward my first objective because they are “particularly useful for exploratory research when rather little is known about the phenomenon of interest” (Stewart, Shamdasani, & Rook, 2007, p. 41).²⁹ I conducted the corpus study in **five online networks of practice** active in the lexical domain.³⁰ I gathered around 70,000 contributions³¹ exchanged between the members of the networks. I approached the first objective of the thesis (i.e., **speakers’ lexical environment**) based on in situ data and posed the following, complementary research question:

Research Question 3:

What sources of information do speakers use
when discussing a lexical item?

Through this third research question, I provided complementary, context-based findings about speakers’ lexical environment. I identified which sources speakers use to support their argumentation when discussing lexical items and, when data were available, the criteria on why they might prefer a given source.

I further used this corpus study to attain the second objective of the thesis (i.e., observing **speakers’ lexical opinions in context**). This part of the corpus study was guided by the following research question:

29 More details are provided on the methods and focus group rationale in Section 6.2.

30 See Chapter 5 for details.

31 See Chapter 7 for details.

Research Question 4:
 What criteria do speakers use
 when evaluating and choosing a lexical item?

Through this question, of a descriptive nature, I sought to gather the largest possible range of criteria that speakers use in effect and verbalize to their peers when discussing lexical items. To this end, I developed a proof of concept based on naturally occurring data.³²

Summarizing all of the above, to reach its three objectives, I used two complementary qualitative methods in the investigation: an exploratory online focus group study, which produced intervention-generated data, and a descriptive corpus study, which collected naturally occurring data. I obtained relevant naturally occurring data through an innovative natural language processing methodology combining research on metalanguage and autonomy, as well as opinion mining and sentiment analysis. Figure 1 shows the overall thesis methodology.

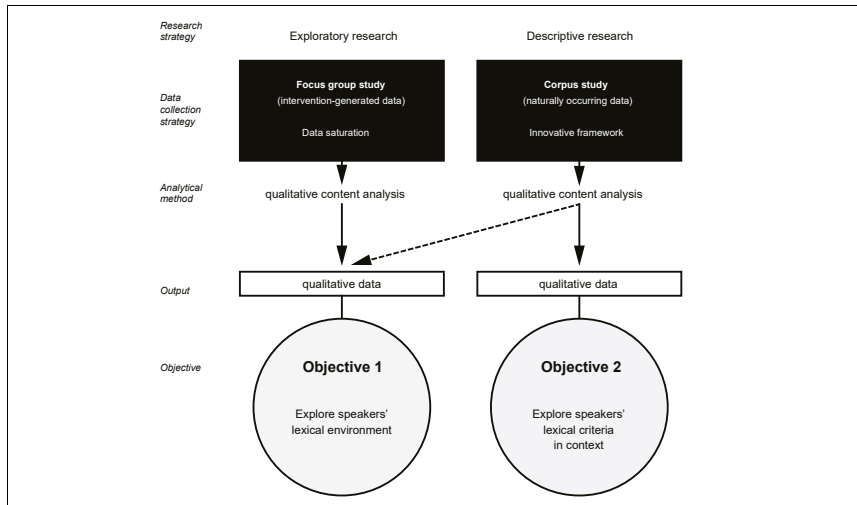


Figure 1. Overview of the thesis methodology.

.....
 32 See Chapter 7 for details.

1.5 Thesis structure

I structured the thesis as follows:

Part 1: Theoretical background

In **Chapter 2**, I present the different forms of deliberate lexical interventions that language managers may make on the lexicon. I discuss the extralinguistic and linguistic goals of these interventions and define the notion of the interventions' effectiveness. I argue that speakers are the source of lexical change and, therefore, that language managers should overcome obstacles at the speaker level if they want their interventions to be effective. Based on existing theories, from aim to achievement, I propose to see the road for language managers as a three-step process taking place at the speaker level.

In **Chapter 3**, I explain the applied science problem that language managers are facing. I review previous studies whose researchers attempted to evaluate and understand lexical implantation phenomena within target speech communities, and I highlight the current research gaps. I propose to deepen research on speakers' lexical environments and to explore lexical opinions in unaided contexts using naturally occurring data.

Part 2: Esperanto for scientific research

In **Chapter 4**, I set the stage in which I conducted its investigations (i.e., the Esperanto speech community). I follow a dual objective: (a) present the main characteristics of the community and (2) show why some of these characteristics make the community a particularly relevant object of study for the present investigation.

In **Chapter 5**, I introduce the five electronic networks of practice examined in the present thesis and, through an online survey, determine the linguistic knowledge and language proficiency of a subgroup of network contributors.

Part 3: Empirical investigation and proposal

In **Chapter 6**, I empirically explore speakers' lexical environments through a focus group study. I address the strategies and sources speakers employ when they need a new lexical item, and I explore the potential of nonprofessional dictionaries among speakers.

In **Chapter 7**, I present the proof of concept created for observing speakers' lexical opinions in context. It explains the aspects related to corpus compilation, detection of metalanguage and autonymy, the filtering of relevant units, and the analysis of metalinguistic and autonymical units.

In **Chapter 8**, I present results obtained analyzing **metalinguistic statement with an opinionated autonym**: a collection of 23 lexical criteria speakers used in arguing in favor or against the use of a specific lexical item.

In **Chapter 9**, I discuss the results obtained in Chapters 6 through 8 and explain what types of data were obtained and why these are relevant for language managers conducting deliberate lexical interventions.

Conclusion

In **Chapter 10**, I remind the reader of the objectives of the present investigation, summarize the main findings, and outline perspectives for future work.

Part 1

Theoretical background

2 Deliberate lexical interventions and their effectiveness

2.1 Introduction

“What’s in a name? That which we call a rose by any other name would smell as sweet” (*Romeo and Juliet*, Act II, Sc. II). How we name things might be unimportant according to Shakespeare. However, human beings have always had the tendency to assess the processes taking place around them and try to manipulate them, and language is no exception to this general principle (Landrø, 2008, p. 88). Language managers in their different forms have tried to *deliberately* choose names for new or existing concepts, an intervention I here call deliberate lexical intervention.

According to Schubert (2014, p. 203), studying interventions on language implies consideration of five aspects: the actors, the instruments, the objects, the goals, and the effectiveness of an intervention. But Gazzola underlined that in language policies some aspects have priority:

Generally speaking, we could say that there is a logical sequence in the questions to be addressed in language policy, and that some questions come first. First, we have to address the questions of “why”—that is, for what reasons a policy should be undertaken—and “what”—that is, what should be done—, before turning to the question of “how” to do it (Gazzola, 2011, p. 114)³³

Accordingly, I begin the present chapter by presenting the “why,” that is, the extralinguistic goals of deliberate lexical interventions (2.2), before addressing what is being done on language, the “what” or linguistic goals of language managers (2.3). In a third Section (2.4), I explain how language managers can turn their *in vitro* lexical items to *in vivo* lexical items—in other words, how they make the speech community use their target lexical items. This explains why speakers can be considered the source of lexical change on which language managers should focus their efforts. In the last section (2.5), I discuss the effectiveness of deliberate lexical interventions.

33 Gazzola mentions the source of this statement is a working paper (Grin & Gazzola, 2007) authored by himself together with François Grin (first author). However, according to the DYLAN project website, this working paper has not been made public (Dylan Project, 2006), wherefore I quote Gazzola and not Grin and Gazzola.

I would like to start by defining deliberate lexical interventions and language managers. According to Wüster, the founder of the General Theory of Terminology (1968, p. 40),³⁴ there are three approaches to language: a) using language, b) describing language, and c) [deliberately] intervening in language. Most people, particularly writers, journalists, and translators, merely *use* language.³⁵ A few others, including lexicographers, terminographers, and language historians, *describe* language. They report the current language situation, or account for changes in language. The last group of individuals, language planners, terminologists, and neologists, *intervene* in language. Here, I group these various actors intervening in language under the term language managers. Language managers can be individuals as well as private or state agencies (D. Blanke, 1985, p. 42; Cobarrubias, 1983, p. 58; Moser, 1967, p. 23). They often work in groups, for instance under the umbrella of a language academy (see Suchowolec, 2018, pp. 46–48).

The concept of interventions on language was extensively discussed in the scientific literature (Betz, 1960; D. Blanke, 1985, p. 1862; Chansou, 2003, p. 19; Glunk, 1967a; Ischreyt, 1965; Moser, 1967, pp. 27–31; Schubert, 2011, 2014; Suchowolec, 2018). Concerning the lexicon, D. Blanke (1985) offered a non-exhaustive list of interventions (p. 42).³⁶ This plurality of intervention types is further reflected in the names under which specific lexical interventions appear in the scientific literature:³⁷

- *elaboration*ⁱ or *linguistic elaboration*ⁱⁱ (Goebel, 2012, p. 279)
- *guided neological practice*ⁱⁱⁱ (Gardin, 1974, p. 68)
- *lexication* (Kaplan, 1997, p. 42; Nahir, 2002, p. 272)

34 Wüster was initially not a linguist but an engineer specialized in electrical engineering (W. Blanke, 2013, p. 85). However, his greatest scientific contribution was in language research (Felber & Lang, 1979, p. 21) and his PhD dissertation on international language standardization is considered the founding work of the General Theory of Terminology (Felber & Lang, 1979, p. 15). The relevance of his work has been challenged, especially in France (see e.g. Humbley, 2004), but to me some of his work is still relevant today, for instance his discussion of effectiveness and actors of language change (Wüster, 1931/1970, p. 123–129).

35 Evidently, by using language, speakers contribute to changing it, as I shall discuss again in Section 2.4.3. However, this action on language from the part of mere language users is unintentional, unplanned and unconscious (Keller, 2003, p. 29).

36 Daoust (1986) also discusses conscious interventions on the lexicon in general terms.

37 Here and later in my work, I am providing equivalents of terms used in other languages. If no other indication is given, the proposed equivalent are mine. Terms in the original languages are given in the endnotes.

- *lexical expansion* (Adegbija, 2014, p. 101)
- *lexical modernization* (Nahir, 1977, p. 107)
- *modernization* (Coluzzi, 2007, p. 132; Haarmann, 2012)
- *neological planning* or *neological assistance*^{iv} (Quemada, 1971, p. 142)
- *planned neology*^v (Candel, 2005; Herrera, 2005; Lorente, 2013, p. 8; Mayar, 2005; Quirion, 2012, p. 131)
- *planned lexical innovation*^{vi} (Boulanger, 1984)
- *planned terminological change*^{vii} (Daoust, 1986)
- *planned terminology*^{viii} (Lorente, 2013, p. 2)
- *term planning* (Bhreathnach, 2011, p. 21)
- *vocabulary expansion* (Kaplan, 1997, p. 38)

In the present investigation, I call language managers' interventions **deliberate lexical interventions** or deliberate lexical interventions: *intervention*, because language managers try to change a given situation; *lexical*, because my investigation focuses exclusively on interventions concerning the **lexicon**, or perhaps more precisely on the lexical usage of a speech community;³⁸ and *deliberate*, because the intervention is undertaken with a specific goal.³⁹ I intentionally avoid the adjective "planned" because planning is an act that refers to the instruments used for interventions, which I am not discussing in the present work.⁴⁰

Lexical interventions have been undertaken with a variety of instruments. At the level of a large speech community, an intervention can for instance consist of active dissemination of target lexical items by language agencies,⁴¹ (i.e., a set of activities seeking to make the target lexical item known to the target speech community) or, more generally, to promote the target lexical item (Hermans, 1994,

38 'Lexicon' can be understood as the theoretical set of all the words of a language (see Cartoni, 2008, p. 17), but language managers intervene on the lexicon that is actually used by a target speech community. Thus, lexicon must here be understood as the set of all the words that are being used or that could be used by the target speech community.

39 I prefer to speak of *deliberate* or *intentional* interventions rather than *conscious* interventions, following Keller's remarks on the psychological character of consciousness and the association between goal and intention (2003, p. 26–29).

40 Evidently, interventions of language managers are usually planned, but 'deliberate' should not be automatically equated with 'planned', see Keller (2003, p. 28).

41 In French it is sometimes also referred to as 'implantation' (e.g. Depecker, 1997, p. VII; Hermans, 1994, p. 40), but I prefer to speak of dissemination to avoid any ambiguity.

p. 40).⁴² Such activities have typically been suggestive (e.g., offering free dictionary copies to members of the target speech community) or prescriptive (e.g., a law prohibiting the use of a foreign lexical item when there exists an approved indigenous lexical item in specific contexts, e.g. advertisement⁴³). Detailing the activities undertaken falls out of the scope of the present thesis.⁴⁴

2.2 Why: Goals of an extralinguistic nature

Language managers try to influence communicative action or communication means in a deliberate manner according to specific goals (see Schubert, 2014, p. 203). Their deliberate lexical interventions can be observed in a large range of sectors such as language planning, terminography, planned languages, controlled languages, computer-aided translation, or content management (Schubert, 2009, pp. 127–128). Deliberate interventions on speakers' lexicon are undertaken for motives of very heterogeneous natures.

Interventions can be undertaken for ideological motives, such as to purify a language (Landrø, 2008, pp. 30–33).⁴⁵ As Cobarrubias notes, in this case the aims of language managers are not philosophically neutral (Cobarrubias, 1983, p. 41). They are driven by ideologies such as linguistic assimilation, linguistic pluralism, vernacularization, or internationalism (Cobarrubias, 1983, p. 63). In Quebec, for instance, language managers have tried to purify terminology (see for example Martin, 1998, p. 15). By intervening on the lexicon, language managers can also pursue functional motives, for instance to modernize a language (Ricento, 2000,

42 In some cases, the intervention on the target speech community must imply much more than merely making the target lexical item known to the target speech community, because especially in cases of conscious multilingual secondary lexical creation, target speech community speakers might not even know the underlying *concept*. As Slodzian mentions, there is no point for instance for an Inuit fisher to get the lexical item 'cancer' if he does not understand the concept to which the lexical item is referring (1995).

43 See for instance Walsh on the Bas-Lauriol and Toubon laws in France (2016, p. 36–38).

44 The interested reader can only refer to few pieces of research (e.g. Ballarin, 2009; Chansou, 2003; Glunk, 1967b; Nissilä & Pilke, 2017). In specialized languages it seems that there has not been much research on the marketing of new lexical items (Nic Pháidín et al., 2010, p. 955) and, in general language, research on normative lexical items appears to be almost inexistent (Freixa i Aymerich, 2015, p. 66).

45 There can be in turn various underlying motives for the purification of a language, see e.g. Lipczuk (2007, p. 19–26).

p. 200) or to revive, spread, or maintain it (Nahir, 2002, p. 273). Lexical interventions may also serve practical motives, for example to allow for cataloguing, comparing, interchanging, or harmonizing concepts (Catalan Centre for Terminology [TERMCAT], 2006, p. 19), to optimize communication (Schubert, 2009; Hermans, 1995, p. 226), to prevent mistakes in documents or their translations (Massion, 2009, p. 28), or to ensure people's security (TERMCAT, 2006, p. 19). They can also be justified by financial motives, such as a decrease in document production or translation time and cost (Herwartz, 2007, p. 5; Massion, 2009, p. 27) using a unique lexical item for a specific concept. This list, though by no means exhaustive, seeks to highlight the great diversity of possible extralinguistic motives. These overall aims of extralinguistic nature depend on the actors involved and their needs and desires.

2.3 What: Goals of a linguistic nature in the speech community

Language managers intervene in two main initial language use situations: one in which a form⁴⁶ is completely missing in the target speech community, and one in which the form is not missing but is considered unsatisfactory by language managers. Accordingly, they follow one of **two linguistic goals: filling a lexical gap or modifying the existing lexicon**.

An explanatory note about the terminology I will be using in the next sections seems necessary here. In a large range of cases, filling a lexical gap or modifying the existing the lexicon implies introducing a new lexical item—that is, a neologism. There exist an impressive number of typologies of neologisms in the scientific literature (Sablayrolles, 2000, p. 71):

46 Here, the focus is especially on interventions seeking a change in the *form* of a lexical item, although, at the level of the lexicon, changes can occur in the *form* of a lexical item, in the *meaning* of a lexical item or in both (Glunk, 1967a, p. 110; see also Ischreyt, 1965, p. 263; Moser, 1967, p. 28; Guilbert, 1975; on lexical changes see e.g. Munske, 2015, p. 22–34). For a discussion of lexical interventions on the semantics of the lexicon, see e.g. D. Blanke (1985, p. 45–46), Glunk (1967a), Martin (1998, p. 53) or Moser (1967, p. 36–37). Since form and meaning are associated, usually both change in parallel through time, see e.g. Picton for specialized languages (2009, p. 41–60) or Mejri's discussion that the distinction between formal and semantic neology has been too rigid (2011). See also Guilbert's remarks on neologisms (1973, p. 18). Thus, one might question whether it is at all possible to intervene *only* on the form or *only* on the meaning of a lexical item. Be as it may, such interventions have been undertaken by language managers. Whether they have been effective is another issue.

Not only are there a great number of typologies that distinguish between a larger or smaller number of subcategories [...], but these are based on criteria that do not fall within the same area: criteria may be radically heterogeneous, which makes it impossible to directly compare one typology with another. (Sablayrolles, 1996, p. 15)^{ix}

What I propose here is an overview of types of deliberate lexical interventions based on a thrichotomic classification:

1. whether the target lexical item is new to the target speech community: I will speak of *creation* if it is and of *selection* if it is not
2. if a target lexical item is being created, whether this new target lexical item originates from another language: I will talk of *multilingual* creation if it does and of *monolingual* creation if it does not
3. whether a lexical item already exists for the concept in the TSC: I will speak of *original* creation if it is not the case and of *alternative* creation if it is

This is a simplification, and in a sense a restructuring of concepts that already exist in scientific works, as I will point out in the following sections (see also Table 1, p. 52, for an overview).⁴⁷

2.3.1 Filling a lexical gap ...

As some of the aforementioned names suggest (i.e., *lexical innovation*, *vocabulary expansion*), language managers may attempt a constructive intervention, usually compensating for a lack of in vivo language development. Their linguistic goal is to fill a lexical gap, that is, the “lack of lexicalization of a certain concept in a given language” (Gregori & Panunzi, 2017, p. 102). To this end, they design a brand new lexical item, a process which is here called *creation* (cases 1, 2 and 3 in **Table 1**, p. 52).

47 Note that I am avoiding the terminology found in Sager (1997) here, because the adjective *secondary* is used for one of two situations, which I find may create confusion: 1) “when a designation is changed at a later date as a result of monolingual revision of a terminology” (1997, p. 27) and 2) “on the occasion of the transfer of scientific and technological knowledge from one linguistic community to another” (1997, p. 27)

For instance, language managers may help in the naming of a scientific or technological innovation, or new processes and events in private industry (Schmitz & Straub, 2010, p. 43), or for finding new product names or descriptions of professional functions.⁴⁸ In the present investigation, this intervention is called **original lexical creation** (case 1 in **Table 1**, p. 52). It occurs when no lexical item previously exists and “accompanies concept formation as a result of scientific and technological innovation or change in a linguistic community” (Sager, 1997, p. 27).⁴⁹ In the late 1980s, for example, the German company Daimler-Benz AG invented a brand-new stroke-controlled wiper system and registered it for a patent under the new lexical item “hubgesteuerte Scheibenwischeranlage für Kraftfahrzeuge” (stroke-controlled wiper system for motor vehicles; Patent-De, 2008).

When the meaning (the concept) considered already exists and has been named in another speech community, this intervention usually receives another name. In the present investigation, it is called **multilingual lexical creation** (case 2 in **Table 1**, p. 52).⁵⁰ This is for instance the case when scientific and technical knowledge is transferred from one speech community to another.⁵¹ It often occurs especially in minority languages, which are influenced by larger speech communities, notably English-speaking communities:

Most of the languages of the world today live in the shadow of English as a dominant language, with concepts, and therefore new terms, reaching them first through the medium of English (Prys & Jones, 2007, p. 7)

An example of a large-scale intervention of language managers with multilingual lexical creation is the case of Malaysia. When the country gained its independence from the UK in 1957 and Bahasa Melayu (Malay) was established as a national and official language, a language agency was created and given the task of coining new lexical items in Malay. With their team, language managers coined about half a million new words by the mid-1980s (Gill, 2013, p. 249) for concepts that mostly

48 See the typical activities of a terminologist in RaDT’s brochure (Rat für Deutschsprachige Terminologie, 2004, p. 2) or in Otman (1991).

49 This is called ‘primary term formation’ by Sager.

50 Again, here I am not coining a new notion, but merely adapting concepts and their denominations found in the scholarly literature more specifically under various names (see e.g. Estopà, Coromina, & Mestres, 2010, p. 19; Hermans & Vansteelandt, 1999, p. 37; Humbley, 2006; Rondeau, 1984, p. 122; Sanz Vicente, 2012).

51 See the second type of ‘secondary term formation’ in Sager (1997).

already existed in other speech communities (mainly English). Malay is not an isolated case. As Ní Ghearáin notes (2011, pp. 307–308, my emphasis),

*Instead of focusing on the selection and standardisation of existing terms, as is possible in more vibrant language communities, **much terminology work in the minoritised language necessarily involves the coining of new terms.***

Another example is the Hawaiian language modernization measures, which included a computer project for which new lexical items were created including the items “hojyouka” (to upload) and “mälama” (to save; McIvor, 2009, p. 3; Warschauer, Donaghy, & Kuamojō, 1997).

2.3.2 ... or modifying the existing lexicon

Language managers may also want to deliberately intervene with the purpose of modifying the lexical items that are in use, i.e. for selecting or replacing existing lexical items. As language is variable, it is not uncommon that a set of multiple lexical items—a **designational paradigm**⁵²—is used to point toward the same referent. This situation is called **formal variation** (see Geeraerts, 1994, p. 80; Geeraerts, Grondelaers, & Bakema, 2010, pp. 155–188) or **denominative variation** (see also Fernández Silva, 2013; Freixa, 2006, p. 51). Formal lexical variation can be tracked back to causes of significantly different nature,⁵³ but it is especially present when new concepts emerge (a phenomenon called “**foisonnement**,” the proliferation of synonyms, in French):⁵⁴

In the period of creation of a new reality and the formation of an adequate vocabulary, it is characteristic of the linguistic situation that a certain, temporary proliferation of neologisms designating the same concept occurs. (Guilbert, 1965, p. 331 my translation)^x

.....
52 On designational and definitional paradigms see e.g. Delavigne (2003), Mortureux (1993) and Reboul-Touré (2004, p. 204).

53 In specialized languages for instance, Freixa (2006, p. 52) identifies five main categories of causes: dialectal, functional, discursive, interlinguistic and cognitive.

54 In the present investigation, quotes from languages other than English are sometimes translated for ease of reading. The quote in the original language, however, can be found in the endnotes.

For instance, in French the concept of *crash cushion*⁵⁵ is designated by multiple lexical items, including *atténuateur d'impact* and *atténuateur de choc*, the former item being used in Quebec whereas the latter is used in France. In a situation of formal variation, language managers may select one of the lexical items in the designational paradigm, an intervention I here call **lexical selection**. In the present case, arguing for an enrichment of the standard French language with variants from Quebec, the Quebec Board of the French Language chose to select the variant *atténuateur d'impact* (Vézina, 2004, p. 6). In a narrow sense, what I call lexical selection is what the literature has called standardization, “the imposition of uniformity upon a class of objects” (Milroy, 2001, p. 531), or harmonization, the “process in which diverse positions are largely reconciled and assimilated into a single unified position” (Gilreath, 1992, p. 138).⁵⁶ I prefer to speak of selection, because in the aforementioned definitions standardization and harmonization equate to a reduction to a *single* item in the designational paradigm, whereas in a lexical selection process one could imagine selecting more than one item. Also, “selection” refers to the consideration of language change as an evolutionary phenomenon, to which I will come back later (see Section 2.4.3).

Finally, if an existing designational paradigm is considered completely unsatisfactory, language managers may choose to create a new lexical item.⁵⁷ This intervention can be either constructive or reductive: The new lexical item might be created as a complement or as a substitute for an existing lexical variant. This type of intervention is made “when a designation is changed at a later date as a result of monolingual revision of terminology” (Sager, 1997, p. 27). In the present investigation, such an intervention is called **alternative lexical creation**.⁵⁸ Such an intervention may be for instance the “result of the discovery of a new entity in the same subject field” (Valeontis & Mantzari, 2006, p. 4)—for example, the coining of “fixed

55 A crash cushion is “a device that is installed in front of a rigid obstacle to absorb the energy of an impacting vehicle.” (Public Works and Government Services Canada, 2016)

56 The meaning of “harmonization” depends on the author and the domain. For instance, in accounting and finance, Fuertes explains that “the main difference between harmonization and standardization processes lies in the degree of strictness of the accounting standards. Harmonization involves a reduction in accounting variations, while standardization entails moving towards the eradication of any variation.” (Fuertes, 2008, p. 327)

57 See the first type of ‘secondary term formation’ in Sager (1997).

58 This is called ‘secondary term formation’ by Sager (1997, p. 27) and ‘néologie d’adaptation’ (in French) by Dury (2013), i.e. using a new lexical item because the existing one is no longer satisfactory.

telephone” to replace “telephone” because mobile telephones were invented. This intervention may also be performed for puristic reasons: language managers may coin ethnic language equivalents for replacing borrowings already circulating in language use (see Martin, 1998, p. 15).⁵⁹ For instance, the concept of marketing was introduced in France in 1950 (Soubrier, 1998, p. 409). The lexical item used by the speech community for this concept was “marketing” (a borrowing from English). Dissatisfied with the use of English language material in French, in 1973 language managers started suggesting the new lexical item “mercatique.”

Table 1 summarizes the two major types of deliberate lexical interventions that have just been discussed according to the initial language situation.

Case number #	Lang. code	Existing designational paradigm	Target lexical item	Type of deliberate lexical intervention	Source of example
1	de	∅	hubgesteuerte Scheibenwischeranlage für Kraftfahrzeuge	1. Filling a lexical gap	original lexical creation (Sturz, 2014, p. 39)
2	haw	∅ [en: to upload]	hojyouka		multilingual lexical creation (McIvor, 2009, p. 3)
3	ca	chaqueta chaquetilla jaqueta peto traje	jaqueta	2. Modifying the existing lexicon	lexical selection (Vila i Moreno & Vila i Moreno, 2007, p. 74)
4	fr	marketing	mercatique		alternative lexical creation (Soubrier, 1998, p. 409)

Table 1. Two major types of deliberate lexical interventions (1. Filling a lexical gap, 2. Modifying the existing lexicon) according to the initial language situation.

As Suchowolec (2018) synthesized, ultimately a deliberate intervention is undertaken with the aim to bring a change in the language usage of a group of people (p. 231). By definition, the general purpose of language managers conducting a deliberate lexical intervention is for a group of people, which I call here the target speech community, to *use* their target lexical items. As illustrated in Figure 2, this

.....
59 I prefer not to give a precise definition of ‘borrowing’. There is a large quantity and variety of taxonomies with which the phenomena linked to borrowing is presented and categorized (Variano, 2014, p. 8), and as Picone notes, it is a hard concept to grasp: To say that *week-end* in French is an Anglicism is uncontroversial. But what of *station-service*, whose elements are French” (1996, p. 1).

means that language managers wish that the language usage A at a given point in time would change through an actualization process to a given language usage B at a later point in time that corresponds to their goal (here, lexical selection, reducing the designational paradigm to *jaqueta* only).

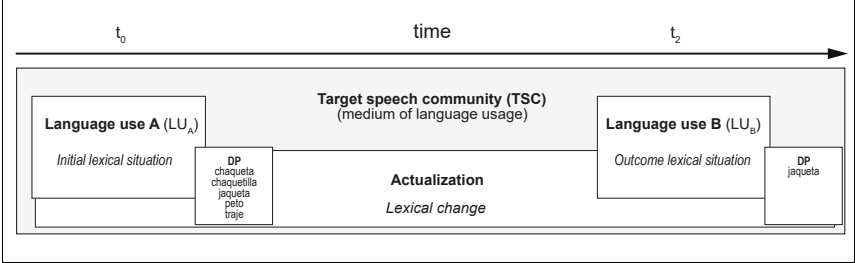


Figure 2. Example of the actualization process toward an outcome lexical situation desired by language managers. After a certain amount of time, a five-item designational paradigm (DP) has been reduced to a single-item designational paradigm by lexical selection.

2.4 How: The route from aim to achievement for language managers

2.4.1 From in vitro to in vivo lexical items

Let me start this section by providing a definition of lexical item: by **lexical item** or lexical item, I mean a unit of description made up of words and phrases that is at least one semantic constituent and at least one word. I am using Sinclair’s terminology here (1998) speaking of “lexical items,” which are also called “lexical units” by other scholars, and I am compiling the definition from Sinclair but also Cruse’s work:

The basic syntagmatic lexical units of a sentence will be defined as the smallest parts which satisfy the following two criteria;

(i) *a lexical unit must be at least one semantic constituent*

(ii) *a lexical unit must be at least one word.* (Cruse, 1986, p. 24)

In Section 2.2.2, I reviewed types of linguistic goals language managers pursue and defined the desired outcomes of language managers. But how do language managers reach their linguistic goals and measure the outcomes of their interventions?

When they create or select target lexical items, language managers do not yet have any kind of influence on actual language use. “Giving names and fixing terms is one thing, but implementing and disseminating them is quite a different thing” (Allony-Fainberg, 1974, p. 495). Target lexical items coined or selected are merely **in vitro lexical items**,⁶⁰ they are outside of their natural medium, which is the speech community. They only live in artificial settings: in the language managers’ laboratory, in a database, on paper, etc.

Ultimately, however, language managers want to see their in vitro efforts turn into an updated in vivo reality of language use (see also the wording in Vila i Moreno & Vila i Moreno, 2007, p. 86). Language managers seek to bring about lasting changes in speakers’ language use, to make speakers use the target lexical items. This requires an intervention on the target speech community. As language is borne by **speakers**,⁶¹ the intervention must take place among them.

This intervention in the speech community is no simple undertaking, because the lexicon is generally assumed to be the most variable and changing component of language (Bossong, 2000; Helfrich, 1993, p. 1; Landrø, 2008, p. 98; Munske, 2015, p. 20). Every **natural language**⁶² is variable (Desmet, 2005) and every natural language changes (Keller, 2003). The reality of language is movement and perpetual creation (Coseriu, 1977, p. 5).⁶³ Planned languages are not exception to this rule, as Ferdinand de Saussure had already pointed out in his course in general linguistics a century ago:⁶⁴

60 See the concept of ‘terme-éprouvette’ in French, e.g. in Bibeau (1983, p. 17).

61 Throughout this thesis, by speakers I mean language users in general, i.e. individuals who use language not necessarily only in oral communication, but also in writing, sign language, etc.

62 **Natural language** is usually defined as a language for human communication that *has evolved* naturally (e.g. D. Liu, Li, & Thomas, 2017, p. 1113). I would give it a broader definition, i.e. a language for human communication that *is or has been* evolving naturally. Planned language that are used by a speech community, such as Esperanto, have properties that make them very similar to natural languages in the strictest sense (see Lindstedt, 2006).

63 See also the concept of ‘perpetual dynamics’ (Beckner et al., 2009, p. 15–16).

64 Language change in Esperanto cannot be fully equated to that of ethnic languages, notably because native speakers are not norm-providing (see Fiedler, 2012). It must also be noted here that Zamenhof, the initiator of Esperanto, did not mean to create an immutable language (see the remarks of D. Blanke, 2010, p. 57–58 (under 3.2)).

Mutability is so inescapable that it even holds true for artificial languages. Whoever creates a language controls it only so long as it is not in circulation; from the moment when it fulfills its mission and becomes the property of everyone, control is lost. Take Esperanto as an example; if it succeeds, will it escape the inexorable law? Once launched, it is quite likely that Esperanto will enter upon a fully semiological life; it will be transmitted according to laws which have nothing in common with those of its logical creation, and there will be no turning backwards. A man proposing a fixed language that posterity would have to accept for what it is would be like a hen hatching a duck's egg: the language created by him would be borne along, willy-nilly, by the current that engulfs all languages. (de Saussure, 1966, p. 76)

Consequently, language managers intervene on a moving object, an object that is constantly transforming even in the absence of a deliberate intervention, of a deliberate lexical intervention. Here, I call factors of language change unrelated to deliberate lexical intervention **environmental factors**.

The intervention in the target speech community serves to guide the *natural* process of lexical change in language, to accelerate a feature of this ongoing process, to oppose it or to reorient it (see Daoust, 1986, p. 249; Steinmüller, 1978). At the lexical level, the aim of this intervention is to foster the use of a target lexical item and/or to discourage the use of competing lexical items in the designational paradigm. The intervention stretches over an extended time frame, sometimes dozens of years when introducing a new lexical item (see the figures in Depecker, 2005, p. 14). In fact, time is an essential factor in lexical interventions: Numerous scholars (Allony-Fainberg, 1974, p. 514; Diki-Kidiri, Joly, & Murcia, 1981, p. 4; Fischer Hubert, 2001, p. 119; Glunk, 1966a, p. 73, 1967a, p. 110, 113; Ischreyt, 1965, p. 217; Martin, 1998, p. 45; Quemada, 1971, p. 138; Quirion, 2004; Soubrier, 1998, p. 409) have warned that interventions on the target speech community may fail if they are made too late in the lexical change process. For example, if for a new concept language managers try to disseminate a specific target lexical item whereas other lexical items have already started to spread within the speech community for a given designational paradigm,⁶⁵ the target lexical item will fail.

65 See Quirion's procedural factors "temps écoulé" (elapsed time) and "livraison juste à temps" (just in time (JIT) production) (Quirion, 2004, p. 198).

2.4.2 The actualization issue

Through their studies, scholars empirically showed that some target lexical items are not used by speakers (not implanted) although they are known to the speakers:

According to our findings, some terms are known that are not used. From a theoretical perspective, this situation is not particularly surprising. It roughly corresponds to the opposition that modern language teachers very well know between passive vocabulary and active vocabulary. (Thoiron et al., 1993, p. 69, my translation and my emphasis)^{xi}

Allony-Fainberg (1974) had already pointed out this reality when she decided to compare claimed knowledge with claimed usage of target lexical items (p. 502).

A question for language managers, thus, is how speakers “decide”⁶⁶ to use or not to use a target lexical item. Here, one thing is certain: Generally, speakers do not simply adopt the language variants that are most common around them, otherwise innovations would never spread within a language community. This is what Nettle calls the threshold problem (see Nettle, 1999, pp. 98–99). This implies that speakers apply some kind of filter, some kind of selection mechanisms to the new lexical items they are exposed to. According to Maras (2015), speakers evaluate expressions using their language feeling,^{xiii} usually unconsciously (p. 77). In doing so, they can choose the “correct” means of expression among those available (Maras, 2015, p. 76). Building on Levelt’s language production model (Bock & Levelt, 1994; Levelt, 1995), Fiehler (2014) explained speakers’ evaluations of language in terms of monitoring. In his explanation, any human act—including a speech act—comes with a certain monitoring from the part of the individual.⁶⁷

The major challenge for language managers on the route from aim to achievement is one of actualization (see Brinton & Traugott, 2005, p. 7): how can the desired lexical change spread through the target speech community, the medium of language use? In his seminal work in the context of technical language standardization, Wüster already questioned whether the interventions had a real impact:

.....
⁶⁶ I enclose ‘decide’ in quotation marks because this is not necessarily a conscious process. In the present investigation, I approach this issue by means of a corpus, thus my focus is not on determining whether speakers are acting consciously or unconsciously. Rather, I observe ‘decisions’ that are explicit in my corpus: See the notion of explicit metalinguistic statement in 3.4.2, p. 114), which I borrowed from Rodríguez Penagos (2004b, p. 1-4–1-5).

⁶⁷ Monitoring can concern various aspects of communication. Levelt mentions e.g. agreement with social standards (choice of register) and lexical choices (1995, p. 461).

Standardized technical language as laid down in published norms represents great progress for the quality of the system. But don't these innovations exist only on paper? How is the call for an efficient language development satisfied? (Wüster, 1931, p. 175, my translation)^{xiii}

Over the years, language managers became increasingly suspicious that deliberate lexical interventions—both those of agencies and those of individuals—were not completely successful. Holz noted that only about 5% of new lexical material suggested by the German writer J. H. Campe made it into the German language (1951; as cited in D. Blanke, 1985, p. 43). Mortureux (1987) stated that probably half of the target lexical items proposed by France's terminology commissions were not adopted by French speakers (p. 250).⁶⁸ For a long time, the outcomes after an intervention were, however, not systematically evaluated.⁶⁹ Ischreyt told us in the mid-1960s in the context of specialized languages:

Hardly anything is known about the successes and the failures of normative terminology. There aren't enough exact studies evaluating whether the normative terminology is actually used as prescribed, in the scientific and semi-scientific literature as well as in factory brochures, nor are there studies considering whether the normative synonym prevails over other synonyms. (Ischreyt, 1965, p. 217, my translation)^{xiv}

Lexical actualization or change has been under the microscope in implantation studies and, on a macro level, in language change studies. In language change studies, the general position is that there is no monocausal explanation for the phenomenon of language change (Landrø, 2008, p. 111), the causes for the propagation of changes in language being very different in nature. Studies on the lexicon arrive at the same conclusion: The diffusion of lexical material is the result of a large range of factors that are *not* related to each other (Hermans, 1994, p. 43; see also the comprehensive framework of potential factors in Costa, 2016). Three

68 It is interesting to note that similar thoughts were expressed in the wider field of language planning. Ricento states that from the early 1970s through the late 1980s “There was a growing awareness among scholars that earlier attempts in language planning [...] were inadequate, purely from a descriptive perspective [...] There was a number of factors that caused the field to reconsider where it was, and where it might be headed.” (Ricento, 2000, p. 201)

69 Probably, also, because it is difficult to evaluate them, as we will see in the next chapter.

principal categories of factors have been identified (see M. Teresa Cabré, 2010, p. 12; Glunk, 1967a, pp. 110–111; Quirion, 2004):⁷⁰

- linguistic factors, both formal and semantic: factors that concern the lexical item itself or language as a system,
- pragmatic/sociological factors such as prestige, ease of use, etc., or language usage as a social phenomenon, and
- procedural/language planning factors, especially dissemination methods.

The first two categories of factors are those I call environmental factors: They are always potentially at work, even in the absence of a deliberate lexical intervention. The third category of factors, which I call factors related to deliberate lexical interventions, is linked to the interventions of language managers. I feel it is essential to make such a distinction to give an accurate definition of the effectiveness of a deliberate lexical intervention, which I will do in Section 2.5.2.

2.4.3 Speakers as the source of lexical change

Language change is the result of human activity (Keller, 2003, p. 85). Croft (2000) proposed to conceptualize it as an evolutionary phenomenon: The entities of language (utterances and grammars) are replicated by speakers (p. 4). From an evolutionary perspective, language change is a *cumulative* process based on mechanisms of variation and selection (Keller, 2003, pp. 195–197).

The macro level emerges from the structure created by the micro level ... the causal consequence from a great number of individual intentional actions that at least partly serve similar intentions. (Keller, 2003, p. 93, my translation)^{xv}

This is a crucial point: The overall system at the macro level emerges from numerous actions at the micro level, the level of speakers as agents:

70 The factors mentioned in implantation studies are comparable to those cited by language change studies. For instance, according to Moser (1967), language change occurs through 1. human-related (psychological, physiological, etc.) driving forces, 2. intralinguistic driving forces as well as 3. circumstances of propagation. Moreover, language change studies often postulate that change occurs through a mix of functional and social selection (Croft, 2000; Nettle, 1999).

I came to understand language as a complex adaptive system, which emerges bottom-up from interactions of multiple agents in speech communities (Larsen-Freem, 1997; Ellis with Larsen-Freeman, 2009), rather than a static system composed of top-down grammatical rules or principles. The system is adaptive because it changes to fit new circumstances, which are also themselves continually changing. (Larsen-Freeman, 2011, p. 49)

The best way to picture the influence of individual actions over the language system as a whole is perhaps to use a model and a simulation. For instance, one can employ the language change model in NetLogo (Wilensky, 2016), which is based on Nettle’s (1999) model using social impact theory. His model is “a framework for simulating language change in social networks derived from Social Impact Theory” (Nettle, 1999, p. 95). This theory was developed for modelling situations in which an individual is influenced by those around him (Nettle, 1999, p. 101).

In Figure 3, the sample simulation starts with 15 speakers (out of 100) using a specific variant (the white dots). This variant progressively spreads. Its use starts by declining, then increases and declines again several times, but finally reaches the totality of speakers after about 180 iterations.

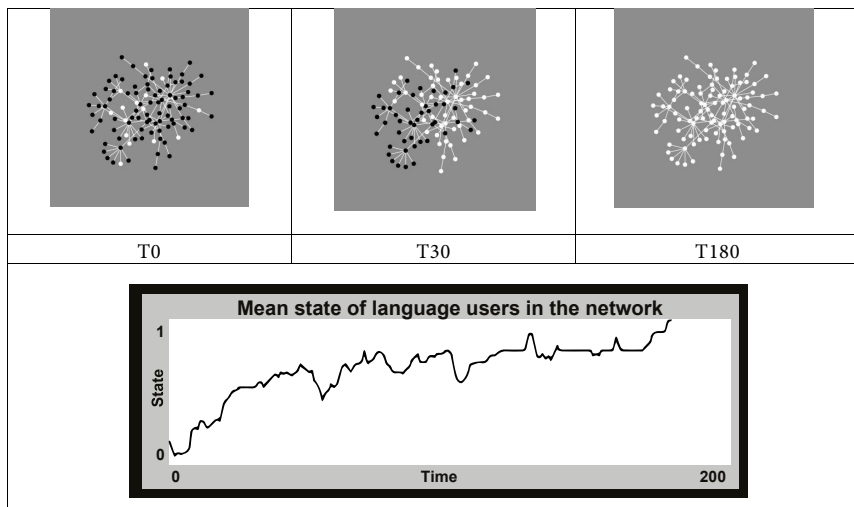


Figure 3. Simulation of language change using the language change model in NetLogo (Wilensky, 2016), based on Nettle’s model using social impact theory (1999). The white variant propagates through the speech community (from 15 to 100 speakers in about 180 iterations).

In other words, language change is caused by *individual* actions of speakers. This implies that lexical change starts at *speaker* level. Thus, implantation of target lexical items can be realized only through the speech community, through its speakers. The speakers are the agents on which language managers should act if they hope to be efficient with their deliberate lexical interventions.

In an eco-system approach to language planning, individual decisions about language use are the ultimate test for the language planner. (Kaplan, 1997, p. 303)

Within the speech community, speakers are individual agents:

Speakers/agents produce tokens of linguistic structure in their interactions; that is, speakers replicate linguistic structures they have heard previously in their utterances, albeit in novel combinations and sometimes in altered forms. (Blythe & Croft, 2009, p. 48)

The question is, therefore, how language managers can potentially act on speakers to trigger the lexical changes they desire. And language managers must act on ordinary speakers, because as already stated, in the end speakers decide whether they will use specific lexical items.

As a matter of fact, “linguists need speakers; the reverse is in no obvious way true” (Moore, 2015, p. 1). In the past, the correct way to speak may have largely been defined by the upper layers of society, by authoritative figures. At the time when France was a monarchic society, for instance, the lexical norm was that of the words used by the French royal court, and only the king and a handful of renowned writers were in a position to coin new lexical items (Guilbert, 1975, pp. 50–51). That era is now history.

In the current age of the collaborative web and with the increasing digitization and uberization of society, the models that survive are often only those that are actively supported by their consumers. This is a general trend: Most individuals no longer trust traditional media, and with the advance of social computing, the power is shifting from institutions to communities (Charron, Favier, & Li, 2006). Language is no exception to this:

... The relationship with traditional sources of authoritative discourse has changed: Wikipedia, “the free encyclopedia”, has replaced printed dictionaries and encyclopedias for many (maybe most) internet users. The

development of new internet-based devices of authoritative discourse is currently taking place in many fields; especially in the field of language.
(Bonnin, 2012, p. 8)

Bonnin (2014) spoke of a *change of sociolinguistic era*^{xvi} (p. 352). The prescriptive, top-down language tradition is actively challenged by ordinary speakers.⁷¹ Nowadays, “there is no need to be a linguist nor an academician to judge on good usage and norms. One merely needs to connect to the Internet” (Osthus, 2003, p. 140, my translation).^{xvii}

All speakers feel entitled to speak about language: they compare their usage with that of others, they look for similarities and differences, they make judgments as to the quality of usage (Rey-Debove, 1978, p. 82). Speakers can publicly talk about language norms, approving or disapproving of them (Reyes & Bonnin, 2016, p. 2), and they do: Ordinary speakers comment not only on the language usage of their fellow citizens (Wilton & Stegu, 2011, p. 12), but also that of individuals that long functioned as traditional language authorities, such as professional writers or teachers. For instance, Web users do not hesitate to post comments explicitly criticizing journalists’ use of language.⁷²

Ordinary speakers’ statements on language can trigger language change. Although ordinary speakers do not *plan* their language per se, their reflections may have an influence on their language usage (Kabatek, 1996, p. 43; see also Beuge, 2014, p. 129; Rousseau, 2007, p. 66), and on that of other speakers: Individuals who explicitly criticize language can consciously or unconsciously become normative (Heringer & Wimmer, 2015, p. 88), and there is growing evidence that metalanguage exerts a considerable influence on language (Mertz & Yovel, 2003). Our modern media and mass communication are accelerating these types of processes (see Ammon, Dittmar, Mattheier, & Trudgill, 2008, p. 1616).

Consequently, ordinary speakers are becoming a new type of normative agent with an increasing influence on language (Reyes & Bonnin, 2016, p. 5), although their influence is often neither conscious nor intentional. Hundt (2010) went so far as stating that “speakers have a greater influence on the emergence, establishment and elimination of language norms than language codices, norm authorities, language experts and model speakers” (p. 50). Traditional language norms are

71 By ‘ordinary speakers’ I mean speakers who are not language professional. This equates to what Preston calls ‘folk speakers’ (2011, p. 11).

72 See e.g. the study undertaken by Jacquet and Rosier (2014) or that of Arendt and Kiesendahl (2015).

furthermore shifting as ordinary speakers unprecedentedly use new sources as their language reference materials: not only imaginary figures of reference (Meunier & Rosier, 2014), but also online forums, dictionaries, wikis, blogs, machine translation systems (Bonnin, 2014, p. 358), and even commercial search engines:

In online discussion forums concerned with language issues, it can be observed again and again that participants use the Internet, and especially commercial search engines like Google, to check the occurrence of certain words, combinations of words or grammatical constructions. Helped by the proofs of usage and/or the calculated frequencies that are generated in this way, they establish theses or verify them. (Kunkel, 2015, p. 201, my translation)^{xviii}

In the era of the participative web and social media, “people use technologies to get the things they need from each other, rather than from traditional institutions” (Li & Bernoff, 2011, p. 9). More than ever, linguistic power lies to a large extent in the hands of the speech community, and its ordinary speakers make use of a set of tools and normative references that are not necessarily those of language managers.

In the current state of knowledge, the effectiveness of lexical interventions cannot be directly measured. However, lexical changes can be expected to result from small, incremental changes at the speaker level (see Beckner et al., 2009, p. 17). There is no achievement of lexical implantation without speakers, because language is carried by a medium: the speech community.⁷³ Thinking in terms of lexical implantation forces us to be concerned with the transition from speech to language (Gaudin, 2007, p. 32), and in this transition the speaker is the interface, or the “interactor” as Croft (2000) expressed it (p. 54). It is the agent through which language is uttered, through which the language system is transferred to language use, through which a language policy is interpreted, appropriated, or ignored. As Ricento (2000) noted,

It seems that the key variable which separates the older, positivistic/technicist approaches from the newer critical/postmodern ones is agency,

73 Language change is a collective phenomenon and is characterized by the fact that populations are involved in it (Keller, 2003, p. 25).

that is, the role(s) of individuals and collectivities in the processes of language use, attitudes, and ultimately policies. (2000, p. 208)

In the end, it is speakers who use or do not use the lexical items that language managers would like to implant (Loubier, 1994, p. 20). Despite the host of new trends (socioterminology, pragmateterminology, etc.⁷⁴), the key variable “speaker” does not seem to have received the central role it deserves. A target lexical item will only become part of the language if speakers use it actively, if they utter it in speech and in writing. I am not alone in deploring this research gap. Depecker (2013), for instance, also regretted that so far socioterminology has not taken the motivations of speakers into account (p. 18).

Speakers replicate items they have previously heard (or read, seen, etc.). This sounds like a statement of the obvious, but is essential because target lexical items are often new forms for speakers.⁷⁵ Speakers must become acquainted with the new form before they may replicate it. This is why it is also essential to examine lexical sources that are present in speakers’ environments. By **lexical source** I mean any person (a fellow speaker) uttering lexical items to the target speaker or any medium (a dictionary, a document, a movie, a website, etc.) to which a target speaker is exposed and from which they receive lexical items.

Speakers will not necessarily replicate what they have heard, so there is also a process of selection at stake. In her model to determine the acceptance of French neologisms, Helfrich explains this selection process by means of accessibility and usability in two steps:

74 In my view, they are complementary approaches to one and the same discipline (terminology) and would not necessarily each justify a separate name, just like Hymes once noted for the field of linguistics: “Ethnolinguistics’, ‘psycholinguistics’, ‘sociolinguistics’—these, and the older standby, ‘language’ and ‘culture’, are the chief terms by which one or another common cause between linguistics and other fields, such as anthropology especially, has come to be known in the period since World War II. ‘Linguistics’ itself would do, of course, if linguists generally would agree to such a scope for the discipline. Such an event seems unlikely, and composite terms are likely to prevail for some time, wherever something of concern both to linguists and others is in question.” (1964, p. 2)

75 Except in cases of conscious lexical selection, the target lexical item will be new to speakers of the target speech community. The target lexical item may not be new at the level of the speech community, but any lexical item that is perceived as new by a speaker—and if they’ve never been exposed to it, it is new—is a neologism to them. See Guilbert (1965, p. 135) and Martin (1992, p. 36).

The first, significant filter is situated on the level of accessibility. It comprises—in decreasing order of importance—the criteria “customary nature,” “familiarity,” “simplicity” and “comprehensibility.” If a neologism passes this first filter, its “passive acceptance” is highly probable. (Helfrich, 1993, p. 292, my translation)^{xix}

The second filter, usability, is linked in particular to a speaker’s evaluation of usefulness of a neologism with regard to its use in a concrete communicative situation, which is closely connected to the criterion of “adequacy.” In this regard, the source that produced the neologism also has its importance. Moreover, on this second level, the criteria “correctness,” “aesthetic qualities” and “normality” have an influence to a lesser extent. (Helfrich, 1993, p. 292–293, my translation)^{xx}

Other linguists (see Martin, 1992, 1994; Quirion, 2014b, p. 110) have suggested drawing a parallel between the process leading to the use of target lexical items and Rogers’ theory of diffusion of innovations (1983), which results in a *three*-step process. As Martin (1992) noted, although the theory of diffusion of innovations has been developed outside of the discipline of linguistics, it entails anthropological as well as sociological aspects: It therefore allows for a broad generalization and is consistent with sociological observations (p. 34). In this theory, the innovator⁷⁶ “passes from first knowledge of an innovation, to forming an attitude toward the innovation, to a decision to adopt or reject, to implementation of the new idea, and to confirmation of this decision” (Rogers, 1983, p. 165). The same applies to language:

*The implantation process of innovations essentially consists in a series of choices and actions that lead an individual or an organization to **take note of an innovation, to develop positive or negative attitudes toward it, to make the decision to adopt or reject this innovation, to give effect to that decision in a concrete way and, finally, to maintain or change this decision.** (Martin, 1994, p. 33, my translation and my emphasis)^{xxi}*

76 In the linguistic literature, speakers introducing new variants have also been called ‘introducers’ (Croft, 2000) or ‘innovators’ (Milroy, 1992). Interestingly, in all the companies she studied, Daoust noted that some specific employees played the role of innovators for terminology (1986, p. 249).

Other models, especially in marketing studies, also propose to illustrate subsequent steps leading to a specific, desired behavior of an individual, such as the “path to purchase” (Jones & Runyan, 2016; Song, Sahoo, Srinivasan, & Dellarocas, 2016), the “buying funnel” (B. J. Jansen, 2011), or the “brand purchase funnel” (Dierks, 2017). Based on Martin (1994), the adoption of a new lexical item can be depicted as a three-step process:

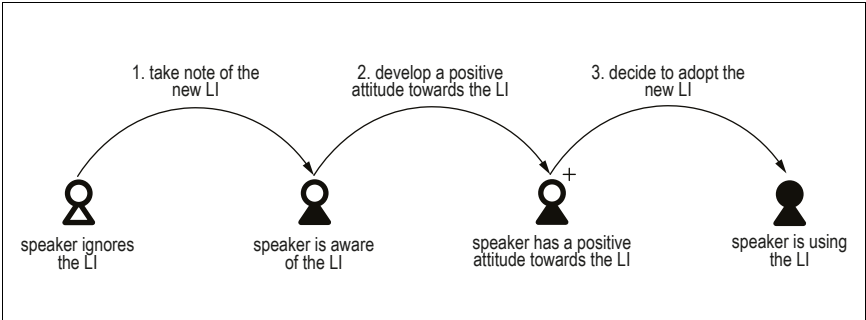


Figure 4. Three-step implantation process of a lexical item (LI) from a speaker’s perspective, based on the ideas expressed in (Martin, 1994, p. 33).

Each step is a necessary precondition for the next one. Using this three-step process, one can represent possible outcomes of deliberate lexical interventions at speaker level, as illustrated in Figure 5 with four scenarios for the implantation of the lexical item *mercatique* to replace “marketing” (an Anglicism) in French. In the best-case scenario for language managers (1.), the speakers will take note of the existence of the target lexical item *mercatique*, will develop a positive attitude toward this target lexical item and will use it in their interactions with fellow speakers. This leads to full implantation of the target lexical item. In a second scenario (2.), the speaker may also come to know the target lexical item, will develop a positive attitude toward it, but will still have a positive attitude toward another lexical item of the same designational paradigm (“marketing”) and therefore will use both lexical items depending on the context. This results in partial implantation of the target lexical item at speaker level. In a third scenario, the speaker becomes aware of the target lexical item, but develops a negative attitude toward it. Therefore, he will not use it, which results in a lack of implantation of the target lexical item. The last scenario proposed here also results in the absence of implantation, this time because the speaker is never exposed to the target lexical item *mercatique*.





Scenario	A. 	B. 	C. - + 	D. 	
	Speaker ignores the lexical item	Speaker is aware of the lexical item	Speaker has an opinion towards the lexical item	Speaker is using the lexical item	OUTCOMES
1.	mercatique	mercatique marketing	mercatique + marketing -	mercatique	Total implantation (B+C+D effective)
2.	mercatique	mercatique marketing	mercatique + marketing +	mercatique marketing	Partial implantation (B+C effective)
3.	mercatique	mercatique marketing	mercatique - marketing +	marketing	No implantation (B effective)
4.	mercatique	marketing	marketing +	marketing	No implantation (B not effective)

Figure 5. Possible outcomes of a deliberate lexical intervention aimed at implanting the target lexical item “mercatique” instead of “marketing” at the level of a single speaker (modification of the existing lexicon: lexical selection).

Thus, there are two major obstacles on the route from aim to achievement for language managers at speaker level: lack of knowledge of the target lexical item (scenario 4) and negative attitudes toward the target lexical item (scenario 3).

Using a management matrix, the three-step implantation process seen from the perspective of speakers (see Figure 1) can be embedded in the perspective of language managers, as the following figure illustrates:

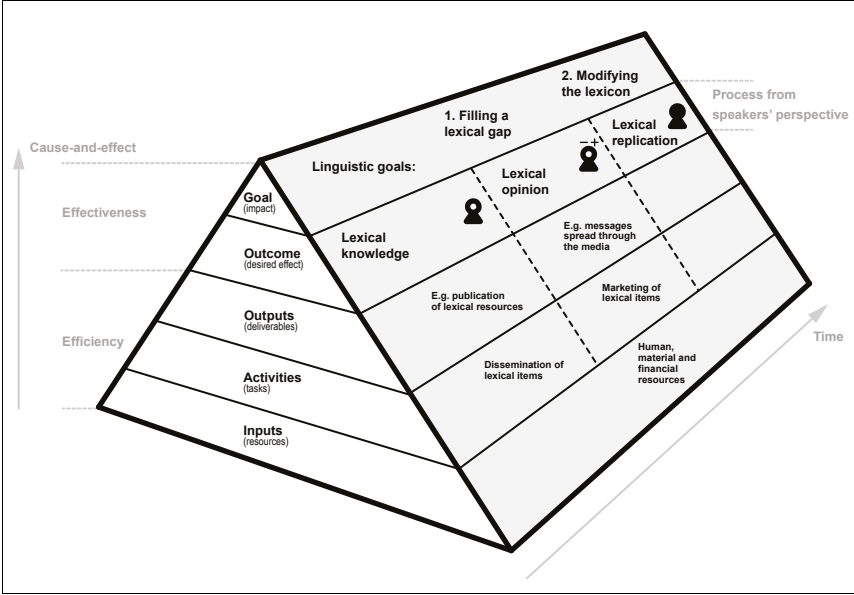


Figure 6. Three-step implantation process of a lexical item from a language manager's perspective, adapted from the matrix in (Crawford & Bryce, 2003, p. 369).

The goal of language managers is to obtain the best possible results for the two milestones in the process, ensuring lexical knowledge of target lexical item among speakers of the target speech community and guaranteeing that speakers develop a positive attitude toward it. From a managerial perspective, language managers may use several activities (or instruments, in Schubert's [2014, p. 203] terms) to increase lexical knowledge and improve the polarization of lexical opinion through time.⁷⁷

.....
 77 The type of activities or instruments are out of the scope of this chapter, but examples are provided in the figure.

2.5 Outcomes and effectiveness of deliberate lexical interventions

2.5.1 Possible and desired outcomes of deliberate lexical interventions

Generally speaking, the linguistic goal of language managers is to achieve the outcome of scenario 1 (see Figure 5) for the greatest number of speakers in the target speech community.

To summarize,⁷⁸ there can be three main states of language usage (three lexical usage outcomes) after a deliberate lexical intervention:⁷⁹

1. Total implantation (standardization⁸⁰): In the outcome state of language use, the target lexical item of language managers is the only one used by the target speech community for the concept considered

2. Partial implantation: In the outcome state of language use, the target lexical item of language managers is used by the target speech community, together with one or more other lexical items of the designational paradigm

3. No implantation: In the outcome state of language use, the target lexical item of language managers is not used at all by target speech community

78 I am here suggesting only three main categories, the second of which is fairly broad, because scholars have not agreed on the degree of implantation necessary for a deliberate lexical intervention to be called a success—see my comments below in the text body.

79 Other scholars proposed finer-grained classifications, e.g. Martin (1998, p. 68) distinguished between forms in conditions of competition ('formes en situation de concurrence terminologique') for target lexical items used less than 50% of the time, forms in the process of being implanted ('formes en voie d'implantation') for target lexical items used more than 50% of the time, and implanted forms ('formes implantées') for target lexical items used more than 80% of the time. It seems particularly arbitrary—and this is explicitly acknowledged by Martin—, however, to set such a precise threshold for partially implanted forms.

80 Standardization can be seen as a special case of total implantation (Moser, 1967, p. 27 states that terminological standardization is a special case of language management). Standardization aims in particular to *reduce* a designational paradigm that initially contains multiple lexical items to one that contains a unique item (Vila i Moreno, 2007b, p. 18). It therefore corresponds to an intervention of conscious lexical selection on the language system, combined with an intervention on the target speech community leading to a total implantation in language use.

What degree of implantation is required to declare the outcome a success is a question that, to the best of my knowledge, has not been answered. Some scholars proposed a minimum threshold (e.g., Martin) with a relative use of 80% (Martin, 1998, p. 68). Quirion (2003a) considered that implantation is not an absolute measure, but rather a series of values distributed over a continuum: He discusses the issue, but admits himself that he does not solve it (pp. 29–31).

Historically, reaching total implantation for a target lexical item was perhaps the first and sole linguistic goal of lexical interventions (see Quirion, 2014b, p. 103). A great amount of research on the topic was conducted within the discipline of terminology, and deliberate interventions on the lexicon find their roots in the development of the discipline of terminology. The former classical approaches of this discipline sought almost exclusively to reach complete standardization: Each lexical item was expected to correspond to only one concept (monosemy) and each concept to be named by only one lexical item (no synonymy; Desmet, 2007, p. 4). This is because the main aim of the practices and the theory of terminology in its initial years was to eliminate ambiguity from technical languages (Cabr e, 2003, p. 165), and standardization was the linguistic means to achieve this extralinguistic objective. There are contexts in which standardization is useful. A prototypical example is that given by W uster (1931) himself in his doctoral dissertation (p. 97), where four distinct objects (four types of wedges) are partly referred to by the same lexical items within a designational paradigm (below DP):

- DP(object1) = {Keil, Einlegekeil, Federkeil, Nutenkeil, Achskeil, versenkter Keil}*
DP(object2) = {Keil, Einlegekeil, Federkeil, Flachkeil, Feder, Einlegfeder}
DP(object3) = {Keil, Federkeil, Flachkeil, Feder, F hrungskeil}
DP(object4) = {Flachkeil, Fl chenkeil}

Some kind of linguistic consensus is needed here if unambiguous communication is to be ensured. **Lexical standardization**, the reduction of a designational paradigm to a single lexical item, is in specific settings still desirable today.⁸¹

81 In the technical translation industry, for instance, if not one but several lexical items are used in the source text to point towards the same concept, the following issues are expected to arise (Schmitz & Straub, 2010, p. 24): translators may make more follow-up inquiries about the source text, the match rate of the translation memory may be lower, and terminology-related translation errors may be more frequent. This has a cost both for the client and the translation provider: increased time is spent on the translation (time is money) and the resulting product might be of poorer quality, i.e. potentially contain mistakes. Depending on the context, translation errors can have

Because of the nature of language, total implantation for a target lexical item is however difficult to reach, is not necessarily desirable in every situation, and can probably only be achieved in very limited contexts through the use of drastic instruments (systematic terminology checkers, controlled languages, a totalitarian regime,⁸² etc.). Today, as lexical interventions are made in contexts other than standardization, lexical variation tends to be taken into account and accepted (Quirion, 2014b, p. 103). In specific settings, a partial implantation of a target lexical item can also be seen as a success by language managers. Quirion (2014b) noted, moreover, that governmental language agencies tend to be satisfied if indigenous lexical items are used (versus borrowings), independently of the items themselves (p. 103). Thus, the linguistic goal of language managers can be either total implantation or partial implantation (with degrees to be defined by language managers for their specific project).

2.5.2 The effectiveness of deliberate lexical interventions

Building on Heckhausen and Heckhausen's model of motivation and action,⁸³ Suchowolec defines the success of a language management measure as follows:

A language management measure is said to be successful when a collective and sustainable change in language usage is both the goal and the outcome.
(Suchowolec, 2018, p. 261, my translation)^{xxii}

I prefer to speak of *effectiveness* rather than *success*, because this is a term used in a large range of publications on project management.⁸⁴ Although relevant, Suchowolec's definition does not explicitly imply any kind of *relationship* between the goal and the outcome. Theoretically speaking, the goal and the outcome could be identical due to reasons *other* than a language management measure. This is a

far-reaching consequences, and not only of financial nature: they may even cost people their lives (Magris, 2009, p. 304).

- 82 As a matter of fact, even the Nazi Party in Germany did not systematically reach full implantation for their target lexical items (see Glunk, 1966).
- 83 Suchowolec refers to the German language edition, but see e.g. the last edition of the English version (Heckhausen & Heckhausen, 2018).
- 84 E.g. above I borrowed (from Eichhorn & Towers, 2018, p. 2) a general managerial definition of effectiveness.

crucial element, because the **effectiveness**⁸⁵ of deliberate interventions in language is still a little-examined topic (Schubert, 2014, p. 203). As D. Blanke pointed out:

The question also arises as to which degree the development of language takes place spontaneously and to which extent this development can be—deliberately or ‘artificially’—guided towards a target by human beings.
(D. Blanke, 1985, p. 23, my translation)^{xxiii}

Chansou believes that “it is not possible to isolate the effects of a policy intervention from other linguistic and extralinguistic factors influencing language change” (1993, p. 137, my translation; see also Moser, 1967, p. 24)^{xxiv} Let us take an example to better illustrate this point: In Quirion’s (2003a) study, for instance, for 38 concepts and corresponding designational paradigms in a given corpus,

- 21 target lexical items reached total implantation
- 11 target lexical items attained partial implantation, and
- 6 target lexical items obtained no implantation at all

Whereas the description of the outcomes is straightforward, the question remains as to why 21 target lexical items are exclusively used and six others are not used at all. As the lexicon varies over time even in the absence of an intervention by language managers, it is not clear to what extent the successes can be attributed to one or more deliberate lexical interventions (and thus how effective each deliberate lexical intervention actually is). Vila i Moreno and Nogué-Pich even go as far as stating that in some circumstances, an intervention in favor of a target lexical item may unfortunately even lead to a reduced implantation (2007, p. 37). In Chansou’s (2003) study, one of the interviewees confirms this assumption, expressing the opinion that any regulations or even recommendations will only reinforce the presence of Anglicisms⁸⁶ (p. 157).

85 ‘Effectiveness’ here is to be understood as the “accuracy and completeness with which [language managers] achieve certain goals”. It should not be confused with ‘efficiency’, which is “the relation between (1) the accuracy and completeness with which [language managers] achieve certain goals and (2) the resources expended in achieving them.” (Frøkjær, Hertzum, & Hornbæk, 2000, p. 345)

86 In this context, Anglicisms are the lexical items in competitions with the target lexical items.

A first difficulty in measuring deliberate lexical intervention effectiveness is to isolate the effects of deliberate lexical interventions from those of environmental factors. Thus, transposing Suchowolec's definition to deliberate lexical interventions, I find it necessary to mention some sort of causal relationship between the goal and the outcome, which I here express using the words "resulting from:"

Effectiveness of a deliberate lexical intervention

Partial or total implantation of the target lexical item in the target speech community resulting from a deliberate lexical intervention, whereas the desired degree of implantation is defined a priori by language managers.

I have now defined the goals of language managers and given a definition of effectiveness to measure how well these goals have been met. In the next chapter, I will state the problems language managers face in trying to reach these goals.

3 Intervening on speakers' lexicon under conditions of uncertainty

3.1 Introduction

In Chapter 2, I explained why language managers wish to deliberately intervene on the lexicon and what they are trying to achieve in the lexicon. I presented the route from aim to achievement and argue that speakers are the source of lexical change on which language managers should focus to enhance the efficiency of their deliberate lexical interventions. I propose considering lexical implantation as a three-step process comprising lexical knowledge, lexical opinion, and lexical replication.

I begin this chapter by stating the applied science problem language managers are facing (Section 3.2). I then review previous studies that have tried to evaluate and understand what was happening in target speech communities after a deliberate lexical intervention (Section 3.3). These highlight the research gaps, especially the fact that the data collected were mostly research-generated and gathered in an aided context. To address methodological shortcomings, I then propose (Section 3.4) to deepen research on speakers' lexical environments (Section 3.4.1) and to explore lexical opinions in unaided contexts using naturally occurring data (Section 3.4.2).

3.2 Problem statement

3.2.1 Language as a complex adaptive system

In Chapter 2, Section 2.4.3, I highlighted two major milestones language managers must pass in their interventions on the lexicon: reaching adequate levels of lexical knowledge for the target lexical items among speakers on the one hand, and ensuring a positive opinion regarding the target lexical item (approval) on the part of speakers on the other. Linguistic power, to a great extent, rests with speakers, and language managers might face substantial ignorance or resistance from the speech community. When selecting or coining lexical items with the goal of achieving (partial) implantation for their target lexical items, they commonly find themselves in the following situations:

- They must make a choice between mutually exclusive options when choosing a target lexical item⁸⁷
- They should make this decision quickly, because time is expected to be an essential factor for successful implantation of a target lexical item,⁸⁸ and
- They face **structural uncertainty**,⁸⁹ because they cannot rely on existing models to predict the speech community's response to their intervention

Language managers tend to coin or select lexical items based on a set of guiding principles (Diki-Kidiri, Joly, & Murcia, 1981; Gendron & Messina, 2015; ISO/TC 37, 2009; Schmitz & Straub, 2010, pp. 45–47; Suonuuti, 2001). However, “despite the use of implantation criteria grids in terminological work, nobody can predict what real usage(s) will be” (Rousseau, 2007, p. 70, my translation).^{xxxv}

Some scholars deplore the fact that “there is no hierarchy of principles to aid the resolution of dispute when different criteria are in conflict, and no one criterion is more important than the others” (Prys & Jones, 2007, p. 38), or that “there is still no definitive conclusion or model of the factors affecting [implantation]” (Bhreathnach, 2011, p. 174).⁹⁰ Depecker (2005) was still stating a little over a decade ago that we “urgently” need principles for coining and selecting lexical items (p. 21).

The problem language managers are facing here is a typical *applied* science problem. As stated in Chapter 2, they have a linguistic goal—reaching (partial) implantation—and their research seeks to uncover the most efficient and effective way to reach this goal. One must remember that “science serves two human purposes: to know and to do. The former is a matter of understanding, the latter a matter of action” (Feibleman, 1961, p. 305). One is called pure or basic science,

87 There are manifest contradictions between some of the common principles for naming (e.g. as presented in ISO/TC 37, 2009; Schmitz & Straub, 2010, pp. 45–47), for instance between the principle of transparency and that of linguistic economy. See also Quirion (2004, p. 198).

88 E.g. Quirion (2004, p. 198) lists just-in-time delivery (French: *livraison juste-à-temps*) as a positive factor for implantation.

89 ‘Uncertainty’ here is to be understood as “the character of situations in which agents [here: language managers] cannot anticipate the outcome of a decision and cannot assign probabilities to the outcome.” (Beckert, 1996, p. 804; on structural uncertainty see also Conroy, Runge, Nichols, Stodola, & Cooper, 2011, p. 1209).

90 In my view, these are overstatements, as e.g. Helfrich (1993) developed an implantation model in which factors are clearly weighted.

the other applied science.⁹¹ Applied sciences “produce new knowledge which is intended to be useful for the specific purpose of increasing the effectiveness of some human activity” (Niiniluoto, 1993, p. 5). This is precisely the aim here: enhancing the effectiveness of deliberate lexical interventions.

Up until now, scholars have approached the lexical intervention problem by working toward reducing structural uncertainty *ex ante*. A widely expressed belief among implantation scholars is that “if we know the variables that influence term usage, we can obtain the conditions for ensuring the implantation of normalized terminology” (Montané-March, 2012, p. 317, my translation)^{xxvi} Depecker (2013), for instance, advocated for an analytical-predictive approach:⁹²

[...] *a terminogram should help to situate the reasons of a language situation at a given moment in time (synchronous perspective) and to explain and foresee terminological language change over a long period of time (diachronous perspective).* (p. 24, my translation and my emphasis)^{xxvii}

Based on the latest studies of language change, however, it is doubtful that a *predictive* model could be at all possible for lexical change. Croft (2000) stated that he is “inclined toward the pessimistic view with respect to language change, which implies that even with perfect knowledge of the initial state, we would not be able to predict a language change” (p. 3). Language change is difficult to grasp because it involves various intertwined complex systems which can all change simultaneously.

To borrow van der Sluijs’ metaphor, trying to reduce uncertainty is a courageous (but often unfortunate) venture, because “for each head [one] chops off the uncertainty monster, several new monster heads tend to pop up” (2005, pp. 89–90). As a matter of fact, recent research on language change continues to tell us a

91 Temmerman regrets that “the scientific study of terminology is confounded with the pragmatic activity of standardisation” because most of the terminology schools have language planning as a motivation (2000, p. 19), but to me both approaches are scientific: descriptive studies belong to basic science, while deliberate lexical interventions (and corpus planning in general) are part of applied research.

92 Building on Gonthier’s concept of reference frame (1975, pp. 142–154), Depecker proposes to compile a map describing speakers’ individual reference frames. He calls this map a terminogram (2013, p. 24). A ‘terminogram’ would be the set of objective references (geographical, environmental, institutional...) and subjective references (cultural, value-based, ...) in which a speaker evolves and grows (see Depecker, 2013, pp. 24–25).

story that rather goes in the direction of an *increase* of uncertainty for language managers. The social networks of speakers, the individuals with whom they interact in private and professional life, can be expected to become more diverse and more complex: We have shifted from “being bound up in homogeneous ‘little boxes ...’ to networked societies, [in which] boundaries are more permeable, interactions are with diverse others, linkages switch between multiple networks, and hierarchies are both flatter and more complexly structured” (Wellman, 2002, p. 10). Furthermore, “the idea that language should be analyzed as a complex adaptive system is gaining currency, particularly among those researchers who are developing models of language behavior” (Blythe & Croft, 2009, p. 47; see also Ellis, 2011). Zarnikhi seemed to share this view with regard to terminology planning:

The system of language of science planning is involved in is, in fact, a socio-linguistic complex system taking on both social and language systems ... It is a complex adaptive system because of having many agents (dynamic forces) and networks and their complicated interaction and interconnection and, at the same time, it is not inflexible to changes. (Zarnikhi, 2014, p. 20)

Assuming that language is a complex adaptive system (CAS) carries enormous implications for interventions on the lexicon. One of these implications is to abandon the idea that lexical change may be largely predictable, to let go of the belief that knowing all the lexical change mechanisms as well as the initial conditions of a (socio)linguistic situation will allow language managers to predict change. *Broad features or patterns* can become knowable, but in a CAS it is probably impossible to predict every detail of an upcoming evolution (Levin, 2002, p. 17).

In a highly interactive system in which nonlinear relationships determine outcomes, system interactions are likely to be highly unpredictable (Bovaird, 2008, p. 326)

Even if initial conditions and generative mechanisms are exactly specified (which they cannot be), prediction of the future often becomes fruitless as specification errors grow exponentially as one progresses into the future (Choi, Dooley, & Rungtusanatham, 2001, p. 357)

Partial or total unpredictability is not necessarily synonymous with uncontrollability, but it requires approaches that do not seek to understand and

predict future processes in detail.⁹³ Considering lexical interventions as actions having a partially *uncertain* outcome is essential, because it paves the way for a reflection on viable strategies for coping with uncertainty. As Christensen suggests,

A crucial planning task is to discover, assess, and address uncertainty. Traditionally, planning has assumed that both means and ends are known. This professional legacy biases planners toward planning processes that address such conditions of certainty and disguise actual conditions of uncertainty. (Christensen, 1985, p. 63)

3.2.2 Introduction of a new lexical item and marketing

If one explicitly embraces uncertainty and acknowledges that lexical interventions do not necessarily lead to a predictable outcome in the speech community,⁹⁴ it becomes possible to start looking at the ways planners and managers in a large variety of other fields (ecosystem management, marketing, economics, health care, city planning, etc.) have dealt with uncertainty. Marketing, for instance, develops **strategies**, where strategy can be defined as “a statement of the action to be adopted **under a state of partial ignorance**, where all the alternatives cannot be recognised and stated in advance of the need for a decision” (Baker, 2014, p. 33, my emphasis).

I am using marketing as an example because a parallel has long been drawn between the introduction of a new product to a market and the introduction of a new lexical item in a speech community. Bourdieu (1977a, 1977b, 1982) in particular spoke of language using market terminology. Helfrich (1993) proposed using a marketing model to explain the introduction of new lexical items in language and compared the creation of a new lexical item with product innovation, the acceptance of the new lexical item with market success, and the consolidation of the lexicon with marketing productivity (pp. 41–45). As mentioned in the previous chapter, other scholars have seen an analogy between the dissemination of lexical items and Roger’s theory of the diffusion of innovations (Martin, 1992, 1994; Quirion, 2014, p. 110).

93 Such as Depecker’s approach that aims to foresee terminological change, which I just quoted above (Depecker, 2013, p. 24).

94 Or that the outcome will always remain at least partly unpredictable.

Quirion (2006) suggested that intervening on the lexicon equates to selling products (lexical items) to clients (speakers) in given markets (domains; p. 826). He mentioned the fact that numerous marketing studies have established principles for market entry in terms of purchasing behavior, purchasing decisions, decision processes, and types of consumers (Quirion, 2006, p. 827). Quirion (2006) also stated that, in most deliberate lexical intervention cases, the lexical innovation does not correspond to speakers' needs, but rather, language managers make speakers aware of their target lexical item (p. 828). In marketing terms, this means that one is mainly dealing with a *supply-side* market.⁹⁵

Predictability and full comprehension of the situation *ex ante* are not the key to success in supply-side markets: In fact, expert entrepreneurs are skeptical of market research (Read, Dew, Sarasvathy, Song, & Wiltbank, 2009). In a supply-side market, the "marketing strategy is highly entrepreneurial—formulated on sketchy market information and on intuition" (Shanklin & Ryans, 1984). The successful entrepreneur is an individual who responds creatively to *unforeseen* changes:

Entrepreneurship is a process that involves some degree of uncertainty, and thus the ability of entrepreneurs to interpret and respond to uncertainty is often what determines the degree of success or failure achieved by the venture. (McKelvie, Haynie, & Gustavsson, 2011, p. 273)

Unlike demand-side marketing, which calls for traditional target market analyses, supply-side marketing requires the need to import knowledge (Couillard, 2006, p. 27). Companies pay large amounts of money to "listen" to their customers (Li & Bernoff, 2011, p. 80), and successful entrepreneurs adapt as they learn from the marketplace. They do not apply causal logic. Dreyfus and Dreyfus (1986) once suggested that experts act *arationally*. They do not use people's feedback to *predict ex ante* a future situation. Rather, they use this feedback to *improve what they offer* on the market or to adapt their strategies. Experienced entrepreneurs are fast on the market, gather fast feedback at low cost, and work in an iterative manner.

Regarding lexical interventions, Auger stated the following:

*It is important that, **throughout the whole process**, an outcome evaluation is conducted. Are the chosen lexical items well received by speakers? What is the feeling of individuals targeted by the change? Does the terminology*

95 As opposed to a *demand-side* market.

reference meet the expectations of the future users? (Auger, 1986, p. 52, my translation and my emphasis)^{xxviii}

However, language managers often wait until much time has elapsed before they start evaluating how well their target lexical items fare in the speech community, following the assumption that “since time is a crucial factor for the implantation of terminologies, the evaluation can take place only long after these have been disseminated” (Quirion, 2003, p. 7, my translation).^{xxix} In Vila i Moreno and Vila i Moreno’s (2007) study on fencing, for instance, the evaluation was undertaken about two years after the dictionary containing the target lexical item was prepared.

This is in clear contradiction with the theory of the diffusion of innovations, in which the primary goal is to reach a minimal critical mass of individuals who adopt the innovation. I believe language managers should concentrate their efforts especially on the *beginning* of the implantation process, on reaching the minimum amount of speakers for their target lexical item; then, the reaction could be expected to be self-sustaining:

Reaching the critical mass is an important point to consider during the implementation phase. After reaching a critical mass of adopters, any innovation, like the introduction of new terms ... becomes self-sustaining. Earlier and later users re-inforce themselves mutually in their decision to continue, abandon or take up this particular innovation. (Drame, 2009, pp. 126–127)

Additionally, theories of diffusion suggest that there are distinct categories of adopters who do not adopt the innovation at the same point in time (Rogers, 1983), and individuals might choose to adopt an innovation for different reasons.⁹⁶ This implies that there cannot be only a single lexical implantation framework that fits every member of the target speech community.⁹⁷ What is particularly striking in previous lexical implantation factor research is the nondynamic nature of the proposed frameworks and models. Helfrich (1993), for instance, designed a one-size-fits-all model for the acceptance of new lexical material (pp. 228–291). However, she had clearly shown in her empirical results that the evaluation of

96 This has been illustrated by empirical research. In Eastin’s study on the adoption of e-commerce activities (2002), for instance, adopters adopt the activities for different reasons.

97 I would be curious to see how Costa’s ‘favourable sociolinguistic conditions framework’ (2016) could be implemented in practice.

lexical material was group specific (Helfrich, 1993, pp. 146–170). The dynamics of implantation factors imply that language managers’ “marketing tactics” for their target lexical item should change as they move along the adoption curve:

Strategies must change to leverage the specific requirements and behaviors of different groups along the diffusion curve. Product offerings may have to be adjusted over time and different adopter groups have to be told different stories about the benefits of the innovation. (Waarts, van Everdingen, & van Hillegersberg, 2002, p. 413)

It thus seems essential that language managers be aware of what is currently happening in the speech community to react and, if needed, adapt their lexical “products” and/or their dissemination strategies.

From a marketing perspective, the use of market information is considered a key success factor for new products:

Although no single variable holds the key to new product performance, many of the widely recognized success factors share a common thread: the processing of market information. (Ottum & Moore, 1997, p. 258)

3.2.3 Deliberate lexical intervention as a cocreation process

To gain information from the marketplace for new product development, marketers have applied approaches ranging from reactivity (monitoring using traditional research techniques) to proactivity (cocreating directly with customers). In particular, the proactivist approach of cocreation “is becoming increasingly popular among companies, and intensive communication with customers is generally seen as a determinant of the success of a new service or product” (Witell, Kristensson, Gustafsson, & Löfgren, 2009). Cocreation existed long before the 21st century, but with Web 2.0 technologies, it has moved to the forefront (Cova, Dalli, & Zwick, 2011, p. 2). Thanks to the advances of technology, customers are now contributing in a way that would not have been possible before (Humphreys & Grayson, 2008). According to Zwick et al. (2008), companies have refashioned their strategies, replacing top-down approaches with a “government” approach, in which customers’ actions are not shaped by marketers’ orders, rules, or norms but rather are taking place in dynamic platforms of practice controlled by the marketers (p. 165).

What marketers call cocreation has long been practiced by some language managers: Speakers' inputs have been used, for instance, in collaborative approaches in private companies (Arndt & Wöllbrink, 2013), in the Hebrew revival (Nahir, 1998, 2002), in TNC's⁹⁸ collaboration schemes with domain experts (Nissilä & Pilke, 2017), or in the testing of new terms within group discussions (Williams-van Klinken, 2008). Quirion (2012) also clearly highlighted the potential of collaborative approaches on a large scale. The power of cocreation approaches can be enormous:

From a managerial standpoint, it is quite interesting to note that co-creating the voice of the customer is likely to result in deeper bonds with customers—more trust, more commitment, and more loyalty—for two reasons. First, because a customer is involved in the process, the customer builds commitment to the resultant offering by the firm. Second, because of the offering is co-developed [sic], it has a higher probability of accurately meeting the customer needs. Finally, co-creating the voice of the customer provides a firm asymmetric information about the marketplace. (Jaworski & Kohli, 2006, p. 116)

However, cocreation comes with a host of challenges:

*Indeed, the challenge in building a more service oriented and customer centred business model relates to the **type of relationships and interactions** to be utilised in co-creating value. Hence, organisations need to decide on **how they can best involve co-creators** and choose appropriate approaches and tools. (Roser, Defillippi, & Samson, 2012, my emphasis)*

Although cocreation is not a new research area, marketers have just begun to scientifically discuss how to best “put customers to work” (Lusch & Vargo, 2006; Tran, 2017).

***Key strategies for co-creation**, which are needed in order to facilitate proactive market-orientated NPD [new product development], **are inconclusive**. While, the outcome of close collaboration with the customers is well documented, understanding what to do in order for these benefits to be realized is more obscure. (Kristensson, Matthing, & Johansson, 2008, p. 477, my emphasis)*

98 TNC was the Swedish national center for terminology.

Theoretical considerations about cocreation have thus just started to emerge (Shamim & Ghazali, 2014).

For lexical interventions, it has become clear that target speakers must be involved to some extent:

It is no longer enough simply to accumulate and to act within official organisations confining speakers to a passive, spectator role, outside the process of planned modernisation of the language. (Samuel, 2005, p. 516, as translated in Bhreathnach, 2011, Vol. I, p. 12)

However, in our current state of knowledge, I believe that cocreating the lexicon with speakers raises more questions than it provides answers: What roles should language managers give speakers? How many of them are needed? How can speakers' participation and engagement be ensured? How should speakers' opinions be weighted?⁹⁹ The Web provides us with examples where I think language managers have failed in introducing cocreation with speakers through a dynamic platform. Let us take one for illustrative purposes: the online Wiki platform for the French language "wikiLF" (FranceTerme, n. d.). On this platform, individuals can suggest a new French lexical item to replace an Anglicism or vote on previous proposals. The idea of involving ordinary speakers sounds brilliant at first, but as I write these lines, the last suggestion displayed by the platform dates back to October 2015. This suggests that the platform is not particularly dynamic.¹⁰⁰

I am deeply convinced that language managers should try, as much as possible, to integrate speakers into their lexical interventions. However, Li and Bernoff (2011) suggested that integrating customers into one's business is the *most* challenging goal a company can have and encouraged managers to succeed with another goal first, such as listening to the customers, talking to the customers, energizing the customers, or supporting the customers (pp. 68–69).

Here, I propose to start by listening because this strategy is "best suited to understand [the] customers" (Li & Bernoff, 2011, p. 68). This monitoring of speakers' voice (ongoing evaluation, as suggested by Auger, 1986) could be used not so much (or not only) for increasing the overall comprehension of implantation

.....
99 Private companies are discussing these questions (e.g. Haag & Brodersen, 2013; Herwartz, 2007; Kurfess & Schmach, 2012), but from a pragmatical perspective, not a scientific one.

100 I have no information as to why this platform is no longer active, but I assume something must have failed in the process of involving enough participants.

factors but rather to recursively make data-based decisions at each step of the process, as is the case in marketing strategies:

*The fact that consumers and their environments are **constantly changing** highlights the importance of **ongoing consumer research** and analysis by marketers to **keep abreast of important trends**. (Peter & Olson, 2010, p. 6, my emphasis)*

Monitoring “provides the feedback loop for learning about the system; learning is sought not for its own sake but primarily to better achieve management objectives” (Lyons, Runge, Laskowski, & Kendall, 2008, p. 1683). Monitoring information directly from the “marketplace”, from the speech community, would open the door to new strategic possibilities for language managers, such as (without any claim of completeness) the following:

- quickly adapting the measures for making the target lexical item known in case of null levels of lexical knowledge
- abandoning the target lexical item (and proposing another one) if the immediate resistance of the target speech community is substantial
- if not, empowering early adopters and opinion leaders
- working toward changing the negative attitudes of “resisters”
- in the longer term, understanding what kind of lexical tasks speakers “naturally” accomplish (commenting on the lexicon, searching for alternatives, etc.) to start reflecting on if/how their work could be enhanced in a dynamic platform (cocreation)

In the next section (3.3), I will introduce indicators that have previously been used for monitoring the speech community during or after a deliberate lexical intervention. Underlining their shortcomings, I will argue for new approaches, which I will present in Section 3.4.

3.3 Evaluating deliberate lexical intervention effectiveness in the speech community

In the previous section, I mentioned that deliberate lexical interventions should be submitted to an evaluation throughout the whole process and that, in marketing terms, it is necessary to gain information from the marketplace. To this end,

language managers need appropriate tools for evaluation. In the present section, I review previous studies in which researchers tried to assess or understand the lexical implantation process, giving an overview of method. I classify these studies into three sections, which correspond to the three-step implantation process proposed in the previous chapter (2.4.3):

0. **Lexical ignorance:** The lexical item does not exist for the speaker.
1. **Lexical knowledge:** The speaker is aware of the lexical item.
2. **Lexical opinion:** The speaker has an opinion about the lexical item.
3. **Lexical replication:** The speaker uses the lexical item.

Two cases are possible: (a) The lexical item might not exist at all at the level of the speech community or (b) the lexical item exists within the speech community but the speaker ignores it.

3.3.1 Previous studies on lexical knowledge

What do I call lexical knowledge? Evidently, lexical knowledge is not cut and dry. It is not as simple as the absence or presence of knowledge. Various disciplines distinguish several levels of knowledge,¹⁰¹ and so does applied linguistics. Although there is “no clear and unequivocal consensus ... as to the nature of lexical knowledge” (Laufer & Paribakht, 1998, p. 366), it seems to be largely agreed upon that lexical knowledge is a *continuum*, from the mere awareness of the existence of the lexical item (vague familiarity) to the full ability to access the lexical item for free, active production. Some scholars have suggested that lexical knowledge is a multidimensional construct (Henriksen, 1999). Generally, a discussion is needed as to which type of lexical knowledge is best suited as a metric for monitoring the progress of the target lexical item.¹⁰²

Although lexical knowledge is known to be a gradual concept, previous studies often remained unidimensional.¹⁰³ This cannot be due to the lack of information

101 Rogers discusses this issue in his theory of the diffusion of innovations: he distinguishes between several types of knowledge (1983, p. 167–168): notably awareness-knowledge (what is the innovation) and how-to-knowledge (how does the innovation work). Several levels of knowledge are also discerned in marketing studies (e.g. Peter & Olson, 2010, p. 68–71).

102 But this is out of scope here.

103 On a parenthetical note, I find it surprising that none of the studies reviewed considered for instance the concept of *lexical availability* (see e.g. Morales, 2014), which

or guidelines for developing some kind of scale because scale development was a “growth industry” within the field of psychology in the early 1990s already and “it ha[d] become axiomatic that (publishable) assessment instruments are supposed to be reliable and valid” (Clark & Watson, 1995, p. 304). Also, the construct of “lexical knowledge” has been largely discussed in applied linguistics, especially in studies concerning second language learning. Henriksen (1999), for instance, underlined that there is a need to be more specific about the construct and proposed three separate dimensions: partial-precise knowledge, depth of knowledge, and receptive-productive knowledge (p. 304).

In the present investigation, I propose speaking in terms of four constructs based on two main lexical knowledge dimensions, at two levels. Here, I call the two dimensions lexical recall and lexical recognition. **Lexical recall** is a top-of-mind awareness, whereas **lexical recognition** is an aided awareness. My concepts and terminology are inspired and adapted from the marketing literature:

For example, if people think of a soft drink, they may spontaneously think of either Coca-Cola, Fanta or Lipton Ice Tea. This is their top-of-mind brand awareness. In an unaided context people may recall several brands spontaneously. This is brand recall or unaided spontaneous awareness. It is also possible that people recognise a brand by its package, colour, logo, etc. This is brand recognition or aided awareness. (De Pelsmacker, Geuens, & Van den Bergh, 2005, p. 76)

My two dimensions of lexical recall and lexical recognition can be applied to either the level of lexical items or that of **lexical sources** (i.e., sources containing lexical items such as official documents, dictionaries, online terminology databases, etc.). In addition, the conditions in which a speaker is prompted to recognize a lexical item or a lexical source (in the “aided context”) can be subdivided into two forms: active recognition or passive recognition. Active or spontaneous recognition occurs when the lexical item or source is not given to the speaker: This is the case, for instance, when a researcher gives a definition of a lexical item to the speaker and invites the speaker to indicate the corresponding lexical item. Passive recognition occurs when the lexical item is directly given to the speaker: For example, a researcher gives a lexical item to a speaker and notes whether this speaker

emerged in France more than half a century ago and has developed a quantitative paradigm.

recognizes the lexical item. For this dichotomy between active and passive, I am borrowing ideas from existing lexical implantation research:

*We consider that a denomination [lexical item] is known **spontaneously** if it is produced by an informant within the context of the study (Vila i Moreno & Vila i Moreno, 2007, p. 77, my translation and emphasis)^{xxx}*

Accordingly, I use the six concepts below.

Lexical item recall: the ability of a speaker to think of a lexical item in an unaided context

Active lexical item recognition:¹⁰⁴ the ability of a speaker to think of a lexical item in an aided context, where the lexical item is not directly provided

Passive lexical item recognition: the ability of a speaker to recognize a lexical item in an aided context, where the lexical item is directly provided

Lexical source recall: the ability of a speaker to think of a lexical source in an unaided context

Active lexical source recognition: the ability of a speaker to think of a lexical source in an aided context, where the lexical source is not directly provided

Passive lexical source recognition: the ability of a speaker to recognize a lexical source in an aided context, where the lexical source is directly provided

I used **lexical knowledge** as an umbrella term covering these six concepts. Needless to say this terminology differs from that used in the studies I am reviewing, but I am using these six harmonized concepts with a view to structuring my review.¹⁰⁵

Several of these studies on lexical knowledge have confirmed low or even missing lexical knowledge for the target lexical item among target speech community members, as illustrated here in Quebec, the Gaeltacht, and Catalonia:

.....
104 There have been proposals and endeavors for assessing lexical item recognition in the Esperanto speech community (see Kück, 2009).

105 I invite interested readers to refer directly to the studies quoted which, for the most part, are not written in English (thus the terminology differs in any case, since these works are written in other languages).

In general, as already mentioned, official lexical items [target lexical items] are unevenly and little known by editors. The majority of respondents (50% and more) can recognize only 20% of officialized forms. (Martin, 1998, p. 189, my translation)^{xxxii}

The Gaeltacht community may not be aware of certain modern terms ... The data suggest that Irish terms [the target lexical items] may not be used by Gaeltacht speakers as they simply do not know them. (Ní Ghearáin, 2011, pp. 312–313)

The diffusion of the Calatan denominations proposed by TERMCAT [the target lexical item] has been zero since none of the interviewees knew the proposed forms nor the terminological resources available for consultation. (Gresa Barbero, 2016, p. 68, my translation)^{xxxiii}

Scholars often measure more than one concept in a single study, as well as concepts other than lexical knowledge. Below, I review studies about lexical knowledge in chronological order. This linear presentation is followed by a summary table for comparison purposes and a discussion of existing methods (strengths and weaknesses).

In her study of Hebrew terms for car parts, Allony-Fainberg (1974) used a questionnaire distributed to four population samples. She asked respondents whether they knew specific target lexical items (passive lexical item recognition).

Heller (1978) conducted a study with the aim of bringing to light the multiple factors that can influence the implantation of a set of target lexical items in the automotive domain. Regarding lexical knowledge, she showed her informants car pictures and asked them to indicate the lexical items they knew for each car part (active lexical item recognition).¹⁰⁶ Although she did not offer systematic results for lexical knowledge, she mentioned that “many interviews began with hesitation from the part of the informant, who stated that they ignored the French terms, the ‘correct terms’” (Allony-Fainberg, 1978, p. 35, my translation).^{xxxiii}

Fugger (1980, see also 1983) was interested in the reactions of speakers to the French language policy and the influence of English on the French language. He explored two aspects of lexical knowledge through a survey: passive lexical item

106 The questions were asked partly through semi-directed interviews and partly through written questionnaires. Informants were also asked to give information about the use of these lexical items, but this is out of scope in this section.

recognition, by asking respondents to provide a definition for specific lexical items (both Anglicisms and frenchized lexical items), and passive lexical source recognition, by asking respondents to indicate whether they had heard of the ministerial orders published in 1977 for replacing words of English origin with French words.

Gaudin and Guespin (1993) carried out semistructured interviews with university lecturers/researchers in the domain of genetic engineering in France. With regard to lexical knowledge, their research technique consisted of creating an informal discussion during the interview and guiding it toward specific topics to assess whether interviewees would spontaneously utter the target lexical item. For each designational paradigm considered, they listed cases in which the interviewee spontaneously produced the target lexical item (lexical item recall) versus cases in which he or she used a rival lexical item. In addition, for a restricted number of target lexical items, they gave the target lexical item to the interviewees and asked them to give either a definition of this lexical item or an equivalent (passive lexical item recognition).¹⁰⁷

Gouadec, Crespel, and Colombel (1993) carried out two surveys with individuals involved in IT in France. Regarding lexical knowledge, their goal was to determine to what degree these persons knew the target lexical item. Two approaches were used: a question of the type “Can you define or describe [the target lexical item]?” (passive lexical item recognition) and a question of the type “How do you point toward or express [definition of the target lexical item]?” (active lexical item recognition).

Concerning lexical knowledge, the study undertaken by Thoiron et al. (1993) constitutes a significant step forward. Not only did the research team highlight the fact that lexical knowledge is a continuum, but they further devised a strategy to try to measure multiple degrees of knowledge:

As far as the term knowledge is concerned, several levels must be considered. It is simplistic to say that a term [target lexical item] is known or unknown, even by a specialist. There are degrees, and here we tried to implement various strategies to assess the level of familiarity with the term. (Thoiron et al., 1993, p. 50, my translation)^{xxxiv}

107 In this case, the equivalent units the interviewees were prompted to utter were Anglicisms.

In their semistructured interviews, they first gave the interviewees a definition. If the target lexical item was uttered by the interviewee (active lexical item recognition), 10 points were given. If not, the test continued.¹⁰⁸ If a lexical item other than the target lexical item¹⁰⁹ was given, the interviewers tested whether the interviewee could give the target lexical item (6 points); if not, the interviewers mentioned the target lexical item to see if the interviewee recognized it (4 points; passive lexical item recognition). If the interviewee did not mention the target lexical item spontaneously, could not utter it when stimulated with a competing lexical item of the designational paradigm, or did not recognize it when it was presented to them, they did not receive any points. As the authors mentioned (Thoiron et al., 1993, p. 54), multiple variables were involved, and their measure could only serve as an indicator to quickly evaluate, in an initial stage, whether a target lexical item (or a group of such units) is propagating.

Guilford (1997) conducted two surveys to examine the attitudes of French people toward English loanwords. In the first survey, he explored passive lexical item recognition in two ways: by asking survey respondents whether they knew the lexical item and, if so, by asking them whether they could define the lexical item. The results are interesting from a methodological point of view because they show that in some cases, respondents claimed to know the lexical item but were not able to give a correct definition of the underlying concept. This again brings up the question of levels of lexical knowledge.

Martin (1998) conducted a survey as well as group interviews with editors in the domain of college education and administration in Quebec. He announced that through his survey, he collected inter alia data on *knowledge* and *use* of target lexical items (Martin, 1998, p. 161). Out of four questions, only the last addresses knowledge. Here, however, knowledge is not understood as knowledge of the target lexical item per se. Rather, the survey prompts respondents to indicate, within 27 designational paradigms given to them, which lexical items *are* target lexical items. It is questionable whether the information collected would be generally useful because speakers of the target speech community in some cases could know and use target lexical items without knowing explicitly that these lexical items are target lexical items. The most interesting aspect of Martin's (1998) study in relation to lexical knowledge is, in my view, his group interviews, which explored in depth not (only) the knowledge of the target lexical item but, in a broader sense, knowledge of the language managers' sources containing the target lexical item.

108 For a detailed account of the methods, see the authors (Thoiron et al., 1993, p. 53).

109 Here: an English equivalent.

Martin (1998) used an interview guide, which, as far as I could read, is not provided in his work. Therefore, I am not sure how he interviewed the groups, but the results show that data were generated on the knowledge of lexical sources. I classify his work as passive lexical source recognition.

In their study on fencing, Vila i Moreno and Vila i Moreno (2007) combined participant observation and interviews. During participant observation, they collected information on lexical item recall: Vila i Moreno fenced in national Catalan competitions, sometimes with recordings. As mentioned in the study, this method is an interesting approach to avoiding the observer's paradox. During interviews, they also tested comprehension of given lexical items, which I classify as passive lexical item recognition (the interviewer suggesting the target lexical item to the interviewee¹¹⁰). The researchers were also interested in the knowledge of the dictionary managers had been using to disseminate their target lexical item (passive lexical source knowledge). An interesting aspect (which is disregarded in my classification)¹¹¹ is the fact that knowledge of the existence of this source on the one hand and the possession of (and thus access to) this source are distinguished.

Nogué Pich and Vila i Moreno (2007a) conducted a study in the field of sport climbing in Catalonia. They questioned climbers about a set of designational paradigms. They used pictures instead of definitions or descriptions. The researchers showed illustrations to the interviewees and assessed the lexical item used spontaneously ("How do you name this?"; active lexical item recognition), other lexical items known by the interviewee ("Do you know other names for this concept?"), and, if the target lexical item had not been mentioned, whether they knew the target lexical item ("Do you know this denomination [the target lexical item]?"^{xxxv}) (passive lexical item recognition). Their multi-question approach is interesting because it captures intraspeaker variation,¹¹² whereas previous studies focused exclusively on the target lexical item. It seems relevant to learn more about the designational paradigm at the level of a specific speaker: If it contains only the target lexical item, it is probable that the speaker will use it. If it contains various other lexical items, the target lexical item will not necessarily be the preferred item.

.....
110 With three possibilities: comprehension of the target lexical item, incomprehension of the target lexical item and ignorance of the concept.

111 I think this distinction is bound to disappear in the future, as speakers seem to rely more and more on lexical sources that are available freely on the web rather than paper dictionaries. To facilitate legibility, I thus prefer not to include this aspect in my review.

112 E.g. Heller (1978, p. 34) notes that intraspeaker variation is essential to understand which direction lexical change is taking.

However, the reliability of the “spontaneous” use of the target lexical item is quite relative.

Nogué Pich and Vila i Moreno (2007b) also undertook a study in the field of hockey, which comprised interviews, a survey, and study of oral and written corpora. During the interview, they started with an informal discussion to see whether respondents would spontaneously mention the target lexical item (lexical item recall); if the interviewees did not mention the term, they asked more explicit questions related to the target lexical item (passive lexical item recognition). They further assessed the knowledge of the sources that had been used by language planners for disseminating the target lexical item (passive lexical source knowledge): “Do you know the dictionary of hockey?” “Have you ever seen a poster with this (teach poster)?” “Have you already seen a brochure with this (teach brochure)?” or “Do you know what TERMCAT is?”^{xxxvi}

Ní Ghearáin (2011) conducted semistructured interviews in Irish-speaking workplaces, mostly community-based organizations such as cooperatives. Her study design allowed her to collect qualitative information about lexical knowledge among respondents. Without going into detail, she mentioned that “findings do in fact suggest that the Gaeltacht community [the target speech community] may not be aware of certain modern terms [target lexical items]” and further that her data “suggest that Irish terms may not be used by Gaeltacht speakers as they simply do not know them” (Ní Ghearáin, 2011, pp. 312–313). She mentioned, for instance, that half of her interviewees had never heard the target lexical item (an Irish term) for “laptop” (passive lexical item recognition).¹¹³ Ní Ghearáin (2011) also collected data about knowledge regarding language managers (i.e., the official terminology planning structure concerned). In the paper, the methodology is not reported in full detail, but Ní Ghearáin (2011) mentioned that she investigated “the informants’ practices, use of terminology resources, and awareness of and attitudes toward the institutional structure for terminology development” (p. 311). Thus, I assume she explored both passive and active lexical source recognition.

Gresa Barbero (2016) conducted interviews with owners or individuals in charge of restaurants with regard to the dissemination methods of the Catalan center for terminology (TERMCAT). Her purpose was to collect data on

113 I’m assuming this is lexical recognition here, but I couldn’t grasp the full methodology from the paper, which regarding the interviews succinctly mentions “Interviews consisted of three stages, the first of which entailed detailed discussion of the informants’ everyday usage of technology, thereby facilitating the observation of their knowledge and usage of technological terms.” (Ní Ghearáin, 2011)

interviewees' perceptions of the dissemination process of target lexical items. Thus, she did not focus on lexical knowledge but qualitative data about passive lexical source recognition appear in her results. Gresa Barbero (2016) mentioned that none of the interviewees were aware of TERMCAT's dissemination attempts in the domain considered, such as the lexical resources made available on the Internet (p. 40).¹¹⁴

In Table 2, I summarize the methods I just presented for comparison purposes:

Aspect	Study	Method	Tool
Lexical item recall	(Gaudin & Guespin, 1993)	Informal discussion	Interview
	(Vila i Moreno & Vila i Moreno, 2007)	Participate in speakers' activities (fencing)	Participant observation
	(Nogué Pich & Vila i Moreno, 2007b)	Informal discussion	Interview
Active lexical item recognition	(Heller, 1978)	Show interviewees pictures and ask for the lexical item	Interview
	(Gouadec et al., 1993)	Give interviewees a definition and ask for the lexical item	Interview
	(Thoiron et al., 1993)	Give interviewees a definition and ask for the lexical item	Interview
	(Vila i Moreno & Vila i Moreno, 2007)	Give interviewees a definition/description and ask for the lexical item	Interview
	(Nogué Pich & Vila i Moreno, 2007a)	Show interviewees pictures and ask for the lexical item	Interview
Passive lexical item recognition	(Allony-Fainberg, 1974)	Ask respondents whether they know the lexical item	Survey
	(Fugger, 1980)	Give the lexical item and ask for a definition	Survey

.....
114 She nonetheless notes that some forms identical with the target lexical item are implanted, but suggests that this is not the results of TERMCAT's dissemination activities (2016, p. 40).

Aspect	Study	Method	Tool
	(Gaudin & Guespin, 1993)	Give the lexical item and ask for a definition or equivalent	Interview
	(Gouadec et al., 1993)	Ask to define or describe the lexical item	Interview
	(Thoiron et al., 1993)	Give interviewees the lexical item and see if they recognize it	Interview
	(Vila i Moreno & Vila i Moreno, 2007)	Give interviewees the lexical item and see if they recognize	Interview
	(Nogué Pich & Vila i Moreno, 2007a)	Ask respondents whether they know the lexical item	Interview
	(Nogué Pich & Vila i Moreno, 2007b)	Ask concrete questions about the lexical item	Interview
	(Ní Ghearáin, 2011)	<i>Details unknown to me</i>	Interview
Lexical-source recall	-	-	-
Passive lexical source recognition	(Fugger, 1980)	Ask respondents whether they know lexical sources of language managers (ministerial orders)	Survey
	(Martin, 1998)	<i>Unclear to me</i> ¹¹⁵	Group interview
	(Vila i Moreno & Vila i Moreno, 2007)	<i>Details unknown to me</i>	Interview
	(Nogué Pich & Vila i Moreno, 2007b)	Ask respondents whether they know specific lexical sources (a dictionary, etc.)	Interview
	(Ní Ghearáin, 2011)	<i>Details unknown to me</i>	Interview
	(Gresa Barbero, 2016)	<i>Not the aim of the study but data emerged</i>	Interview

115 This is an assumption. I did not find Martin's interview guide in his work, therefore I was not able to tell clearly how he interviewed the groups.

Aspect	Study	Method	Tool
Active lexical source	(Ní Ghearáin, 2011)	<i>Details unknown to me</i>	Interview

Table 2. Summary of methods used to investigate six aspects of lexical knowledge.

As it appears in the summary table, a large majority of studies were undertaken in the presence of researchers and in an aided context.¹¹⁶ These methods are based on research-generated data and suffer from the observer’s paradox.¹¹⁷ Several implantation researchers underlined the drawbacks of elicitation in the presence of the researcher (Heller, 1978, pp. 3–4; Ní Ghearáin, 2011, p. 111). Vila i Moreno (2007) himself acknowledged that the spontaneity of the lexical item knowledge he was assessing was influenced by the interview setting:

The interviewees can perfectly know various alternative [lexical items] for a single concept, and produce the one that seems to be the most normative—or, inversely, the one that seems to be the least so—depending on the dynamics of the interview. (p. 78)

Heller (1978) gave a straightforward example of the researcher effect at the level of lexical items, where an interviewee uses different lexical items in real life and in the interview setting:

For instance, when I opened a garage door I heard someone say “Va chercher les tires!” [Go get the tires!]. Later on I interviewed this person (it was the owner), and he told me that he only used pneu. (p. 38, my translation)^{118, xxxvii}

The only interesting exception in the studies I reviewed is that of Vila i Moreno and Vila i Moreno (2007), who worked with participant observation. The presence

116 This is acknowledged by some researchers. Heller, for instance, states that informants were hesitant (see above), suggesting an observational bias.

117 See Cukor-Avila (2000) and Labov (1972, p. 109) for a discussion of the observer’s paradox.

118 ‘Tire’ is the Anglicism for the French lexical item ‘pneu’.

of researchers to explore the level of lexical items therefore constitutes an initial problem.

Another debatable issue in the works reviewed above relates to the degree or scale of knowledge. Gouadec et al. (1993), for instance, announced that in the third part of their study, they sought an “analysis of the **degrees** of knowledge of official terms [target lexical items] and of their meanings” (p. 239, my translation and emphasis).^{xxxviii} As psychology teaches us, “[a] primary goal of scale development is to create a valid measure of an underlying construct” (Clark & Watson, 1995, p. 309). The study by Gaudin and Guespin (1993) constitutes an improvement because two levels were considered: an *active* level, where study participants might produce the target lexical item spontaneously, and a *passive* one, where they merely recognize the target lexical item. However, although the researchers claimed that in the third part of their study, they aimed to “assess the **degree** of knowledge of recommended terms [target lexical items]” (Gaudin & Guespin, 1993, p. 12, my translation and emphasis),^{xxxix} it is deceiving because no scale was proposed or even discussed in the paper, and no figures were provided. Also, studies on lexical item knowledge are sometimes based on self-reported knowledge. The problem with self-reported measures of knowledge is that “they may reflect generalized self-confidence more than any actual state of knowledge because people who are self confident may report more knowledge than those with less confidence” (Cole, Gaeth, & Singh, 1986). Some researchers, such as Allony-Fainberg (1974), acknowledged this issue but did not discuss a solution. The scale problem is thus a second problem with lexical item knowledge.

From the above, I see two problems with exploring lexical knowledge at the level of lexical items: elicitation in the presence of researchers and the question of scales of knowledge. However, in addition to these methodological problems is the fact that the studies reviewed generally resulted in unsatisfactory knowledge of lexical items: As mentioned at the beginning of the present section, figures can be as low as 20% or even zero. This is why I propose starting at a higher level, that of lexical sources, to try and understand why speakers have not come into contact with the target lexical items. It might sound like a statement of the obvious that if the source containing the target lexical items is unknown to speakers, the probability that speakers will use the target lexical items is approximately zero. As is evidenced by Table 2 (p. 94), the reviewed studies either did not focus on lexical source knowledge or took target lexical item sources as a starting point. This fails to capture the bigger picture: If target lexical item sources are not known, from which source do speakers draw their lexical items? In Section 3.4.1, I will argue

for an exploration of speakers' lexical environment, which does not use target lexical item sources as a starting point.

3.3.2 Previous studies on lexical opinion

The second step in the three-step implantation process depicted in Section 3.3 (p. 83) is for a speaker to make the decision to adopt or reject a lexical item. However, at a higher level, speakers can also decide to adopt or reject a lexical source, for instance, choosing not to use a specific dictionary they know. Rejecting a lexical source *de facto* means limiting access to the lexical items it contains. Accordingly, two aspects seem important to me here: speakers' opinions about lexical items proper and speakers' opinions about lexical sources.¹¹⁹

Here, I use opinion or **lexical opinion** as an umbrella term to aggregate a large range of concepts: attitudes, sentiments, feelings, emotions, preferences, and so on. I feel the need for a generic term and concept because researchers have approached the issue under different perspectives, which I am presenting in this section.

The concepts linguists have used in their research were oftentimes borrowed from psychology and behavioral sciences. Gervais and Fessler (2017), for instance, provided the following definitions:

- attitudes: "enduring affective valuations that represent relational value"
- emotions: "occurrent affective reactions that mobilize relational behavior"
- sentiments: "higher-level functional networks of attitudes and emotions that serve critical bookkeeping ... and commitment." (p. 3)¹²⁰

Applied linguists have been interested *inter alia* in the attitude construct because there is a potential relationship between attitude and behavior: The attitudes of speakers might be an indicator of the lexical items they (will) use. A good review of the relationship between attitude and behavior in language planning can be

119 Opinion cannot be completely separated from knowledge, as opinions are influenced by the acquisition of new knowledge.

120 Bookkeeping here is to be understood, in my comprehension, as a term in psychology, referring to Rothbart's cognitive model that proposes "a gradual modification of stereotypes by the additive influence of each piece of disconfirming information" (Johnston & Hewstone, 1992, p. 361).

found in the work of Triano López (2007, pp. 15–37). However, as Vandermeeren (2008) underlined, measuring attitudes is not unproblematic:

It was made clear that to date, both social psychologists and sociolinguists must face the fact that quantifying attitudes still entail conceptual and methodological problems. The attitude concept itself and especially the attitude-behaviour-relation are indeed controversial issues. (pp. 13–26)

Below, I review methods used to approach lexical opinion. Several researchers that I do not mention here (e.g., Daoust, 1995; Karabacak, 2009) also tried to grasp the bigger picture in a specific domain, for instance, speakers' general attitudes toward francization. However, in my view, such an approach neglects the fact that speakers are often unaware of the sources from which the lexical items they use originally come.¹²¹ In what follows, I start by explaining what researchers did in relation to lexical opinion, then provide a summary table of methods for comparison purposes, and finally, I comment on the methodological problems posed by existing studies.

In her study on parts of the car, Allony-Fainberg (1974) investigated the attitudes of speakers toward Hebrew target lexical items. In this paper, “attitude” refers to whether a speaker finds a lexical item semantically appropriate and convenient for use. The questions thus concern both the lexical item itself and potential use of the lexical item. It would be hard to draw any conclusion from her findings because results for both questions are mixed in a single table (Allony-Fainberg, 1974, p. 500), and, as she herself noted, “the results yielded by this research can hardly provide any final attitudinal conclusions for all types of words as there were not enough questions asked concerning the words studied in this connection” (pp. 500–501).

In the third part of their study, Gaudin and Guespin (1993) collected opinions on target lexical items (p. 12). Their qualitative approach using metalinguistic statements during interviews seems particularly interesting to better understand why target lexical items might be accepted or rejected by the target speech community. The authors did not seem to define “opinion,” so it is not clear exactly what kind of statements they were trying to obtain from the interviewees. Nevertheless, they were able to uncover why some target lexical items were not accepted

121 If a speaker generally has a positive attitude toward francization, for instance, it does not necessarily mean that they will prefer a French term over an Anglicism, for instance. Thus the studies I reviewed concentrate on the level of (target) lexical items.

by target speakers: The French abbreviation ACP, for instance, was rejected because it is homonymous with acid carrier protein.

In a long survey, Gouadec et al. (1993) asked respondents (question 7) to provide examples of lexical items with specific characteristics (e.g., terms that are expressive, correct [precise], surprising, complicated, ridiculous, etc.; p. 240). They were able to collect some examples of lexical items that fit into those predefined categories, but it should be noted that the nonresponse rate here was above 50% (27 respondents out of 50).

Thoiron et al.'s (1993) primary goal was not to study speakers' lexical opinions. Nonetheless, in the first part of their semistructured interviews, through two questions, they tried to capture the interviewees' preferences for English lexical items (borrowings) versus French ones (the target lexical items).¹²² They asked speakers, for cases in which they knew both an English and a French term for a given concept, which one they preferred and why; they further asked if interviewees would replace the English term with its French equivalent if they were given the French equivalent, and why. From my understanding, these questions were general and did not concern specific terms.

Triano López (2007) investigated, at the lexical level, "how speakers' attitudes and language planning can interact in shaping the course of language development" (p. 1). His study provides insights on the use of Castilianisms in Catalonia (borrowing) versus indigenous Catalan lexical items (here, the target lexical items). Interestingly, he sought not only to assess speakers' attitudes but also to establish whether there was a relationship between these attitudes and speakers' actual lexical behavior.¹²³ The author developed four indexes in relation to speakers' attitudes: a language loyalty index, a lexical loyalty index, a corpus planning index, and a status planning index. Some of the questions related to his lexical loyalty index are relevant for lexical opinions (e.g., he asked respondents to assess whether Castilianisms that replace a Valencian word that already exists are acceptable).

In their study of implantation for climbing sport in Catalan, Nogué and Vila (2007a) asked respondents directly what they thought about the target lexical item proposed by the standardizing body (TERMCAT). In this way, they partly collected positive, negative, or neutral opinions on the target lexical item (e.g., "I do not like it"), as well as metalinguistic statements of the type "strange word,"

122 This is comparable to Triano-López's construct of 'lexical loyalty' (2007, p. 1).

123 Generally, research on language attitudes focuses mainly on the search of a causal relationship between (language) attitudes and language behavior (Lenz, 2003, p. 264).

“good sonority,” “more correct,” “too long,” “complicated,” “inadequate,” or “it does not correspond to the concept.” In several cases, the researchers noted that for the same item, the answers varied from one respondent to the other.

Nogué Pich and Vila i Moreno (2007b) conducted another study on field hockey in Catalonia. One of the aims of their semistructured interview was to “capture opinions about the [lexical items] and standardization efforts in the field of sports in general” (Nogué Pich & Vila i Moreno, 2007b, p. 177, my translation).^{xi} Based on their interview guide, it seems that the sole question they used to collect these opinions was “Do you think it [the target lexical item] could end up spreading?” (Nogué Pich & Vila i Moreno, 2007b, p. 232, my translation).^{xii} They were able to collect isolated opinions (e.g., target lexical items were sometimes said to be confusing or not adequate for the concept in question). They used a mixed-method approach, in which lexical filtering remained only a peripheral aspect.

Leblanc and Bilodeau (2009) examined the adoption in spoken communication of a sample of target lexical items in the IT domain (p. 167). The authors held semistructured interviews with individuals in Quebec and New Brunswick, ranging from the occasional IT user to the professional. The authors announced that the data would be collected using interviewees’ *epiterminological discourses*.^{xiii} They used predefined questions to collect statements. In the introduction and the conclusion (2009), they mentioned that the type of *epiterminological* discourse they are considering are *evaluative* in nature and they speak of linguistic *judgements* (p. 168, 178). However, in their analysis, they distinguished three types of *epiterminological* discourse (Leblanc & Bilodeau, 2009, pp. 171–172), none of which appears to be of evaluative nature: explicative discourse (definitions, paraphrases, etc.), discourse that shows traces of the English language, and silent discourse (hesitations, etc.). I fully agree that metalinguistic statements can be of evaluative nature and that it would be worth collecting such statements, but there exist other types of metalinguistic statements as well,¹²⁴ which are *not* of evaluative nature, and from reading the paper results, I have the impression that this is what the authors collected most in their study.

In his doctoral dissertation, Remysen (2009) analyzed “how language columnists identify certain usages as French Canadian [and] look[ed] at the ways these columnists describe the usages they comment on. It also examine[d] the value judgments they make about them as well as the arguments they put forward to

124 This will be discussed in Chapter 6.

justify these judgements” (p. ii). Remysen (2009) classified the results into arguments to support the acceptance or condemnation of lexical items:

	Arguments to support acceptance (Remysen, 2009)	Arguments to support condemnation (Remysen, 2009)
Objective norms	1.a) 1 ⁰ Semantic proximity	2.a) 1 ⁰ Semantic gap
	1.a) 2 ⁰ Compliance with morphological rules	2.a) 2 ⁰ Noncompliance with morphosyntactical rules
	1.a) 3 ⁰ Lexical gap	2.a) 3 ⁰ Lexical overlap
Subjective norms	1.b) 1 ⁰ Linguistic authorities/references	2.b) 4 ⁰ Linguistic authorities/references
	1.b) 2 ⁰ Cultural or identity-related significance	–
	1.b) 3 ⁰ Gallo-Roman or French origin	2.b) 1 ⁰ Foreign origin (English)
	1.b) 7 ⁰ Frenchness	2.b) 3 ⁰ Absence of Frenchness
	1.b) 4 ⁰ Clear or expressive	2.b) 7 ⁰ Imprecision or ambiguity
	–	2.b) 6 ⁰ Impeding intercomprehension
	1.b) 5 ⁰ Aesthetic character	–
	1.b) 6 ⁰ Established in Canadian language use	2.b) 2 ⁰ Established in France or in the Francophonie
	–	2.b) 5 ⁰ Unacceptability due to diachronic, diatopic, diastratic or stylistic variation

Table 3. Translated and adapted from Remysen (2009).

Ní Ghearáin (2011) investigated “the hypothesis that official terminology planning is not well received by the Irish language speech community in the Gaeltacht” (p. 306). Throughout her semistructured interviews, she gathered

metalinguistic statements concerning target lexical items.¹²⁵ This qualitative study was *explorative* in nature rather than systematic about lexical opinion.

Saint (2016) collected 330 Twitter messages mentioning a specific target lexical item in the months following its publication by the language managers concerned. She found metalinguistic statements expressing a personal preference in about two thirds of them (polarized opinion about the target lexical item itself or the language managers concerned).

As I write these lines, Nissilä and Pilke (2017) are studying the correspondence between subject experts and language managers in Sweden between 1941 and 1983. In the letters exchanged between the two parties, aspects such as information on the “right” lexical items for specific technical concepts are discussed. One of the aims of the authors’ research is to characterize the content of experts’ opinions and to review their position and arguments regarding proposed target lexical items and their definitions. Final conclusions cannot be drawn here because Nissilä and Pilke’s (2017) project is still a work in progress, and the 3000+ letters have yet to undergo a more in-depth analysis (p. 253). However, because their preliminary observations suggest that the most common type of question language managers asked experts was about term acceptability (Nissilä & Pilke, 2017, p. 247), this project seems to be a promising future source of data for the study of lexical opinion.

125 E.g. “I don’t like it... it’s too... prissy or something,” “strange new terms that no one understands,” “it’s not very professional,” “it’s easier to just use the English word,” “If you used [...] a brand new word ... maybe the other person wouldn’t understand that word.”

Aspect	Study	Method	Tool
Lexical opinion	(Allony-Fainberg, 1974)	Ask speaker about semantic appropriateness of lexical items and convenience of use of lexical items	Survey
	(Gaudin & Guespin, 1993)	Collect metalinguistic statements of speakers about lexical items	Interviews
	(Gouadec et al., 1993)	Ask speakers for examples of lexical items with specific characteristics	Survey
	(Thoiron et al., 1993)	Ask speakers general questions about their preferences between Anglicisms vs. French lexical items	Interviews
	(Triano López, 2007)	Attitudes of speakers (various dimensions, indexes)	Survey
	(Nogué Pich & Vila i Moreno, 2007a)	Ask respondents directly about lexical items and collect metalinguistic statements about lexical items	Interviews
	(Nogué Pich & Vila i Moreno, 2007b)	Ask respondents about the use of lexical items	Interviews
	(Leblanc & Bilodeau, 2009, p. 167)	Ask predefined questions to make interviewees utter terms and metalinguistic statements	Interviews
	(Remysen, 2009)	Collect metalinguistic statements in naturally occurring data	Written corpus (language columns)
	(Ní Ghearáin, 2011, p. 306)	Collect metalinguistic statements by investigating interviewees' practices and attitudes	Interviews
	(Saint, 2016)	Collect metalinguistic statements in naturally occurring data	Written corpus (social network)
	(Nissilä & Pilke, 2017)	Collect metalinguistic statements in naturally occurring data	Written corpus (letters)

Table 4. Summary of methods used to investigate lexical opinions.

As illustrated in Table 4, up until recently, speakers were asked questions directly (observational bias) or statements were collected in the presence of the researcher through surveys or interviews. This again poses the problem of the observer's paradox and the fact that data are elicited for the lexical item level (for details, see previous section).

Of the authors reviewed above, Triano López (2007) is the only one who provided a convincing definition of the attitude construct he was examining and a sound theoretical framework for his concepts. Thus, there is again an issue regarding the constructs researchers are trying to measure, especially that of attitudes. As Broermann (2008) noted (and it is not different here), with but few exceptions,¹²⁶ researchers in linguistics have directly borrowed and adapted the attitude definitions and models from social psychology and have not developed constructs that are specific to linguistics (p. 26). Several scholars agree that, despite a large range of studies on language attitudes, sound research approaches are still missing or incomplete (Arendt, 2010, p. 10; Casper, 2002, p. 95). Generally, Neuland (1993) noted that the experimental methods used in attitude research lead to artificial test situations for which the validity of the results is questionable (p. 729).

The last two studies reviewed (Nissilä & Pilke, 2017; Saint, 2016) and that of Remysen (2009) are interesting because they are based on real language data. The former two are rather limited in scope (Saint's [2016] study, for instance, concerns only one lexical item), and that of Remysen (2009) analyzes the opinions of language columnists, not of folk speakers, but they open promising perspectives to work with naturally occurring data and allow the researcher to escape the problem of the attitude construct. Using corpora indeed turns the problem around: Instead of starting off with a problematic theoretical construct, here, the idea is to collect and observe naturally occurring data and to determine whether these data can be useful for language managers. This eliminates the problem of defining a construct a priori and that of the observer's paradox. Thus, this is the approach I want to develop in the present thesis and for which I will further argue in Section 3.4.2.

126 Vandermeeren's attitude-manifestation-model is one of these exceptions (2008, p. 1320), combining theories and variables from different disciplines. As Casper notes however (2002, p. 114), there is no systematic relationship between the components of the model.

3.3.3 Previous studies on lexical replication

In the three-step implantation process described in Section 3.3, the last step is lexical replication. Unlike the previous two steps, this one will not be part of my proposal (see Part 3 of the investigation). This is because, as I am about to explain, an operational monitoring tool that could be used by language managers has already been developed.

Previous researchers, such as Allony-Fainberg (1974), in the 1970s were already interested in this aspect.¹²⁷ From the 1980s on, studies flourished, especially under the French term *études d'implantation* (implantation studies; see the review in Quirion, 2003, pp. 51–66). A measurement protocol was developed by Quirion (2003a) to evaluate to which degree the target lexical items are used in effect by the target speech community. As he himself explained, Quirion (2003) was not the first researcher to work on the challenging issue of implantation, and one of the drawbacks of previous proposals is the fact that they were not reproducible (pp. 56–58). Quirion's (2003) protocol, in contrast, is operational, has been reused by other scholars (e.g., Saint, 2013), and could very well be used by language managers. It is based on naturally occurring data (corpora) and reflects how speakers behave in real life, whether they utter the target lexical item in actual interactions. The metric used in this protocol is the **implantation coefficient**, defined as the ratio between (a) the number of times the target lexical item is used for a given concept and (b) the number of times this concept is mentioned (Quirion, 2014a, p. 288).¹²⁸

The metric as developed by Quirion (2003) has become the standard for evaluation during the step of lexical replication. I will not discuss its details,¹²⁹ but I would like on a parenthetical note to make a short comment on the wording used by some scholars. Some researchers (not Quirion) claimed in their work that they used the implantation coefficient to measure the impact of deliberate lexical interventions:

*The objective of the article is to **measure the influence of the linguistic recommendations** produced and circulated by the Office québécois de la langue française. For this study, I ... **measured the degree of implantation***

127 She, for instance, asked respondent to self-report whether they used specific terms.

128 Such a metric had already been semi-explicitly mentioned by Wüster (he speaks of a certain measure, see 1931, p. 177), who compared the 1913 state of language usage for designational paradigms in the aviation domain with that of 1929.

129 I invite the reader to refer to Quirion's work directly, in English see (Quirion, 2003b).

of officialized terms for computer peripherals and that of other French and English equivalents. (Saint, 2013b, p. 167, my emphasis)

To approach the study of terminological implantation, that is, of the impact that the diffusion of certain terminological standardization proposals has had. (Gresa Barbero, 2016, p. 21, my translation and emphasis)^{xliii}

I would like to underline that the implantation coefficient metric measures a *state* of language usage. Evidently, this state of language usage might, to a greater or lesser extent, be *related to* deliberate lexical interventions, to the interventions of language managers. However, the implantation coefficient is at best only an indirect indicator of an intervention's effectiveness because environmental factors also influence language actualization. The implantation coefficient thus measures the "post-intervention outcome" or "post-intervention state of language use," not the "impact of a deliberate lexical intervention."

3.4 Proposals for language managers

In the present investigation, I focus on using listening strategies for contending (see Section 3.2). I do this to help language managers overcome low levels of lexical knowledge on the one hand and negative lexical opinions on the other. To this end, I do not adopt an approach centered on the target lexical item but rather a comprehensive one. As Peter and Olson (2010) underlined, "[M]arketers have to analyze and understand not only consumers of their products and brands but also consumers of competitive offerings and the reasons they purchase competitive products" (p. 13, my emphasis). Previous studies regarding lexical knowledge and lexical opinion have often preimposed their sets of questions upon speakers and concentrated on target lexical items, failing to observe how individuals actually "buy their products," (i.e., how speakers select the lexical items they would like to use and why they may sometimes opt for rival items instead of target lexical items).

This thesis makes two proposals: (a) explore speakers' lexical environment to better identify the sources from which speakers actually might draw their lexical material and why (see Section 3.4.1), and (b) explore speakers' lexical opinions in context and develop a proof of concept showing this can be done using natural language processing and why the types of data observed in a corpus might be useful for language managers (see Section 3.4.2).

3.4.1 Exploring speakers' lexical environment

In Section 3.3.1, I explained that language managers face low levels of lexical item knowledge. Therefore, it seems essential to gain a basic understanding of where speakers learn new lexical material and through which channels they acquire new lexical items. Depecker (1997) called for a study of speakers' linguistic environments (pp. XXXIV–XXXV), and Heller (1978) mentioned that

efforts should be made to find out where terms are learned, where they come from, and how they spread within a generation. (p. 39, my translation)^{xliv}

Vila i Moreno (2007) mentioned the potential value of knowing where speakers learn new lexical items:

Thus, for example, knowing a priori the real diffusion channels for innovations provides a very valuable perspective when it comes to proposing innovative forms. (p. 247, my translation)^{xlv}

Also, Drame (2009) underlined that the “marketing” of the lexicon should by no means be limited to traditional channels:

Traditional methods like disseminating wordlists—which sometimes are the only channel used by terminologists in South Africa—are still useful tools; however, they must be regarded as one in many distribution channels. Embedding the information into local stories and topics that are relevant to the target group is crucial. (p. 18, my emphasis)

Evidently, lexical knowledge can be acquired passively or actively by speakers: It can happen by chance but also as a result of a speaker's expectations or needs (Martin, 1994, p. 34). In the present investigation, I propose to concentrate on cases where speakers actively search for lexical knowledge in a situation of perceived need (more will be said about the context in Chapter 6). Do speakers search for external information at all? If so, where? Online or offline? From traditional or alternative resources? Little is known.

Sociolinguistics is not particularly helpful for investigating this question because its theoretical considerations almost never explain where the variants come from in the first place (Croft, 2000, p. 55). The study of neology, which has become a well-established research area (Mejri & Sablayrolles, 2011, p. 3), also does

not seem to provide much help as to where new lexical items come from in the speech community. It has largely focused on the description of neological formation processes.

The most valuable and comprehensive approach so far to tackle this question is perhaps that of socioterminology:¹³⁰

Socioterminology is a convenient term that can be used to describe the relationship between society and terminology and especially the actual social use, whether by specialists or by ordinary people, of the terms coined by terminologists. (Maurais, 1993, p. 121)

Socioterminology is particularly concerned with the circulation of knowledge and the movement of terms. It is, above all, a theoretical interrogation, a scientific proposition in which variation is seen as central to negotiations, including, of course, terminological negotiations. Socioterminology is engaged in areas where standard usage is disrupted, where meaning has to be crafted in order to re-establish mutual understanding. (Delavigne, 2017, p. 33)

According to an ISO (2007) technical recommendation,¹³¹ one of the tasks of socioterminology is to identify the networks for disseminating target lexical items:

In identifying networks for disseminating terms, there are two types of work to be carried out: on the one hand, to describe the factors, the situations that favor or not circulation and implantation (how terminologies are infused or diffused into the professional environment), on the other hand, to list the methods, supports for terminology creation, for transmission (oral discourse, texts, databases, etc.), by using the possible logic of mediatization. (p. 9)

However, Depecker (2013) mentioned that so far, socioterminology has not taken the motivations of speakers into account (p. 18). To him, one needs a more detailed approach that would consider sociological, cultural, psychological, and ethnological dimensions. Depecker (2013) has pleaded for an *ethnoterminology*,

130 I am quoting contributions in English here, but the interested reader should refer to Gaudin's founding works (1993a, 1993b, 2003, 2005, 2007).

131 This 2007 recommendation has been withdrawn.

which he defines as the “area of terminology that studies terminologies from an ethnographical (field study) and an ethnological (generalization of observed facts, comparisons of human groups) perspective” (p. 27, my translation).^{xlvii} Hermans (1994) had also previously suggested using ethnographical methods and mentioned that lexicology could be seen as a sociological discipline because “using such or such language or such or such lexical unit is a speech (or language) act composed by the interaction of several components, language being just one of them” (p. 41, my translation).^{xlviii} The valuable component here is the consideration of the extratextual context in a comprehensive manner.

Pointing out the weaknesses of classical terminology and socioterminology, de Vecchi (2016) has argued for a pragmatic approach to terminology that takes into account differences in language use between companies (sociolects of corporate companies) or between communities within one and the same domain as well as the internal dynamics of lexical units that are actually in use (p. 132). According to de Vecchi (2004), four dimensions can be used for approaching language in corporate environments: (a) a terminological dimension, (b) a socioterminological dimension, (c) a pragmaterminological dimension, and (d) a temporal dimension (pp. 79–80). De Vecchi (2004) evidently contributed the pragmaterminological dimension to terminology approaches. The pragmaterminological dimension that he introduced must highlight the purposes and the usefulness of lexical items (terms) in accordance with actions to be undertaken and practices in a specialized domain as well as with the environment in which these actions occur. The foundation of the pragmaterminological approach is the description of lexical units within a company’s sociolect, including, for example, information on where it is used in the company and by whom (de Vecchi, 2016, pp. 132–133). Similar approaches have also been adopted by other researchers under the term “corporate lexicography” (Leroyer & Møller, 2006, p. 99). Two additional perspectives from seemingly unrelated fields (marketing and lexicography) also seem helpful for approaching the question of lexical environments: studies on information searches in marketing¹³² on the one hand and Tarp’s (2009) concept of “extralexical situations” on the other. I believe the following assumption from Tarp (2009) is essential:

132 “Because consumer information search behavior, in one fashion or another, precedes all purchasing and choice behavior, it has been a perennial topic of research. Consequently, the literature on consumer information search behavior is voluminous and possesses a long and rich history.” (Peterson & Merino, 2003, p. 101)

A person may, in an extralexical situation, have a lexicographically relevant need which he/she does not recognize and therefore does not try to solve consulting a dictionary, although dictionaries designed for this specific type of need may already exist. (p. 281)

This implies that if we want to understand where speakers actively search for lexical items, we must not start with specific lexicographical products (language managers' dictionaries, for instance, as was done in several studies reviewed above), but rather from speakers' extralexical situations. As previously stated, in our current sociolinguistic era, speakers tend not to get their information from traditional, institutional sources, but rather from alternative sources, such as websites, search engines, and their peers. This means that a speaker may have a lexicographically relevant need that he or she recognizes but does not try to solve using a dictionary (let alone using the dictionary of language managers).

Here, marketing again can provide us with insights and methods, because it has considered consumers' environments and in particular the way consumers search for information:

*Nearly every introductory marketing and consumer behavior textbook depicts the consumer purchase decision process as a series of steps progressing from problem recognition, to **information search**, to evaluation of alternatives, to purchase decision, and finally to postpurchase behavior. In the information search stage, consumers actively collect information to make potentially better purchase decisions. (Schmidt & Spreng, 1996, p. 246, my emphasis)*

In marketing studies, it is generally accepted that individuals can search for information in one of two ways: Either they conduct an internal search, extracting information stored in their memory, or they carry out an external search, seeking information from their environment (Schmidt & Spreng, 1996, p. 246). These two concepts are distinct but related: External search behavior is dependent on memory, and the overall search process can be iterative (Peterson & Merino, 2003, p. 101). The choices individuals make are constrained by the sources they consult:

The study of external search for information is critical in understanding the decision-making process of consumers, to avoid the otherwise necessary, though oftentimes unrealistic, assumption that choices are made under full information conditions. (Srinivasan, 2012, p. 153)

During external search, individuals may use various types of sources, some of which are controlled by marketers, but also third-party information (e.g., a magazine article), interpersonal sources (e.g., friends; Olshavsky & Wymer, 1995, p. 20; Schmidt & Spreng, 1996, p. 247), and in our digital era, online sources such as websites, online word-of-mouth, and online bulletin boards and forums (Bickart & Schindler, 2001; Klein & Ford, 2003).

Information search behavior studies have been conducted in various fields: for instance, in health care for understanding where cancer patients seek information about their condition (Finney Rutten, Arora, Bakos, Aziz, & Rowland, 2005) or in tourism research to understand how tourists search for information about potential travel destinations (Sparks & Pan, 2009). In these studies, the official source of information (the health care professional or the travel agent) is far from being the only source consulted, and often it is not the most influential.¹³³ I believe we can draw a parallel here because language managers and their official resources today clearly are not the only source of lexical information that speakers use.

As reported above, several scholars have investigated lexical source knowledge (Gresa Barbero, 2016; Martin, 1998; Ní Ghearáin, 2011; Nogué Pich & Vila i Moreno, 2007b), but they have only or mostly focused on lexical sources containing the target lexical items of language managers, failing to see the bigger picture, to question whether speakers consulted resources at all, and, if such is the case, which ones. As mentioned above, it is, however, essential to be aware of competitive products on the market.

The paths individuals take in their search for information can be diverse and complex. In marketing studies, various models have been developed to illustrate the information search behavior of customers (see, for example, Ratchford, Talukdar, & Myung-So, 2001; Schmidt & Spreng, 1996).¹³⁴ In the present investi-

133 In the two studies mentioned, the identified sources were respectively for cancer patient “health professionals, printed materials, media, interpersonal, organization/scientific” and for Chinese outbound tourists “television programs, friends, fashion magazines, travels books, newspapers, tourist brochures, chinese websites, television advertising, travel agent, family members, australian tourism websites, exhibitions/travel shows, radio, online chat with Chinese in Australia or have visited Australia, other personal channels (wom), previous personal experience, online chat with local Australians, other”.

134 According to Srinivasan, there have been three theoretical foundations to explain the variation in external search for information: “(1) the economics approach using the cost-benefit framework, (2) the psychological approach of motivation and person/product/situation related variables, and (3) the consumer information-processing approach which stresses the role of memory and human information processing limitations.” (Srinivasan, 2012, pp. 153–154)

gation, I will not design a model but rather try to gain first qualitative insights about sources of lexical knowledge and the various search strategies employed by speakers, mainly through focus groups (see Chapter 6), although I will also show in Chapter 9 (see Section 9.2) that information on lexical-source recall and on lexical-source opinion can be found in naturally occurring data, in contexts containing opinionated autonyms.

3.4.2 Exploring speakers' lexical opinions in context

Let me start this section with a definition of naturally occurring data. By naturally occurring data, I mean here that the recording of the data is "situated as far as possible in the ordinary unfolding of people's lives, as opposed to being prearranged, set up in laboratories, or otherwise experimentally designed" (Hutchby & Wooffitt, 2008, p. 12). In Section 3.3.2, I mentioned that previous studies often used elicited information about lexical opinions. But this type of information can also be obtained from naturally occurring data:

This discourse about languages can be explicitly elicited, using surveys that place speakers in a situation in which they react to utterances and produce judgments on spoken or written languages in a given speech community. But this discourse can also be identified in various discursive productions which are spontaneous or not (political speeches, unions' discourses, various texts such as newspaper articles, literary texts, pedagogic texts...), written or spoken, in particular in conflict situations in which languages are a factor of power. (Morsly, 1990, p. 80, my translation)^{xlviii}

As already stated, one of the obstacles language managers face during their deliberate lexical intervention is the negative lexical opinions of speakers. But naturally occurring data can provide information about how ordinary speakers chose the lexical items they use and how they filter lexical items.

Clearly, in most cases, speakers' choices are automatically made, with no conscious deliberation. Individuals do not reflect on each and every lexical item they use. Communication would be greatly impeded by such a process. As Selten (2002) underlined, to a great extent, individuals follow automatized patterns of behavior:

Much of human behavior is automatized in the sense that it is not connected to any conscious deliberation. In the process of walking, one does not decide after each step which leg to move next and by how much. Such automatized routines can be interrupted and modified by decisions, but while they are executed they do not require any decision-making. They may be genetically preprogrammed (e.g., involuntary body activities) or they may be the result of learning. (p. 16)

In the present investigation, however, I explore two special cases where lexical choices are made with a certain degree of metalinguistic awareness on the part of speakers:

- Lexical choices in collaborative groups where individuals work together on producing a lexical resource, and
- Lexical choices when an individual's lexical knowledge is insufficient and they thus seek information from a group of fellow speakers.

By **metalinguistic awareness**, I mean “the general level as the ability to think about and reflect upon the nature and functions of language” (Tunmer, Pratt, & Herriman, 1984, p. 2).¹³⁵ The groups, which are Internet-based, will be presented in more detail in Chapter 5 (p. 149).

I propose to take advantage of the fact that in the current Web 2.0 era, consumers are “leaving clues about their opinions, positive or negative” (Li & Bernoff, 2011, p. 81). As Saint (2016) showed in her recent paper, the Internet is replete with comments of speakers about lexical items. Individuals and language professionals (linguists, teachers, etc.) but also ordinary speakers, use metalanguage (i.e., they talk about language; see, for example, Jakobson, 1980; Rodríguez Penagos, 2004b, pp. II–26). Here, I use **metalanguage** as language about language.¹³⁶ In a natural language, one can say there is a **common language** or **object language** describing the world: tangible elements (a chair, a house, a school) as well as abstract thoughts (democracy, freedom, war). There is also a metalanguage, which speakers use to talk about language. A linguistic metalanguage can be anything on a wide range from a purposefully designed, formalized language to a completely natural, unaware language. Rey-Debove (1978) suggested the dichotomy of

135 Speakers have varying degrees of metalinguistic awareness (Mertz & Yovel, 2003).

136 For other definitions of metalanguage used in the scientific literature, see Berry (2009).

scientific-educational use on the one hand and common use on the other (p. 22). The scientific-educational use is the use linguists, language learners, language teachers, or any specialist of language makes of formal metalanguage. In the scientific-educational use, metalanguage can be natural or partly or wholly formalized. The common use is the use ordinary speakers make of a more natural metalanguage.

Speakers may or may not have metalingual awareness¹³⁷ when they use metalanguage (see Mertz & Yovel, 2003). In fact, speakers may talk about the norm or about correctness without even knowing the term *metalinguistic* (Houdebine-Gravaud, 2004a, p. 12).¹³⁸

Metalanguage can be used as a starting point for observing speakers' representation of language. As Marcellesi (1971) noted, the metalinguistic activity of speakers regarding lexical items can be observed in their speech (both written and spoken):

*The corpus study demonstrates to me that, as we will see, **subjective factors or metalinguistic considerations intervene in the choice and use of certain lexical items.** Metalinguistic considerations, in that vocabulary users to some extent distance themselves from this vocabulary and do not fully assume responsibility for it yet; also in that they take certain precautions when employing English vocabulary, which they mark from the corresponding French term, or when **they comment on the validity of the use of the terms.** And subjective factors, which involve certain underlying connotations in the choice of terms. (p. 66, my translation and my emphasis)^{xlix}*

137 In the present investigation, following Berry I use the adjective 'metalinguistic' to refer to awareness of language and the adjective 'metalingual' to refer to the awareness of metalanguage (2009, p. 12).

138 Some scholars are interested in the distinction between conscious and unconscious metalinguistic activity and coin distinct terms for each of them. Culioli, for instance, makes a distinction between metalinguistic (conscious) and epilinguistic (unconscious) (see Beaulieu-Marianni, 2012, p. 113). Some scholars, before all in the francophone world, reject this distinction and use the term epilinguistic or epilinguistic discourse to refer to speakers' judgements and evaluative statements about language (Canut, 1998, p. 70). For a discussion of *metalinguistic* versus *epilinguistic* see for instance Beaulieu-Marianni (2012, pp. 113–116). Lucy in addition discusses consciousness in relation to reflexive language (1993, pp. 21–29). In my investigation, this distinction does not appear to be relevant since my starting point is what is visible in corpora, regardless of the degree of metalinguistic awareness of speakers. Therefore, I only use the term 'metalinguistic'.

Here, I call these “sentences where discourse reflects upon itself, where language itself is the subject, where language is creating and manipulating the elements and rules that make it possible” explicit metalinguistic statements, borrowing the notion and its definition from Rodríguez Penagos (2004b, pp. I-4–I-5).¹³⁹

For language managers, the fact that a target lexical item that is neological for speakers can be expected to generate explicit metalinguistic statements in speakers’ discourse is interesting.¹⁴⁰ This is because at the level of a speaker, a norm is not merely a convention, a set of regularities. A norm implies an expectation of regularity (Bartsch, 1987, p. 157) on the part of a speaker’s interlocutors. Innovations in language and thus most target lexical item of language managers are by definition infractions to existing norms (objective norms). If a target lexical item is neological for a specific speaker, it violates the status quo—neologisms are abnormal with reference to their objective lexical norms. A new lexical item is always an infraction to the speaker’s expected norm, at first a mismatch compared to the norm that must be positively sanctioned by the speech community to enter the actual lexicon (Juhász, 1970, pp. 33–34).

Far beyond the linguist’s constructs of norm, speakers have their own subjective norm, their own representation of language, and their own expectations of how language should ideally be. Because language(s) are part of human life and they shape social life, it is common for speakers of a language to have personal views about language, in a similar way that they have opinions about other aspects of life (Wilton & Stegu, 2011, p. 4).

Value judgements on language form part of every competent speaker’s linguistic repertoire. One of the things that people know how to do with words is to evaluate them. (Cameron, 1995, p. xi)

Using speakers’ metalanguage, it becomes possible to observe speakers’ opinions about lexical items. I already mentioned the necessity to gather data directly from the “marketplace,” from the speech community, not only about the target lexical item but also about rival lexical items. The decisive advantage in my investigation is the use of naturally occurring data, visible on the Internet, which can be gathered in the form of a corpus. As Arendt and Kiesendahl (2015) mentioned,

139 But not the exact term: Rodríguez Penagos speaks of ‘Explicit Metalinguistic Operations (EMOS)’.

140 Explicit metalinguistic statements, as their name indicates, are explicit, i.e. they are visible manifestations found in language, but they are not necessarily conscious.

Via forum communication, it is possible to access a large spectrum of comments which fall under language criticism. In these comments, ordinary speakers' norms become visible. (p. 165, my translation)¹

Rey-Debove (1979) pointed out that natural metalanguage can be treated as a kind of language data (p. 15). As Paveau (2011) stated,

Folk propositions are not necessarily false beliefs that must be eliminated from the sphere of science, but constitute perceptive, subjective and incomplete forms of knowledge that need to be incorporated into the scientific data of linguistics. (p. 41)

Today, this may appear to be a truism, but for years, linguistics ignored data from ordinary speakers. Modern linguistics long considered exclusively speech acts of informants as objective data and ignored informants' statements about language because, as these statements were subjective, they were seen as subjective data (Neuland, 1993, p. 723). Although the interest for what is said by nonlinguists about language is not new,¹⁴¹ for a long period of time it was not the object of study of a specific scientific field. The interest in what ordinary speakers say about language is usually dated to the early 1960s in relation to the UCLA Sociolinguistics Conference and Hoenigswald's talk entitled "A Proposal for the Study of Folk-Linguistics" (as cited in Niedzielski & Preston, 1999, p. 2). Back then his proposal had very limited impact at first (Achard-Bayle & Paveau, 2008, p. 6). The field developed mainly in the Anglo-Saxon¹⁴² world and later in Germany¹⁴³ but remained nearly absent in other areas, for instance in the francophone world¹⁴⁴ (Achard-Bayle & Paveau, 2008, p. 4). Today, there is an increasing interest for folk data.

Explicit metalinguistic statements have been studied from several perspectives, for instance:

.....
141 According to Niedzielski and Preston, from a linguistic viewpoint it is at least as old as the 19th century German-language publication about the people's opinion of language (see Polle, 1889).

142 According to Coupland and Jaworski (1993), most of the sociolinguistic research into metalanguage was done under the headings "language attitudes", "folk linguistics" or "language awareness".

143 Mostly under the headings *Volklinguistik* and *Laienlinguistik*.

144 Where it is found under the terms *linguistique populaire, linguistique ordinaire, linguistique profane, linguistique des profanes...* (Achard-Bayle & Paveau, 2008, p. 5)

- by Rodríguez Penagos for mining knowledge from texts (e.g. 2004b, 2004a, 2005)
- by folk linguistics studies (e.g. Achard-Bayle & Paveau, 2008; Antos, 1996; Niedzielski & Preston, 1999; Paveau, 2011; Paveau & Rosier, 2008; Preston, 2004, 2011; Stegu, 2008; Wilton & Stegu, 2011)
- by language criticism¹⁴⁵ studies (e.g. Arendt & Kiesendahl, 2015; Griesbach, 2006; Heringer & Wimmer, 2015; Kilian, Niehr, & Schiewe, 2010; Meier, 2015; Schütte, 2015)
- by linguistic imaginary studies (e.g. Canut, 2000; Houdebine & Fodor, 2013; Houdebine-Gravaud, 2004; Obreja, 2012, 2012; Remysen, 2009, 2011; Rheault, 2004, 2010; Tsekos, 1996), and
- by studies concerned with lexical interventions (e.g. Gaudin & Guespin, 1993; Leblanc & Bilodeau, 2009; Ní Ghearáin, 2011; Nogué Pich & Vila i Moreno, 2007a; Saint, 2016)

In the present investigation, I will adapt Rodríguez Penagos' methods (2004b) to fit the objectives of language managers (see Chapter 7 for methods, p. 210). I will use, as an entry point to the lexical marketplace, autonyms (i.e., elements of language that are used to refer to themselves; see Carnap, 1934, p. 109).¹⁴⁶ Lexical autonyms (underlined in the examples below) are an interesting entry point for gathering lexical market data because they are surrounded with information (highlighted in grey below) on both the lexical item standing in autonymical condition¹⁴⁷

“Selektiva” is a bad word, although it's probably tempting for some users, and all inhib-words (inhibi, inhibicio, inhibitorio) are completely unnecessary and only blur the matter.

[excerpt from my corpus]
my translation^{li}

.....
145 German ‘*Sprachkritik*’.

146 Any sign of language (phoneme, syllable, quotes...) can stand in autonymical condition (see Authier-Revuz, 2003, p. 76), but here, the scope is restricted to lexical items (lexical autonyms) since I am concerned with lexical interventions. Thus I will not deal with other types of autonymical forms, such as reported speech, as in “Mary said to me yesterday at the station, ‘I will meet you here tomorrow’” (Lee, 1993, p. 369).

147 As we will see, explicit metalinguistic statements containing a lexical autonym can be of different natures, such as descriptive, prescriptive, evaluative, or authoritative (see e.g. the typology in Rheault, 2004, p. 28). Here, I am only giving a single example for illustrative purposes.

... as well as on the source where a lexical item was found:

How do you say “walkie-talkie” or “CB radio”

I only found the words *portebla radiotelefono* in relation to *walkie-talkie* in the bilingual (French-Esperanto) dictionary published by SAT-Amikaro in 2000.

For *walkie-talkie*, I found the word *promenradio* (Minnaja).

[excerpt from my corpus]
my translation^{li}

The information on the sources of lexical items will allow me, in Chapter 9 (see Section 9.2), to complete the data obtained through the focus group study (see Chapter 6). The data on lexical autonyms themselves will allow me to monitor what speakers are saying about lexical items and specifically to offer a comprehensive panorama of the aspects speakers discuss when choosing lexical items (see results in Chapter 8). Instead of starting from a theoretical construct (language representation, language attitudes, linguistic imaginaries, etc.), which varies from theory to theory and has been shown to be difficult to measure in practice for language-related data, I will “let the data tell the story” (Schwartz & Ungar, 2015, p. 79), adopting a data-driven approach.

Before moving to the empirical part of the investigation, comprising the focus group study (see Chapter 6) and the corpus study (see Chapter 8), in the next two chapters I will present the speech community in which the empirical investigation has been conducted (see Chapter 4) and the electronic networks of practice considered for the investigation (see Chapter 5).

Part 2

Esperanto for scientific research

4 Using the Esperanto speech community to study lexical phenomena

4.1 Introduction

In this chapter, I follow a dual objective: (a) presenting the main characteristics of the Esperanto speech community and (b) showing why some of these characteristics make the Esperanto speech community a particularly relevant object of study for the present investigation.

The first section (see Section 4.2) briefly recounts the history of the language, explaining how a fully functional decentralized speech community emerged from the plan of a single individual. It presents the main features of the speech community. The second section (see Section 4.3) explains in more detail how the original plan has been and is being completed from the perspective of the lexicon. It shows that Esperanto speakers, to a great extent, present a high level of metalinguistic awareness and explains why this trait makes the speech community especially interesting for the study of lexical opinions.

4.2 The emergence of a decentralized speech community

4.2.1 From a plan ...

It all started with a plan. This plan was the *Unua Libro*, published in 1887:

In July 1887 the first publication was printed, the Russian fundamental learning book, which had received permission the previous month by the Russian censor....The same year Dr. Zamenhof also published Polish, French and German translations of this first booklet, always according to the same plan. (Privat, 1912, p. 18, my translation)^{liii}

Esperanto is a **planned language**, a “system which has been consciously created according to certain criteria by an individual or a group of individuals for the purpose of making international communication easier” (Fiedler, 2006, p. 67). The term *planned language*^{liv} refers to the origin of the language. It can be opposed to

ethnic languages¹⁴⁸ (e.g., English, Russian, or Korean), because a planned language is not the creation of an ethnic group (Fiedler, 1999a, p. 25). Esperanto differs from ethnic languages in its history of origin, but nowadays it functions like other natural languages (Duin, 2006).¹⁴⁹ Because the language is still being used, language changes can be observed (see Philippe, 1991, pp. 136–264; Wood, 1979, p. 444).

As far as the lexicon is concerned, Esperanto—like other natural languages—has, for instance, developed idiomatic expressions over time (Fiedler, 1999a, pp. 28–30). Though with different frequencies, idioms occur both in specialized and general language texts (Dasgupta, 1993, p. 370; Fiedler, 1999a, p. 261), and the language has an evolving phraseology (Fiedler, 1999b, p. 51).¹⁵⁰ As in ethnic languages, identifying where changes originate in the Esperanto speech community remains difficult:

Just like in ethnic languages, dating new language phenomena in Esperanto is an issue that is usually difficult to solve. Every innovation is propagated only in successive steps and starts with a speech act of a given individual or in the case of a primary interference with the language habits of a specific group of speakers sharing the same first language.... Only exceptionally is it possible to deduce especially at what point in time and from which speakers the innovation originates. This is because the innovation can usually only be noticed once it has already been universally adopted. (Philippe, 1991, p. 100, my translation)^{iv}

Esperanto is only one of many languages that have been consciously planned. There have been hundreds of them: According to Sakaguchi (1998, p. 271),¹⁵¹ Duličenko (1990) listed more than 900 of them. The total number of planned

148 Ethnic languages can be considered the languages of the nations, nationalities and tribes (see A. D. Duličenko, 1989).

149 I consider Esperanto to be a natural language (a language for human communication that has evolved naturally). See my note above (62, p. 56).

150 Phraseologisms in Esperanto have historically been developed both consciously and spontaneously (Fiedler, 1999a, pp. 333–334): consciously in order to give the planned language more expressive power and cultural wealth, and spontaneously based on the inner structure of the language and reflecting the culture of the speech community.

151 I prefer to rely on Sakaguchi's German monography rather than quoting Duličenko's Russian work directly, because as I write these lines my knowledge of Russian is relatively limited.

languages evidently depends on the definition chosen for the concept of a planned language. Back (1996, p. 884), for instance, proposed a more restrictive number of about 300 planned languages, excluding items such as mere mentions of planned languages, outlines, revised versions, and pasigraphies listed by Duličenko (1990).

There also exists a great variety of planning strategies and resulting plans for a language. This variety can be classified based on several criteria. One of the most widely cited classifications of planned languages is the etymological classification,¹⁵² comprising the following:

- **a priori planned languages**, where lexical items bear no similarities to the lexical items of ethnic languages,
- **a posteriori planned languages**, where lexical items have been borrowed from ethnic languages and adapted, and
- mixed systems (see also Fiedler, 1999a, p. 26)

As Liu (2001) pointed out, “The creator of a planned language has the right to select source languages for his project. He can also construct some new elements which do not exist in any ethnic languages” (p. 126).¹⁵³ Esperanto belongs to the second category above, that of a posteriori planned languages: When the language plan was devised, morphemes were borrowed from the main European languages (D. Blanke, 2006, p. 203). Gledhill (2000) estimated that Esperanto comprises Latinate words (70%), Esperanto words (12%), Germanic words (10%), Indo-European words (5%), Greek words (< 2%), and Balto-Slavic words (< 1%; p. 122–126).

In Moch’s classification, a posteriori planned languages are further placed on a scale from naturalness to schematicity. Naturalistic planned languages pursue a receptive ease of use (direct understandability of the lexicon for its target users) and schematic ones a productive ease of use (explicit mechanisms of lexical formation; as cited in Schubert, 2015, pp. 2214–2215). Esperanto is typically a planned language of the schematic type. Its schematicity is characterized by the following:

.....
152 This classification is often attributed to Couturat/Leau (1903) and/or Moch (1897) by scholars, although, as Schubert recently noted (2018), they are not the originators of this classification.

153 As Liu also points out, a planned language may also find its lexical sources in a single language, such as Latino sine flexione (Latin) or Basic English (English) (2001, p. 144).

- its uniform structure¹⁵¹ (Fiedler, 1999a, p. 26; Janton, 1977, p. 12)
- its regularity in word formation (Schubert, 2015; Tauli, 1968, p. 168)
- its active or productive ease of use (Schubert, 2015)¹⁵⁴
- the immediate recognizability of parts of speech in its lexical items (Monnerot-Dumaine, 1960, pp. 48–49)
- its autonomy from ethnic languages (Tauli, 1968, p. 168)¹⁵⁵

These characteristics make Esperanto considerably different from naturalistic projects such as Otto Jespersen’s Novial or the Interlingua project, to which linguists of renown like Jespersen and Edward Sapir contributed.

4.2.2 ... to a fully functioning language ...

In a similar way, language managers’ target lexical item remain in vitro lexical items as long as they are not used by the target speech community, and a language plan on paper is not a language; it is only a project.

Esperanto was made public in 1887 but initially was used only in writing. The first conversation in Esperanto took place between Ludwik Lejzer Zamenhof, the medical doctor who initiated the language, and Antoni Grabowski, a chemical engineer. It was later followed first by club meetings at the local level and eventually an international congress in France in 1905 (Wood, 1979, p. 440). Progressively, a speech community developed.

Other projects for languages were not as successful. Of all the languages that have been planned, very few have become functional.¹⁵⁶ According to the scale developed by D. Blanke (2006) to classify planned languages according to their degree of practical use or realization, planned languages can range from a **planned language project** (a planned language with no practical application) to a **planned semilanguage** (a planned language with limited practical application) and finally to a **planned language proper** (a planned language with well-developed communicative functions; pp. 49–72). The scale comprise 28 criteria or steps, starting

-
- 154 Word formation in Esperanto is based predominantly on compositionality. Basic elements (roots) can combined to form units that are more complex.
 - 155 Although it may borrow lexical material in the form of roots from ethnic languages, Esperanto possesses an autonomous derivation system based on a regular scheme which does not depend on an existing ethnic language (Monnerot-Dumaine, 1960, p. 51. 71–72).
 - 156 Although at a completely different linguistic level, it is somewhat reminiscent of many target lexical items planned by language managers.

with the elaboration of a planned language manuscript, its publication, and the development of teaching materials up to the use of a family language, the development of a separate culture, and language change phenomena (D. Blanke, 2006). According to Liu (2001), “Less than 10 planned languages can be considered and studied like a real language sociolinguistically” (p. 130).¹⁵⁷

Esperanto is used in various spheres of activity, as a family language (Lindstedt, 2010) but also in specialized communication (D. Blanke & W. Blanke, 2012, 2015; Maradan & Mueller, 2012) and in professional settings (Chrdle, 2013). Although relatively European (Parkvall, 2010),¹⁵⁸ Esperanto has grown to be spoken on every continent of the globe, including Africa (see Goes, 2007). It has been used uninterruptedly for more than a century now by a stable community and has become a natural language from a linguistic viewpoint (Duin, 2006). Esperanto has shown that a planned language project can become a fully functioning language, and it has outlived all of its competitors (D. Blanke, 2009, p. 252). In the classification by D. Blanke & W. Blanke (2015), Esperanto is the only planned language that has become a planned language proper (i.e., a fully planned language with well-developed and various communicative functions; p. 219). Thus, it is the only planned language that can be used to examine processes of a sociolinguistic nature such as the ones considered in the present investigation.

4.2.3 ... spoken mainly by a decentralized community

Today, the Esperanto speech community is increasingly decentralized and grounded in informal Web-based communication (Tonkin, 2015, p. 182). It exists similarly to a diaspora (Fiedler, 1998, p. 27) and lacks a politico-geographical

157 It is possible (and, I think, interesting) to study e.g. the manuscript of a planned language *project*, but if the language has not been used (e.g. because the manuscript has not been published), aspects such as language change or variation are inexistent and thus cannot be studied.

158 The Europeanness of Esperanto concerns especially the lexical roots it borrows from other languages (until now borrowings have occurred mostly from European languages, see the figures given by Gledhill above), but not such much lexical grammar, which is often said to be of agglutinative nature, a characteristic largely absent in Europe (with a few exceptions such as Uralic languages).

base (Wood, 1979, p. 436). Some Esperanto speakers live in remote areas and do not regularly participate in congresses or get-togethers.¹⁵⁹

As Zamenhof published the *Unua Libro* in 1887, Esperanto has become much more than a language with a speech community. It has also become a social movement with a history (e.g., Drezen, 1930; Forster, 1982; Gobbo, 1997, pp. 184–246; MGA, 2011; Privat, 1927; Sikosek, 2006) linked to several—sometimes contradictory—ideologies. As Wood (1979) noted,

The decision to acquire a knowledge of Esperanto and thereby to obtain membership in the Esperanto-using sociolinguistic community is not solely, and for most speakers not chiefly, a purely linguistic decision. As has been pointed out, the Esperanto speech community is a social movement. (p. 435)

Because Esperanto speakers are found in all corners of the world, a fair share of their communication occurs online in written form. This applies in particular to volunteer or collaborative activities on wiki-like platforms or e-mail discussion lists. This generates a large collection of written speech acts that can be studied with the constructs of corpus linguistics and appropriate computer tools, making Esperanto particularly suited to explore speakers' lexical opinions in context.

4.2.4 Defining the Esperanto speaker

From a communicative perspective, there are two characteristics that make Esperanto different from ethnic languages (Waringhien, 1980, p. 253): It is mostly a second language, and it is an international language used by individuals with dissimilar linguistic and cultural backgrounds. The Esperanto speech community is thus mainly a community of second language, or L2, speakers (Fiedler, 2012, p. 76), of individuals who have acquired Esperanto as a foreign language, although there is a small minority of speakers who grew up with the language from birth.¹⁶⁰ Individuals speaking Esperanto from birth are called **denaskuloj** (persons from

¹⁵⁹ Zadzilko showed it empirically through a survey on the use of the Web 2.0 by Esperanto speakers (2011, p. 88), in which about 28% of respondents answered they do not go to Esperanto events, i.e. 66 respondents out of 239.

¹⁶⁰ Sometimes as a result of marriages between people of different languages or nationalities (Sherwood, 1982, p. 184), but according to Corsetti in a majority of cases Esperanto-speaking families do not consist of partners of different nationalities (1996, p. 266).

birth), and have been recorded since the beginning of the 20th century (Lindstedt, 2010, p. 2).¹⁶¹

Speaking of *denaskuloj* is a good opportunity to illustrate that Esperanto studies often require specific methodologies that are not common in traditional linguistic research literature.¹⁶² Because Esperanto has no native speakers in the traditional sense (D. Blanke, 1998a, p. 61) but only a minority of *denaskuloj*, the traditional methodologies of descriptive synchronic linguistics that are based on the extralinguistic concept of *native speaker* (e.g., for discovering the sentence-grammar of a language) cannot be employed in Esperanto studies (Miner, 2011). This can be seen both as an obstacle and as an opportunity: an obstacle if studies are abandoned because of the nonapplicability of existing methods or, in fact, a considerable advantage if the challenge is taken up, because it allows for new, creative methodologies to develop. For solving the lack of native speakers proper, Cramer (2016, see also 2018), for instance, proposed a recursive fact-finding approach to help find which Esperanto speakers can serve as reliable sources for determining the grammaticality of the language. His suggestion is innovative from a methodological viewpoint. Thus, the Esperanto speech community is an interesting challenge for linguistics that, as a special case, can open the door to new, as yet unexplored territories. Esperanto is “a unique model for monitoring and exploring many ideas in general linguistics” (Duličenko, 1997, p. 67).¹⁶³

Every Esperanto speaker is at least bilingual: There is no known instance of children speaking Esperanto as their sole language (Versteegh, 1993, p. 541). Unlike native speakers of ethnic languages, *denaskuloj* do not have a special status in the community. It “is not intuition or linguistic imitation, but knowledge of linguistic rules that forms the criterion of language use, which makes productive and creative use possible and gives speakers of Esperanto a high degree of security and self-confidence” (Fiedler, 2012, p. 74).

161 For more details on *denaskuloj* see Fiedler (2012).

162 The phenomenon of *denaskuloj* is also relevant for theories about language acquisition (Versteegh, 1993, p. 540). Furthermore, as Blanke notes (1998a, pp. 59–60), there are in fact a legion of reasons to be interested in planned languages, among which: 1. an ideal-based motivation: the elaboration of an ideal of a neutral, universal planned language, 2. a pragmatic motivation: the use of an existing functioning planned language in practice, and 3. a scientific motivation: the study of an existing functioning planned language as a sociolinguistic experiment in laboratory conditions.

163 In particular, Duličenko suggests that the socialization of Esperanto is a model for linguistic creation through collective action (1997, p. 68).

As mentioned above, both a speech community and a movement have developed from the planned language Esperanto. Historically, the term *Esperantist* was meant to be used to mean *Esperanto speaker*. In the “Declaration About the Essence of Esperantism” or “Declaration of Boulogne,” which was endorsed by the participants of the first Esperanto World Congress, the concept of *Esperantisto* (**Esperantist**) was agreed upon:

An Esperantist is a person who knows and uses the language Esperanto with complete exactness, for whatever aim he uses it for. Membership in an active Esperantist social circle or organization is recommended for all Esperantists, but is not obligatory. (“Declaration About the Essence of Esperantism,” 1905, as translated in “Declaration of Boulogne,” 2014)

The term *Esperantist* was consciously chosen in widespread agreement to mean exclusively *Esperanto speaker*, independent of one’s degree of involvement in the Esperanto movement. Nowadays, however, *Esperantist* can refer to numerous concepts, both in the mind of Esperanto speakers and individuals who are not familiar with the Esperanto speech community. Puškar (2015)¹⁶⁴ empirically demonstrated that Esperanto speakers associate the term *Esperantist* with a large variety of possible definitions: an Esperanto speaker, an Esperanto speaker with the ideology of a world language, an Esperanto speaker connected to the Esperanto movement, an Esperanto speaker who actually uses the language, a cosmopolitan, an Esperanto speaker embracing the ideal of fair communication, a language freak, an Esperanto speaker with multicultural traits, or an Esperanto speaker adhering to the movement and its ideology.

Evidently, the term *Esperantist* in people’s minds may mean much more than just a speaker of the language. Furthermore, what is particularly apparent in Puškar’s (2015) categories are the two distinct components: a language component (Esperanto speaker, language freak, etc.) and an ideological component (a cosmopolitan, someone embracing the ideal of fair communication, etc.). In line with these two aspects, Gobbo (1997) proposed a two-dimensional scale of Esperantism: on the one hand an individual’s ideological involvement and on the other, the individual’s linguistic competence for Esperanto (pp. 53–55).

164 Puškar (2015) conducted a questionnaire (n=108) among Croatian Esperantists. He asked respondents the following open question: “Who would be an Esperantist according to your definition?”

		ideological involvement			
		1. Sympathizer	2. Borderline	3. Attendee	4. Activist
linguistic competence	A. Beginner	+	+	+	+
	B. Skillful	{}	+	+	+
	C. Fluent	{}	{}	+	+
	D. From birth	{}	+	+	+

Table 5. General typology of Esperantism adapted from Gobbo (1997, pp. 53–55: the sign “{}” stands here for combinations that do not exist).

The first dimension, ideological involvement, is a continuum from the mere sympathizer to the activist.^{lvii} Sympathizers are not Esperantists per se because they are not involved with the language or with its ideology. They simply personally know one or more Esperantists. At the other end of the scale, activists are personally involved in organizations of the Esperanto movement and usually edit publications or organize congresses (Gobbo, 1997).

The second dimension, linguistic competence, comprises the following stages: the beginner, the skillful, the fluent one, and the speaker from birth (*denaskuloj*).^{lviii} The beginners can read and write simple texts but have difficulties speaking. The skillful ones have decent reading and writing abilities and can make themselves understood when speaking, but they always think in the native language before speaking. The fluent ones can read any texts and think directly in Esperanto before speaking. Their style is largely independent of the mother tongue.

Finally, the native speakers have learned the language in their family and are at least bilingual (knowledge of at least one ethnic language; Gobbo, 1997).

What is of interest here is that the degree of activism is not a direct function of the degree of proficiency in the language. Although Esperanto would probably have disappeared if it did not have a social movement to promote it, that does not mean that the community and the movement are not distinct notions (Lindstedt, 2010, p. 5). For instance, Hungarian students who chose Esperanto as a third option (university minor) do not feel a special relation to the language or to its community (Fiedler, 1998, p. 28). There are probably numerous Esperanto speakers who do not take part in the Esperanto movement:

The number of Esperanto organizations is only the “tip of the iceberg”: There are many more speakers of Esperanto who are outside the movement. That some such speakers will exist cannot be denied. Not all or even most of those who take Esperanto courses or teach themselves Esperanto through generally

available textbooks end up joining the movement. There are also many competent speakers who lapse in their membership. (Forster, 1982, p. 17)

If this was true in 1982, it should be all the more true in our current era of online technologies, with large online learning platforms and applications offering Esperanto courses and learning materials (Duolingo, Drops, etc.). Thus, it seems essential to distinguish the two concepts from a terminological point of view. Following Gobbo (2009a, p. 107), in the present investigation I use **Esperanto speaker**^{lix} to refer to an individual's language proficiency per se (someone capable, to a certain degree, of speaking the Esperanto language) and **Esperantist**^{lx} if I wish to say something about the degree of involvement of this individual in the community or in support of the language (someone involved, to a certain degree, in the activities linked to the Esperanto movement).

4.2.5 Counting Esperanto speakers

The Esperanto speech community presents several of the typical features of **hard-to-reach populations** (as presented in Marpsat & Razafindratsima, 2010, p. 4): The population has relatively low numbers (compared to the total world population), members of the population are hard to locate because they blend with the rest of the local population given that the community is by nature nonterritorial (Wood, 1979, p. 434), and there is no adequate sampling frame. Existing qualitative and quantitative studies on the matter should be taken with a grain of salt. Piron (1989) suggested, for instance, that some aspects of the community are overlooked by traditional research methods (p. 171). He mentioned that it “is not impossible, for instance, that housewives and gainfully employed women frequently using Esperanto are more numerous than the sources suggest” (Piron, 1989, p. 171). Gledhill (2000) also stated that “the figures do not take into account local activists who are not members of national associations” (p. 10). This suggests that some studies might have restrictively been counting Esperanto-speaking Esperantists rather than Esperanto speakers.¹⁶⁵

Quantitative estimates of Esperanto speakers range from a few thousand to several million. Lindstedt (2010) estimated that there are some 10,000 fluent speakers and some 100,000 active users of the language (p. 2), Vendelbo Nielsen (2016) estimated that some 63,000 individuals around the world would “answer

.....
165 I.e. counting individuals in categories A3 through D4 in Gobbo's typology (see Table 5, p. 131), failing to capture individuals in the categories A1, A2, B2 and D2.

Esperanto when asked about spoken languages by the authorities”, and Pool and Grofman (1989) mentioned the figure of 500,000 speakers (p. 146), Wandel (2015) 2,000,000, and Piron (1989) 3,500,000 (p. 157). Unfortunately, I think we must conclude with other authors that there exist no reliable statistics on Esperanto speakers, that the number of speakers may largely depend on the degree of proficiency chosen as a criterion (Lindstedt, 2010, p. 2), and that estimates have often been arbitrary (Sherwood, 1982, p. 183). Actually, counting Esperanto speakers is not much different than giving figures of L2 speakers of English, for example:

So, if you are highly conscious of international standards, or wish to keep the figures for world English down, you will opt for a total of around 700 million, in the mid-1980s. If you go to the opposite extreme, and allow in any systematic awareness, whether in speaking, listening, reading or writing, you could easily persuade yourself of the reasonableness of 2 billion. (Crystal, 1985, p. 9, as cited in 2008, p. 3)

In the figures mentioned by Crystal, L2 speakers of English range from 700,000,000 to 2,000,000,000, an increase of about 186% from one figure to the other. Needless to say, this constitutes a substantial difference.

There are palpable language data from censuses in some geographical areas, both for English (e.g., in the United States; Shin & Rosalind, 2000) and for Esperanto (e.g., in Hungary; Hungarian Central Statistical Office, 2011), but censuses take place in isolated regions of the world, not on the scale of the planet. For an international L2, furthermore, the problem also comes down to defining the concept of an L2 speaker. This methodological difficulty is not specific to Esperanto but applies to any language used as an L2 (English as a foreign language, etc.). Existing estimates seem to mostly fail to provide the necessary theoretical framework.

How much language must someone have in order to be recognized as a speaker of that language? For Esperanto and its worldwide community there are no censuses, no school systems; there is no geography. And as a second language for virtually all its speakers, it is spoken imperfectly by many, less imperfectly by a few. (Tonkin, 2015, p. 184)

In addition to language census data, some certain figures can be cited, such as the number of individuals using the Esperanto language on Facebook (320,000 in

2014; see Wandel, 2015, p. 319) or the number of individuals currently learning the language on an online platform like Duolingo (more than 1,200,000 as I write these lines; see von Ahn & Hacker, 2017), but they neither are directly related to the actual L2 competence of the individuals nor constitute a significant representation of the total number of speakers throughout the world.

4.2.6 The status of Esperanto in society and science

Esperanto is not recognized as an official language by any state, but large international organizations such as the former League of Nations, the United Nations, and UNESCO have acknowledged its existence (see Fettes, 1997).

Often, Esperanto does not have a good reputation among people who are neither Esperanto speakers nor Esperantists. Piron (1994) noted that Esperanto triggers psychological reactions that are illogical but typical of human behavior and that misinformation about the language has been reproducing itself for decades. The mass media might be reinforcing this trend. Through a corpus study of selected newspapers, Gubbins (1997) showed that a fair share of references to the language in the press carry a negative connotation.

But the Esperanto movement has to take some responsibility for its somewhat negative image among the general public. Zaki (2015) suggested that the internal divisions within the movement have a considerable impact on Esperanto's image and urges Esperanto organizations to work on a common communication policy (p. 94). Von Wunsch-Rolshoven (2013) noted that the Esperanto movement has rarely applied marketing methods and professional advertising methods to support the diffusion of the language (p. 88).

Esperanto has also been largely disregarded by scholars. Kimura (2012) suggested that Esperanto shares characteristics with some **minority languages**,¹⁶⁶ and like other minority languages, it might be subject to prejudices. He mentioned, for instance, Edwards and his view that minority languages like Esperanto have been neglected by sociolinguistics:

The relatively short history of Esperanto itself should not deceive us into thinking that constructed languages per se are a recent phenomenon. On the contrary, they too have a very lengthy historical pedigree and so, like Irish and Gaelic, Esperanto—the study of which has been unjustifiably neglected

166 Minority languages: “languages dominated by larger languages in a given context, usually a state” (Kimura, 2012, p. 168).

by sociologists of language—can also tell us something of the forces bearing upon maintenance and shift, revival and loss. (Edwards, 2010, p. 3)

Liu (2001) shared a similar point of view:

Planned languages are often ignored by orthodox linguists because of their plannedness (artificiality) and unnaturalness; in the same way, pidgin and creole languages also have a similar fate in the eyes of linguists due to their imperfections and incompleteness. (p. 161)

But the limited presence of Esperanto studies in mainstream linguistics might also be attributable to several methodological difficulties rather than prejudice.¹⁶⁷ The scholarly literature on Esperanto is not easily accessible. Library collections on the topic are poor (Tonkin & Fettes, 1996, p. 5), but scholars also do not necessarily have the language knowledge to explore the field. According to D. Blanke (2004), 60–70% of the literature about planned languages has been written in planned languages. More recently, Fiedler (2015) also noted that out of the 210 Esperanto and interlinguistics studies referenced by Tonkin for the year 2006, 72.9% were published in Esperanto, 10% in German, 5.7% in Russian, 5.2% in English, 2.9% in French, and 3.3% in other languages (p. 99). About two thirds of the literature about the Esperanto language can be expected to be published in Esperanto, which ironically poses another problem: The profound knowledge of the Esperanto language and membership in its speech community may disqualify the researcher (Fiedler, 2015, p. 99), which Lindstedt (2010) called a “special kind of ‘observer’s paradox’” (p. 5). This seems true especially if the Esperanto speaker actively takes part in the Esperanto movement.

From a purely rational viewpoint, the potential advantages of the Esperanto language have been noted (see also Fiedler, 2015): for instance, Grin (2005) mentioned Esperanto as the best-case scenario for Europe from an economic perspective (yearly net savings of some 25,000,000,000 euros; p. 7), and several studies have underlined the potential of Esperanto for facilitating the learning of foreign languages other than Esperanto and for increasing students’ metalinguistic awareness about their own language (see Commissione Sulla Lingua Internazionale [Detta Esperanto], 1995).

To summarize, the study of Esperanto has a long history but only recently has it entered mainstream research (Tonkin, 2007, p. 169). **Esperantology**, or

.....
167 There are also other (political) reasons that I will not mention here.

Esperanto studies, is the branch of interlinguistics that studies the sources, construction principles, development, functions, domains of use, communicative aspects, speech community, and history of the language Esperanto founded in 1887 by Zamenhof (D. Blanke, 2006, p. 34). It comprises both descriptive and prescriptive components (D. Blanke, 1998b, p. 21). Tonkin and Fettes (1996) and more recently Fiedler (2015) gave an overview of Esperanto studies.

4.3 Lexical change in Esperanto

4.3.1 Democratically completing the plan

The Esperanto language belongs to its speakers:

But it is not Zamenhof [the initiator of Esperanto] who further develops Esperanto, and as long as we rely on Zamenhof's writing, we will not be equipped to study the development of Esperanto (Lo Jacomo, 1981, p. 341, my translation)^{ki}

Dasgupta (1993) said that the Esperanto speech community is an “experiment in communitarian, nonauthoritarian language planning” (p. 369). Esperanto speakers have a tradition of lexical collaboration. As previously mentioned, for Esperanto, it all started from an incomplete plan, the *Unua Libro* (*First Book*). As Welger (1999; see also Schubert, 2010) pointed out, Esperanto is only “partly planned.” What was planned in the language was especially its initial state and basic rules for development. It follows that the formal norms of Esperanto allow for need or completion; many of them are not strict rules but rather pieces of advice, and specific organs are entitled to add to the original language norms.

This makes Esperanto particularly relevant for lexical research because the language can be studied from its very birth; the language situation is particularly interesting for any scholar interested in phenomena of diachronic linguistics. A picture of time zero is available, and later variations of the language can thus be compared to a relatively sharp snapshot of the language plan as it was launched in 1887. The first moments of the language have been captured by various authors (Privat, 1912; Waringhien, 1980). Esperanto can be seen as an

extraordinary linguistic laboratory, because a lot of the fundamental variables of the language—its birth date, the basic vocabulary, the number of

speakers at a certain point in time—may not be certain but can at least be better approached in comparison to other historically natural languages. (Gobbo, 2009b, p. 85, my translation)^{lxiii}

The fact that the language project was open to further development was one of the factors that led to the relative success of Esperanto (D. Blanke, 2009, p. 256). From the start, it was clear that Esperanto would belong to its speakers and its speech community (Fettes, 2013). Its initiator, Zamenhof, never considered the Esperanto language his property and tried to set the language free (Gobbo, 2017, p. 3). He did not present a full grammar guide but rather a general framework, and he invited the general public to contribute and turn the parts into a living language (Manders, 1976, pp. 236–237). Schubert noted that “Zamenhof’s general principle [was] to give information about the [language] system, not about its use” (2010, p. 360, my translation).^{lxiiii} This, however, does not mean that language change was necessarily expected to be random:

Was Zamenhof only the builder of the language system, did he leave language usage to chance? He did not. He put in his language system the potential for evolution and established a set direction for its ability to develop. This direction is implicitly hidden in the agglutinative word formation mechanisms and in the unprecise relation with ethnolinguistic models. (Schubert, 2010, p. 361, my translation)^{lxiv}

Zamenhof’s original language plan was collectively deployed and expanded by himself and his followers:

Once the scheme was published, the language had to go through a further period of development, during which it was learnt by a few early followers and then advertised to the wider community (corresponding to Haugen’s process of “acceptance”) at the time as being used for increasingly diverse forms of communication (the process of “elaboration”). (Gledhill, 2014, p. 326)

Lexical structures were not clearly defined in the early stages of the language. Grammar was first described by René de Saussure¹⁶⁸ in response to criticisms about Esperanto about 20 years after Zamenhof’s language was made

.....
168 The brother of the linguist Ferninand de Saussure.

public (R. de Saussure, 1908, 1910, 1985). Throughout the history of the language, several theories describing Esperanto's grammar and formation processes were formulated (e.g., Kalocsay & Waringhien, 1980; Kiselman, 1991; Malovec, 1999; Miner, 2006; Schubert, 1989, 2015). Unfortunately, some aspects of Esperanto grammar are not easily accessible for linguists who are not Esperantologists because grammar authors did not always follow the main tendencies of modern linguistics for structuring and presenting their work, and Esperanto grammar in languages other than Esperanto is partly lacking (Brosch, 2008, p. 5). There are noticeable exceptions, such as Willkommen's grammar (2007) or Gledhill corpus-based grammar (2000).¹⁶⁹

The *neologismo* is a straightforward example of a false friend to traditional linguistics. Although Esperanto's lexical formation processes can be described using widely accepted terms from the field of linguistics (Brosch, 2008, p. 38), in the Esperanto speech community, a **neologismo** (neologism) is most commonly understood exclusively as a new lexical item created under the influence of another language¹⁷⁰ (i.e., by borrowing on the basis of a foreign model or **loan creation**). In comparison, traditionally in linguistics and in the discipline of terminology, a **neologism** is a new form, a new meaning that has arisen recently in the language¹⁷¹ and may or may not be a borrowing. An example to illustrate the difference between these two concepts: When blog technologies appeared a few years ago, several synonyms developed in Esperanto, including "blogo" and "retjurnalo." To Esperanto speakers, only "blogo" is a neologismo because it is formed from the root "blog-" using lexical material from a foreign language. "Retjurnalo" is not considered a neologismo because it is formed from existing Esperanto roots: *ret-* (online), *ĵurnal-* (journal) and the ending *-o* for nouns. To mainstream linguists, both "blogo" and "retjurnalo" are neologisms, as they are both new forms in the language. This is perhaps best clarified with an illustration; in Figure 7, I have adapted a Munske's classification. "Blogo" and "retjurnalo" would both belong to formal neology in this figure, the former being created based on loan rendition and the latter on indigenous word formation.

169 And studies on particular aspects of grammar, concerning the lexicon e.g. Blanke's description of word formation from a comparative perspective German-Esperanto (1981, pp. 60–72) and more recently Werneck Dias' thesis (2007) on lexical formation mechanisms in Esperanto.

170 See Sablayrolle *et al*'s French article about "création 'sous influence'" (2011).

171 This is a broad simplification with two main categories, but it is largely accepted. On the classification of neologisms see Sablayrolles (1996).

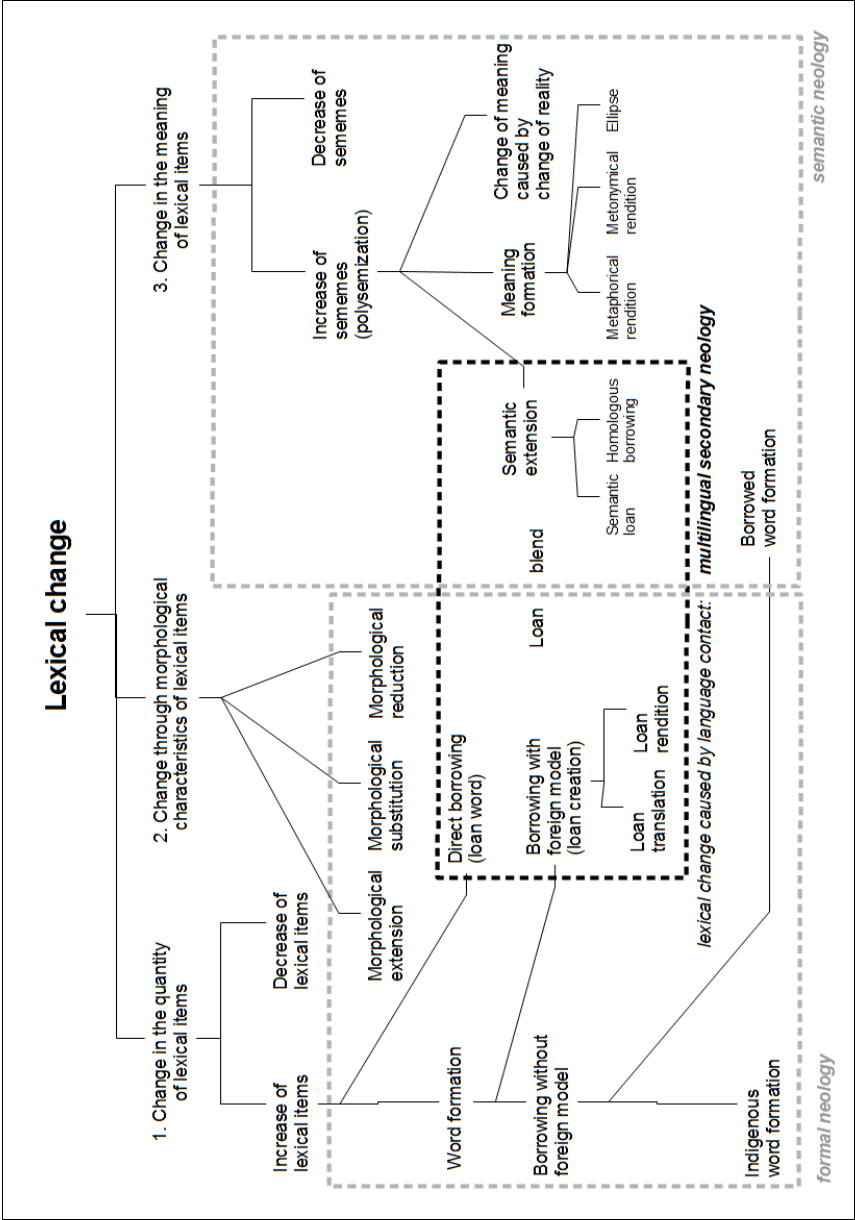


Figure 7. Typology of lexical change adapted from Munske (2015, p. 26).

The Esperanto lexicon might have been the least complete part of the plan. According to Lapenna (1974), historically, the most prominent change in the history of the Esperanto language was the growth of word roots (p. 42).^{172,173} The first Esperanto dictionary, elaborated by Zamenhof himself, contained only 904 lemmas (Gobbo, 2017, p. 2; Werneck Dias, 2007, p. 20). Just a few years later, in 1894, the dictionary *Universala Vortaro* was published, containing twice as many lemmas (Gobbo, 2017, p. 3).

1905 was an important year for Esperanto speakers; the foundations of the language, the *Fundamento* (Zamenhof, 1905),¹⁷⁴ were published, and an official language body, the *Lingva Komitato*, was founded (Gobbo, 2009b, p. 86). According to Brosch (2008, pp. 21–22), four word formation principles appeared implicitly in the *Fundamento*:

1. Lexical items must be transparent (*encentreco*).
2. Lexical items must fill a lexical gap (*plenumo de vortfarada sfero*).
3. Lexical items can have only one categorical ending (*kategorieco*).
4. Complex lexical items must be reversible (*principo de renverseblo*).

The missing but needed vocabulary was developed by members of the speech community through interpretations of the basic rules of the language and occasional collective discussions with other speakers and Zamenhof himself. At first, the community was relatively small, and the first magazine—*La Esperantisto*—was used for discussions about the language somewhat similarly to the way a forum or wiki would be used today (Gobbo, 2009b, p. 85).

Esperanto speakers have been cultivating their language since its early days. This is probably due to sociological motives; the members of the speech community know that the overall social, psychological, and axiological structure of the community would be endangered if language abilities were not fostered (Rašić, 1994, p. 165). Esperanto speakers tend to show significant metalinguistic

172 Esperanto being a schematic language (see Section 4.2.1), lexical items can be created by combining roots. The total number of lexical items in the language, therefore, is considerably bigger than the number of roots.

173 In selected dictionaries, from around 931 elements in 1887, to 2600 in 1893, 7866 in 1954, 16,000 in 1970 and 17,000 in 2002 (Fiedler, 2006, p. 81; Lapenna, 1974, p. 42). Needless to say the count depends on the dictionaries' editorial policy, but it seems safe to state that there is a clear tendency of lexical expansion.

174 For a detailed description of the *Fundamento*, see e.g. Pabst (2014).

awareness and often have clear opinions on language matters.¹⁷⁵ Moreover, Esperanto speakers often explicitly express their opinions regarding the lexicon. The *neologisma diskuto* (debate on *neologismoj*),¹⁷⁶ for instance, is as old as the language (Schweder, 1999, p. 45).¹⁷⁷ A century after the language was launched, publications discussing *neologismoj* continue to appear regularly (Camacho, 1999; Kadoja, 2013; Mayer, 1987).

For almost the entire ninety-year lifetime of the language, there has been a public debate, often acrimonious, about control of growth of the lexicon. Roughly speaking, the disagreements result from a desire on the one hand to keep the number of roots small to benefit the new learners and on the other a need felt by writers, especially poets, to enrich the vocabulary for literary purposes. While the debate has typically been conducted along the dimension utilitarian/literary, the arguments are often reminiscent of the question of purism in many national languages, which in Esperanto usually takes the form of contrasting “homey” compounds created from the internal, autonomous resources of the language. (Sherwood, 1982, p. 185)

The *neologisma diskuto* is not only a question of purism or understandability of lexical items. In fact, it questions the very autonomy of the lexical formation system of Esperanto (Lindstedt, 1983).¹⁷⁸ The fact that Esperanto speakers are meta-

175 A sociological study conducted in the 1990’s (Rašić, 1994, p. 97–185) can serve to clearly illustrate this point. Within the framework of this study, around 170 individuals responded to a questionnaire that comprised four questions related to the language norm (opinion on the use of diacritics, the rules for spelling proper nouns, the creation of new words and the eurocentrality of the vocabulary). The interesting point here is not the actual answers given to the questions, but rather the low percentage of respondents that chose the answer “I did not think about this” (respectively 3,85%, 3,85%, 7,69% and 10,26%). This seems to imply a very low percentage of Esperanto speakers—perhaps less than 10%—do not hold a personal opinion on specific language matters.

176 I.e. mostly on roots borrowed from other languages, see the definition of *neologismo* above.

177 Many Esperanto speakers try and resist *neologismoj*. Waringhien summarized them in three categories (1980, p. 287): 1. old Esperanto speakers who feel offended if they encounter a new lexical item, 2. promoters of the language and language teachers, who advocate for a simple grammar and a restricted vocabulary, and 3. individuals who do not agree that an international language should be used for literature and claim Esperanto should remain some kind of Basic English.

178 Political arguments are also at stake here: some Esperanto speakers try to show that Esperanto can develop its lexicon on the basis of its internal resources in order to

linguistically aware and that they express their opinions on language matters is a great opportunity for linguists interested in better understanding lexical criteria: Esperanto speakers often discuss *neologismoj*, and more generally lexical items at large, and as mentioned, they often do so online, offering researchers a large sample of text materials that can be studied covertly.

Some Esperanto speakers are very active (folk) lexicographers through private initiatives or group work. Early on, Esperanto speakers began actively tracking Zamenhof neologisms that had not been included in the official dictionaries. During the 1907 World Congress, a commission led by Wackrill was given the task of compiling a corpus of the lexical items Zamenhof had been using that were not in the *Universala Vortaro*, and in parallel, Boulet (as an isolated researcher) compiled a similar list (Waringhien, 1980, p. 166). The *Lingva Komitato* then democratically decided which of the listed neologisms would be added to the official dictionary. Leading the *Lingva Komitato*'s section on the common lexicon, Cart sent the two lists of Zamenhof neologisms to the members of his section and had them cross out all the roots that were not justified by language usage or that went against the principles of the *Fundamento* (Waringhien, 1980, p. 167). Eventually, the amendment to the official dictionary contained only those roots that had been *unanimously* accepted by all members of the *Lingva Komitato*'s section (Waringhien, 1980, pp. 166–168).¹⁷⁹

The *Lingva Komitato*, which later became the *Akademio*, continued publishing amendments to the *Universala Vortaro* and remained relatively conservative. From a quantitative perspective, its additions to the official dictionary were not particularly numerous over the years.¹⁸⁰ The *Akademio* itself does not create new lexical items (Bormann, 1999, p. 38). In fact, the *Akademio* sometimes used its

prove it is a fully functioning, independent language, and simply because secondary neologisms are needed if one wishes to discuss scientific topics in Esperanto. Along with literary language, often arguments have thus been given that neologismoj are needed in languages for special purposes because scientific terminology is in essence mostly international and to the domain specialist it would be the indigenous lexical item that would seem 'foreign'. This opinion was expressed for example by the Danish agronomist Paul Neergaard (1955).

179 Although the original lists of Zamenhof neologisms contained respectively 843 and 2021 new lexical items, after the review round the addition to the official dictionary comprised only 864 new roots.

180 Based on the data in Wennergren's article (2008), a mere 2159 roots were officially added between 1909 and 2007.

resources to fight the intuitive solutions of Esperanto speakers (Corsetti, 1999, p. 57). In some cases, it capitulated after fierce combat.¹⁸¹

There is an Academy of Esperanto, but it has historically played a very minor role in the development of the language. Even in lexical matters the Academy has limited itself to occasional listings of words which have been around for enough decades to seem “official.” Major growth in the lexicon has occurred through decentralized individual suggestions and use. (Sherwood, 1982, p. 187)

Despite its marginal role in lexical development, the *Akademio* should not be neglected today because it remains an absolute lexical reference for some Esperanto speakers, a prestigious source that must be obeyed without question, as the following anonymized speaker illustrates:¹⁸²

I accept the decision of the Akademio, which with the BRO [official basic set of roots] changed the category of trondr-. Personally, I doubt that they did it by means of a formal decision about each one of the modified roots, but we must accept the result. [Translated example from my corpus]^{lxv}

However, the *Akademio* followed too conservative a path and failed to respond to all of Esperanto speakers' needs, often losing its leading role as regards the lexicon in favor of other resources (see Gobbo, 2009b, p. 88). From the early years of the language, unofficial lexical resources were published in parallel to official ones: in general language, for instance, the *Plena Vortaro Rusa-Internacia* in 1889 (see Lapenna, 1974, p. 42), Boirac's dictionary in 1909, and Kabe's in 1911 (see Gobbo, 2017, p. 4). Later, the 1930 monolingual dictionary *Plena Vortaro de Esperanto* (Grosjean-Maupin, Esselin, Grenkamp-Kornfeld, & Waringhien, 1930) in particular stole the lexical limelight.¹⁸³ Although published by an unofficial source

181 Fiedler gives the example of verbalization of adjectives, which was first rejected by the *Akademio* but later made official, when it became ever more popular among Esperanto speakers (2006, p. 80).

182 Generally, Philippe states that an overwhelming majority of Esperanto speakers show a strong spirit of resistance against any kind of change in the language (see 1991, pp. 90–91), which partly explains why the conservative position of the *Akademio* still finds its share of supporters.

183 Today it is the 'Plena Ilustrita Vortaro de Esperanto' and is also available online (www.vortaro.net). It is a monolingual, comprehensive dictionary of Esperanto. It is

(the World Anational Association), it “became the de facto monolingual standard dictionary of Esperanto for decades” (Gobbo, 2017, p. 5). Lexical resources can also emerge from private initiatives; for instance, Cherpillod listed lexical items that were absent from the *Plena Vortaro de Esperanto* (1988).

For the first few years, Esperanto was used not only for general language but also for specialized communication. Thus, it needed scientific terms in addition to general language expressions. In 1904, the journal *Internacia Scienca Revuo* (*International Scientific Journal*) began “creating and fixing terms” through the edition of texts, discussions among a large specialized public, and selection and acceptance by a competent commission (W. Blanke, 2013, p. 37). Over the years, principles such as clarity, stability, regularity, and internationality for the coining of new lexical items have been published and disseminated among Esperantists (W. Blanke, 1997, 2013, pp. 99–149; Eichholz, 1995; Neergaard, 1955; Suonuuti, 1998; Verax, 1911; Werner, 1986, 2004). In the first half of the 20th century, there were very close links between the development of Esperanto and that of the discipline of terminology (W. Blanke, 2008a, pp. 62–84; Slodzian, 2005).¹⁸⁴ Many Esperanto speakers are well aware of principles for coining new lexical items and have a consistent record of discussing and negotiating lexical items:

Due to the manifold influences of speakers’ ethnic languages and cultures as well as the ambition to respect all people’s interests and feelings, permanent negotiation regarding meanings and uses of lexical items seems necessary in Esperanto ... similar negotiations can also be observed for everyday objects.

a desk dictionary that aims to give a broad picture of the lexical items which are in general use. It is rather prescriptive in nature and tends to consider exclusively items that are deemed to be part of the standard. PIV is chiefly synchronic. It does not in any manner portray the historical development of the lexical items of Esperanto, neither does it describe the morphological or semantic changes these units have undergone since the creation of the planned language. It contains many terms which have been coined and/or defined by domain specialists. It is, however, not a specialized dictionary per se since these terms have been defined and presented like any other unit of the language—apart from a domain label.

184 There is also an article from Samain on the relationships between Wüster, Esperanto and the development of terminology (2010), which I would recommend reading with a critical eye. I express some reservations about it, because of the appalling spelling mistakes the author makes, e.g. (p. 281) *Enciclopedia Vortaro* instead of *Enciklopedia Vortaro*, *Maŝinfaka Esperanta Vortaro* instead of *Maŝinfaka Esperanto-Vortaro*, *lingua boneco* (p. 286) instead of *lingva boneco*. Also, Samain mentions (p. 282) that a terminological commission was initiated by *Louis de Saussure*. Did he mean *René de Saussure*?

Here, Esperanto speakers constantly seem to fear that names are chosen which are too close to their own language. (Fiedler, 2002, pp. 69–70)

The speech community has been consciously and collectively completing the original implicit principles for lexical development. Esperanto speakers are usually accustomed to creating new lexical items according to a plan and probably hold a much stronger metalinguistic awareness than speakers of other languages:

Those who need terminology for their professional work, and who deal with it in technical committees, are mainly technicians, engineers. It can happen that they know one, two or even several foreign languages and understand each other without difficulties. But on the meta-level, concerning theoretical problems of linguistics or even interlinguistics, they are generally not competent. They are accustomed to accept words as they are. The idea that word-building can happen according to a plan, consciously, following previously established rules, is mainly foreign to them and (as history shows) only after ten or twenty years of unsatisfying work did they feel the need for “codified principles.”

For users of planned languages, contrary to what can be said about engineers using national languages, this idea is quite natural. Esperantists (at least, those who have seriously concerned themselves for some years with the language, its properties and rules, who fluently speak and think in the international idiom) are fully accustomed to “word-building” as one means of expressing one’s thoughts. For them it is self-evident that this free word-building requires guiding principles, to guarantee successful communication. (W. Blanke, 2008a, pp. 32–33)

A great part of activities carried out in the Esperanto speech community are voluntary (and, for the most part, unprofessional), which has an influence on the lexicography of the language (D. Blanke, 2006, p. 225). Due to the lack of governmental support, speech community members can only rely on themselves (Sakaguchi, 1998, p. 296). Among the many voluntary activities that Esperanto speakers undertake, there are several networks of practice concerned with the lexicon. Many language resources have been produced through a large-scale collaboration (Schweder, 1999) and thus are the result of **folk lexicography** (i.e., the creation of dictionaries by ordinary speakers; Šumarina, 2014, p. 293). Esperanto speakers tend to engage voluntarily to address lexical needs and solve one

another's lexical issues. Esperanto networks concerning the lexicon thus usually engage a fair number of speakers who collaborate as equals.¹⁸⁵

Again, this is an opportunity for the researcher to study, in a covert way, how speakers discuss and select lexical items to retain (lexical criteria), because Esperanto speakers tend to use explicit metalanguage to express their opinions of lexical items. Speakers of Esperanto took advantage of the development of the Web and the collaboration possibilities offered by the Web 2.0 (Žadziško, 2011). Over the past 20 years, several groups of Esperanto speakers have been using Internet technologies to carry out lexicographical tasks, particularly to discuss lexicographical contents and lexical items. These lexical collaboration activities, particularly the online networks discussing lexical questions, are extremely rich material for researchers interested in speakers' lexical criteria, because in group settings, members often explicitly explain their lexical choices to their peers.

Teamwork and democracy are generally important in the community. To prevent the influence of one's native language and personal subjective views, it is commonly accepted that the development of (specialized) vocabularies must be a collective action:

It is important to find a solution for the composition of the collective. Its leader should be a person with domain and linguistic knowledge, and their character should be compatible with the precondition of democratic treatment within the collective. Members would belong to various ethnic communities with distinct languages. (Werner, 2004, p. 57, my translation)^{lxvi}

There exist several active groups in the Esperanto speech community that are concerned with lexicographical projects. In the history of the language and the community, these principles of collective action and democratic language development have been put into practice on a large international scale. An eloquent example is the *Esperanta Bildvortaro*, an Esperanto "translation" of the 1958 edition of the illustrated Duden dictionary of the German language. This 880-page dictionary was compiled through the collaboration of around 144 domain specialists from 25 countries between 1974 and 1988. The leader of the project, Rüdiger Eichholz, prepared the basis and gradually sent his proposals to other contributors for a free, world-scale discussion. The exchanges took place in writing, using index cards displaying lexical items in English, German, French, and Esperanto with

185 Unlike for instance FranceTerme's wiki, in which the official language body makes decisions while ordinary speakers can only make proposals.

definitions, sources, and often illustrations. Corrections were made on the cards, which were exchanged between the project leader and contributors. A total of 1440 index cards were exchanged in the preparation of the dictionary (see W. Blanke, 2008a, pp. 124–126; Schweder, 1999, pp. 57–58).

Esperanto speakers have also contributed to multilingual lexicographical projects outside of the Esperanto speech community; for instance, in the development of railway terminology (W. Blanke, 2008a, pp. 127–128; Hoffmann, 2001; Schweder, 1999, p. 52). A terminological newsletter is regularly published by the Esperanto Railway Association (IFEFF) in which Esperanto equivalents are proposed (see Internacia Fervojista Esperanto-Federacio [IFEFF], 2015, 2016). Proposals are discussed via this magazine or during congresses. As a result of this collaborative group effort, the RailLexic CD-ROM with 16,000 specialized lexical items is available in 22 languages, including Esperanto.

The current version of the *Plena Ilustrita Vortaro (PIV)*, often seen as a reference dictionary by Esperanto speakers, is also the result of collective action: *PIV2002*, followed by *PIV2005*, are collaborative revisions of the 1970 edition, in which even ordinary speakers participated, as Duc Goninaz mentioned in the foreword to *PIV2002*:

The echo around the revision work encouraged numerous specialists and ordinary users of the dictionary to spontaneously send contributions, suggestions and remarks. (Waringhien, 2005, p. 21, my translation)^{186vii}

In 1987, the World Esperanto Association (UEA) officially founded a terminology center (TEC) during the world congress in Warsaw. The TEC was intended to coordinate discussions about new lexical items in language for specialized purposes.¹⁸⁶ As in other speech communities (and much like the FranceTerm wiki example), the engagement of ordinary speakers (specialists in technical or scientific domains, but not in linguistics) was an important unresolved issue for the TEC:

The three organizers of TEC were and remained alone with their oversized plans: A leader who would coordinate was always missing and missing were

186 The main aims of this center were (W. Blanke, 2008a, pp. 59–60): the study of the international terminology standardization, the organization of international discussions about terminology proposals, the organization of a procedure of standardization, the publication of terminological standards and the representation of Esperanto in bodies concerned with terminology.

especially a large crowd of specialists who would be willing to collaborate and not only take care of their own specific domains, who would agree to take on co-organization within the framework of the planned structures. (W. Blanke, 2013, p. 213, my translation)^{lxviii}

Unfortunately, the TEC office in Zagreb was definitively closed in 1992, mainly because of the Yugoslav wars (W. Blanke, 2008b, pp. 212–214).

With the advances of new technologies, part of the collaborative activities concerned with the lexicon are conducted online in private and public spaces. This is where the researcher can observe Esperanto speakers speak about lexical items and collect their opinions and preferences about the lexicon to gain a better understanding of the lexical processes occurring within the speech community.

4.3.2 Lexical norms in Esperanto

Today, the creation of new lexical items in the Esperanto speech community is generally not guided by official bodies; the community does not depend on a functioning standardizing body or normative dictionaries (Schweder, 1999, p. 69). The *Akademio* does not create new lexical items,¹⁸⁷ and the UEA's TEC has closed.

Zamenhof's *Fundamento* has not always been followed by speakers (Philippe, 1991, p. 75),¹⁸⁸ but despite the geographical repartition of Esperanto speakers and unguided changes, there are no dialects in Esperanto.¹⁸⁹

It is often claimed that under widespread use Esperanto would break up into mutually incomprehensible dialects. This may be an invalid conclusion, since it is based on the way languages developed before there existed rich

.....
187 According to Golden it is not at all the task of the *Akademio* to develop the lexicon (1990, p. 197), and Olganov suggests the *Akademio* is not supposed to obstruct the natural evolution of the language, but should rather ascertain the supremacy of one variant over the other that have developed through *natural* evolution (1985, p. 91).

188 Mattos for instance regrets 'major accidents' (*gravaj akcidentoj*) which took place in the history of the language (1999, p. 33), one of which is the Esperanto grammar *Plena Analiza Gramatiko de Esperanto* (Kalocsay & Waringhien, 1980), which in his view did not apply the *Fundamento* to the letter. However, as Fiedler points out, the "aim of the *Fundamento* was not to prevent any development but rather to protect the language against arbitrariness." (2006, p. 78).

189 The risk that a language will be dialectized exists only if a group of individuals use that language exclusively within a closed group (Lo Jacomo, 1981, p. 345).

modes of global communication. A language breaks up into dialects when there is isolation, and under present conditions of mass literacy, global electronic communication, and mass global travel, isolation is increasingly uncommon. (Sherwood, 1982, p. 192)

Because in Esperanto, communication takes place almost exclusively between individuals with various native languages, the interference of one's mother tongue is counterbalanced by the need to be understood by someone from another linguistic and cultural community (Leyk, 1980, p. 463). Fiedler (2006) explained the stability of Esperanto by a set of "self-regulation techniques" or "stabilizing factors" (pp. 82–83).¹⁹⁰

Phillipe (1991) proposed three categories of factors (p. 80–99) impacting language change in Esperanto:

- Factors favoring language change: bilingualism, geographic and temporal discontinuity, tendency towards wordplay and innovation (German: "Spieltrieb"), need for complexity, literature
- Factors impeding language change: need to be understood by discussion partner, use of written language, explicit knowledge of the language system, ideological objectives
- Linguistic economy and redundancy

Other scholars have attempted to define norm components affecting usage. Summarizing Jansen's (2007, pp. 43–48)¹⁹¹ and Gledhill's (2014, pp. 322–323) recent work and adapting it to the lexicon, the following lexical norm-giving components can be suggested:

- An authoritative component: The *Fundamento* and *Akademio de Esperanto*, as well as authoritative grammars, textbooks, and dictionaries may be perceived as lexical references by speakers (see also the example above).
- A democratic component: The correct lexicon is that used by the majority of the speech community.

190 For the diachronic perspective, the interested reader may also refer to the factors affecting Esperanto language change as discussed by Philippe (1991, pp. 80–99), Sakaguchi (1998, pp. 253–261) and Detlev Blanke (D. Blanke, 2006, pp. 236–238).

191 Jansen builds his norm concepts on Manders' work (1976) and St. Clair's addition (1978).

- A usage component: The lexicon as used in existing written genres (books, magazines, novels, creative texts, etc.), spoken genres (both formal and informal) and hybrid genres (blogs, Wikipedia, etc.) may be norm-giving.
- A geographical component: Lexical development is influenced by Esperanto speakers' mother tongues, which for the most part are European languages.
- A clarity component: The correct lexicon is expected to be clear and transparent.
- An aesthetic component: correctness equates to beauty.

In Chapter 8, regarding lexical criteria, I will describe these language changes and normative components at work in Esperanto speakers' explicit metalinguistic statements based on naturally occurring data (through a corpus), and I will propose a more nuanced classification of speakers' lexical criteria. It will be clear that Esperanto, like other natural languages, is subject to variation and that speakers often disagree on which criteria should have priority.

5 Electronic networks discussing Esperanto lexical items

5.1 Introduction

Chapter 4 set the stage on which I have conducted investigations (i.e., the Esperanto speech community). In Chapter 5, I introduce the specific groups on which the research was conducted within the Esperanto speech community: five electronic networks of practice. These networks are not run by official entities but by dedicated self-administered individuals discussing language-related matters. In Section 5.2, I present these networks in detail. In Section 5.3, I present the results of a survey conducted among a subgroup of contributors to understand the general profile of contributors and, more specifically, their actual linguistic knowledge and language proficiency.

5.2 Five online networks under the microscope

The empirical investigation (starting from Chapter 6) is centered on what I call **electronic networks of practice**, “a self-organizing, open activity system focused on a shared practice that exists through computer-mediated communication” (Landqvist & Teigland, 2005, p. 360). The networks I am exploring in the present work are self-organized groups of speakers that help each other and share perspectives on specific issues (here: language issues) on the web; for instance, through a discussion list or a web interface. Due to the geographic repartition of Esperanto speakers throughout the world and thus the lack of physical proximity, members of these networks are not expected to discuss topics face-to-face. Once a communication topic is set, therefore, all comments can be expected to take place on the web. This is a clear advantage for my research, because the discussion can be observed in its entirety.

In the present investigation, five networks came under the microscope. The choice of the networks was guided by a heterogeneous purposive sampling strategy, meaning that each network was “selected according to predetermined criteria relevant to [my] particular research objective” (Guest, Bunce, & Johnson, 2006, p. 61). Here, these criteria are:

- The network must be concerned, in whole or in part, with the lexicon
- Each of the networks selected must pursue a different objective (heterogeneous sampling)
- The network must be open for participation to any interested individual and participants all have the same participation possibilities¹⁹²
- The network data must be accessible for covert observation (avoidance of the researcher effect) and allow for anonymization (research ethics)

For comparison purposes, to avoid presenting Esperanto research using Esperanto-specific terms or classifications (although general references are available), I present the five networks based on existing scholarly typologies. Table 6 is a combination of a typology of user contributions (Abel & Meyer, 2013, 2016); one concerning the lexical process (Klosa & Tiberius, 2016) and one dealing with Web collaboration systems (Doan, Ramakrishnan, & Halevy, 2011).

Two networks work on solving one-off language-related problems. The other three networks collectively work on producing (or updating) lexical resources. Thus, I will further present these three according to existing lexical resource (dictionary) typologies. Wiegand (2010) and his colleagues distinguished between four types of dictionary typologies (pp. 82–93):

- Typologies based on users (function, situations of use, etc.)
- Typologies based on the subject of the dictionary (e.g., pronunciation dictionary)
- Typologies based on the form of the dictionary (alphabetical sorting, onomasiological dictionary, etc.)
- Typologies concerning characteristics of the storage and publication medium (digital storage, computer-based data model, presentation of entries on digital media, etc.)

However, these existing typologies do not seem best suited to describe (a) the characteristics of lexicographical products on the Web 2.0 as a relatively recent media, (b) the collaborative aspects of lexicographical products developed in a group or a network, and (c) the specificities of lexicographical processes in Web dictionaries.

¹⁹² Networks such as *AdE-diskuto* or *piv-grupo*, for instance, were intentionally excluded.

NAME OF NETWORK	ASTRONOMIA TERMINARO	RETPOSHTA RONDO POR REDAKTANTOJ	ESPERANTO-TRADUKISTOJ	VIVO-VIKIO	LINGVA KONSULTEJO
NATURE OF COLLABORATION	Explicit	Explicit	Explicit	Explicit	Explicit
TARGET PROBLEM	Create specialized dictionary of astronomy and astrophysics	Expand a retrodigitized paper dictionary	Solve punctual lexical issues related to translation	Perform lexicography of current neologisms	Give advice on the structure and use of Esperanto (complicated issues)
TASKS OF USERS	Dictionary preparation Data collection	Data collection Data preparation Data evaluation	Respond to questions, including: Suggesting secondary neologisms	Data collection Data preparation	Respond to questions
COMBINATION OF USER INPUTS	Through discussions	Through discussions	Through discussions	Through discussions	Through discussions Mini-surveys Likes
EVALUATION OF USER INPUTS	Through discussions	Through discussions	Through discussions	Through discussions	Through discussions Mini-surveys Likes
DEGREE OF MANUAL EFFORT	From little to substantial, depending on the task undertaken Content provider Perspective provider	From little to substantial, depending on the task undertaken Content provider Perspective provider	From little to substantial, depending on the task undertaken Content provider Perspective provider	From little to substantial, depending on the task undertaken Content provider Perspective provider	From little to substantial, depending on the task undertaken Content provider Perspective provider
ROLE OF HUMAN USERS	Contributions to open-collaborative dictionaries Stand-alone (Yahoo group)	Contributions to open-collaborative dictionaries Stand-alone (Yahoo group)	– Stand-alone (Yahoo group)	Contributions to open-collaborative dictionaries Stand-alone (Wiki)	– Stand-alone (Facebook)
ARCHITECTURE					

Table 6. Type of collaboration in the five electronic networks of practice. Classification based on Abel and Meyer (2013; roles of contributors as human users); Doan, Ramakrishnan, and Halevy (2011); and Klosa and Tiberius (2016; tasks of users in the lexical process).

A few remarks are in order here. Some scholars claim that lexicographical work in collaborative settings or the participation of ordinary speakers in lexicographical activities is a completely new phenomenon. Murano (2014), for instance, suggested that thanks to the revolution brought about by IT technologies, non-lexicographers can henceforth elaborate dictionaries in collaborative settings (p. 148). Meyer and Gurevych (2012) stated that “collaborative lexicography is a fundamentally new paradigm for compiling lexicons” (p. 259). This does not hold true for the Esperanto speech community; collaborative lexicography existed long before computers, as clearly illustrated in the previous chapter.¹⁹³ Another fallacy in the scientific literature is that the rise of collaborative dictionaries has *recently* started to drive language change (Creese, 2013, p. 392). In the Esperanto speech community, this is also a long-known phenomenon, at least in specialized domains:

Because of the way in which specialized lexical items are created in Esperanto (see Chapter 5.1), in most cases one cannot unequivocally tell whether the use of specialized lexical items follows dictionaries or, conversely, whether dictionaries record the use of these lexical items in specialized texts. Both probably play a role, but one can recognize certain influences of dictionaries on language use in specialized texts. (Schweder, 1999, p. 169, my translation)^{lxix}

What is relatively new is not the collaborative lexicography phenomenon but rather its current prevalence on the Internet and the growing interest of researchers in collaborative lexicography. Also fairly recent is the fact that media changes have altered the way of contributing contents (Cristea, Forăscu, Răschip, & Zock, 2008).

One has been speaking of “cooperative terminology work” only since the end of the 1990s. But the concept is not new, because terminology work has always required a collaboration between domain specialists and language specialists, i.e. terminologists. What is new in this regard is the fact that new technologies have allowed teams that are geographically dispersed to work

193 Furthermore, this assumption is false for the English language as well. The *Oxford English Dictionary* has collaborated with the public at large since the 19th century and the OED dictionary is partly based on the contributions of a large number of occasional unpaid volunteers (Thier, 2014).

on the same terminology, and go on working with the results of others in “real time,” so to say. (Massion, 2009, p. 15, my translation)^{lxx}

Taking into account lexicographical products on the web, Engelberg and Storrer (2016) completed the typologies listed by Wiegand with a typology specifically oriented toward web dictionaries and dictionary portals. This is the typology I am using for the three networks linked to a lexical resource, as shown in Table 7.

NAME OF NETWORK	ASTRONOMIA TERMINARO	RETPOSHTA RONDO POR REDAKTANTOJ	VIVO-VIKIO
NAME OF DICTIONARY	Astronomia terminaro	Reta Vortaro	Vivo-vikio
TYPE OF DIGITALIZATION	From the start	Mixed Expansion of retrodigitalization	From the start
DEGREE OF COMPLETION	Open to new contents	Open to new contents	Open to new contents
OPENNESS TO USER PARTICIPATION	Yes	Yes	Yes
NUMBER OF LANGUAGES	Multilingual	Multilingual	Multilingual
SCOPE	Concepts and specialized lexical items of astronomy	General	(Secondary) neologisms
DICTIONARY PORTAL	No	No	Yes

Table 7. Type of internet dictionary for electronic networks of practice linked to a dictionary, based on (Engelberg & Storrer, 2016).

In the following sections, I present each one of the networks in more detail and give sample discussion excerpts to illustrate the networks’ activities in relation to the lexicon.

5.2.1 Astronomia Terminaro

The first network selected was *Astronomia Terminaro* (specialized dictionary of astronomy). It is a Yahoo group of around 40 members¹⁹⁴ that was active during the years 2000–2012.¹⁹⁵ The aim of the network was to conduct detailed discussions about the contents and technical aspects of the resource *Astronomia Terminaro* (Astronomia Esperanto-Klubo [AEKO], 2000). The *Astronomia Terminaro*, or terminological resource for astronomy, was meant to be the first version of a specialized dictionary that would collect every astronomical and astronomy-related lexical items in Esperanto and suggest lexical items missing from the international language (Astronomia Esperanto-Klubo [AEKO] & Rocha-Pinto, n. d.).

The *Astronomia Terminaro* is a multilingual specialized dictionary. Esperanto is the main language, and equivalent terms are proposed for 7 other languages, namely Czech, English, French, German, Italian, Portuguese, and Russian. The *Astronomia Terminaro* is a restricted dictionary focused on terms related to astronomy and astrophysics. It is exclusively synchronic. At the time of preparation of the dictionary (predominantly 2000–2002), contributors considered only those lexical items that were in actual use or created new units that did not exist in the planned language.

From the end user's perspective, the *Astronomia Terminaro* is mainly an active dictionary for domain specialists who wish to look up an Esperanto equivalent of a term in their preferred language.¹⁹⁶ The dictionary serves to establish the meaning of this notion in the international planned language and set the term to be used between domain specialists.

Each dictionary article in the *Astronomia Terminaro* contains at least a term (if there are several terms for the same notion, the preferred term is used), its source (either existing dictionaries or the author of the new proposal), the definition of the corresponding notion, and one or more foreign language equivalents. Articles may also contain orthographic variants or synonyms of the term, cross-references to related terms, and notions, usage examples, and usage notes.

194 Figure from the year 2017.

195 The network was especially active during the first three years of the project. The central activity linked to the project took place between the years 2000–2002, but the network remained active and a few messages were sent in the following years.

196 By “active” use, I mean that the speaker is in a situation of production of the language (e.g. text production), as opposed to “passive” use, in which the speaker would be in a situation of reception of the language (e.g. reading). See also the communicative situations in Tarp's work (2008, p. 147).

5.2.2 Esperanto-tradukistoj

The second network is *Esperanto-tradukistoj* (Esperanto translators). It is also a Yahoo group with about 280 members.¹⁹⁷ It is not linked to a particular lexical resource and its activity does not lead to the creation of any lexicographical resource or entries but rather serves to address punctual problems. Its conversation list contains more than 9,000 email contributions covering 1600 topics. The aim of the network is threefold (Esperanto-tradukistoj, 2003):

- Publish calls for translation (of documents, web pages, brochures, articles, songs, stories, etc.)
- Discuss terminological and other linguistic problems encountered in translation projects
- Ask other questions linked to translation (translation prices, Esperanto characters, online specialized dictionaries, machine translation, etc.)

In the present investigation, the second aspect—the discussion of lexical items—is of particular interest. In discussions, network members often suggest several solutions and, although they do not necessarily agree on the lexical items to retain, they discuss their choices and deploy lexical opinions and arguments to justify their views.

5.2.3 Lingva Konsultejo

The third network is Lingva Konsultejo (place for language consultations).¹⁹⁸ This Facebook group has been active since September 2011 and has more than 2,000 members. According to the group’s own description,

Lingva Konsultejo is a forum to discuss and give advice on the structure and use of Esperanto. Its objective is to deal with rare, unusual, complicated and unclear aspects of the language, for which answers might be difficult to find. Moreover, it aims to provide information and answers from individuals who have a good knowledge of Esperanto and have various experiences of the language (writing, teaching, reading, etc.). The goal is not to clarify simple

197 Figure from the year 2017.

198 Although it does have administrators, it was included in the investigation because it is a public group in which anyone can post content.

questions, like in a textbook or a basic course, nor to define words that are in a dictionary.

Thus we ask:

1. That you do not ask about things you will easily find in a dictionary, such as *vortaro.net* or *reta-vortaro.de*

2. That you do not post something in the group if you do not have questions related to language

(Lingva Konsultejo, n. d., my translation)^{lxxi}

Like Esperanto-tradukistoj, it is not linked to a particular lexical resource but aims to address punctual problems. Many of the group's discussions concern the lexicon.

5.2.4 Retposhta rondo por redaktantoj

The fourth network is the *retposhta rondo por redaktantoj*¹⁹⁹ (e-mail circle for editors). It is also a Yahoo group, which has been active since 1999 and comprises around 300 members.²⁰⁰ The network is linked to the *Reta Vortaro* (ReVo; online dictionary), an Esperanto language dictionary that records lexical items in use along with their definition and ethnic language equivalents (*Lingva manlibro pri verkado de Revo-artikoloj*, n. d.).

ReVo, which is of the open-collaborative type, covers general naming needs for the Esperanto language. It comprises general language expressions and lexical items in specialized domains. At the dawn of the project, one of the major sources for the ReVo was the *Plena Vortaro de Esperanto*, a print dictionary; the current work of the network falls under what lexicographical studies have called an “expansion of a retrodigitized paper dictionary”^{lxxii} (Klosa & Tiberius, 2016, p. 77). Tasks completed by voluntary editors of the ReVo are, for instance, (*Informoj por redaktantoj kaj redaktantoj*, n. d.):

- Reading new or updated entries and commenting on potential mistakes
- Correcting typos
- Adding equivalents in ethnic language,

199 “retposhta” is spelled with ‘sh’ rather than ‘s’ supposedly because Yahoo Groups did not the full range of Unicode characters when the group was created. Since this can be considered a proper name, the spelling is maintained.

200 Figure from the year 2017.

- Indicating whether a lexical item is official
- Checking lexical items that belong to specialized domains (adding domain indication and correct definition)
- Drawing or taking a picture for an entry in a specialized domain

The objective of the network is for interested individuals (especially dictionary editors) to share information and discuss topics related to the dictionary (Informoj por redaktantoj kaj redaktontoj, n. d.).

As far as dictionary usage is concerned, ReVo is both a passive and an active dictionary. Users can look up the definition of a particular lexical item (passive use) or find an Esperanto equivalent and usage examples from an ethnic language word (active use). Anyone who is interested in participating can join the project and contribute new entries, modify existing ones, or make comments through the discussion list, Revuloj.

The dictionary includes equivalents in more than 60 languages, but many of them are incomplete. Equivalents are offered in 16 languages for at least 10% of the dictionary entries. In decreasing order of completeness, these languages are French (88%), Russian (76%), German (69%), Hungarian (69%), Dutch (63%), Belarusian (56%), English (56%), Polish (48%), Portuguese (40%), Spanish (38%), Catalan (34%), Italian (25%), Bulgarian (14%), Swedish (14%), Breton (11%), and Persian (10%). Furthermore, 50 additional languages have been added to the dictionary interface, but the number of entries they cover has not reached 10%.

In ReVo, each article rests upon a base root and its grammatical ending. Lexical items that are based on a root are listed under the base root. An entry is divided into three main sections: lexical items (lemmas), equivalents in other languages, and sources. The lexical items section comprises compounds and lemmata derived from the base root. Lexical items that belong to a domain of knowledge are assigned a domain label. A definition (usually starting with a hyperonym) and example sentences are provided. Compounds and derivations that are considered to be self-explanatory are not included as lexical items. The equivalents section displays all the foreign language equivalents added by contributors. The sources section lists all the sources used in the dictionary article. The sources can be original or translated literature as well as existing dictionaries.

5.2.5 Vivo-Vikio

The fifth and last network is ViVo-vikio (ViVo-wiki, a wiki for words that live). With around 30 members, the ViVo-vikio is a place for discussion and preparation of new lexical items before their addition to other dictionaries (Demeyere, 2010, n. d.). The activities of the wiki correspond to what Lemnitzer (2010) has called “lexicography of current neologisms”^{lxxiii} (pp. 69–70), a type of lexicographical practice that can provide speakers with useful information about lexical items that have not yet entered a dictionary proper. As stated on the dictionary website, it does not seek to offer a comprehensive description of the Esperanto language:

ViVo is meant for words that “live;” that is, words that can still “die” or disappear. Many of them will deserve eternal life and will enter the heaven of words, which is the Reta Vortaro de Esperanto.

The ViVo-vikio is a place for discussion, where one can anticipate the work needed for other general language and specialized dictionaries such as ReVo, Vikivortaro, and the dictionaries of Lernu. (my translation)^{lxxiv}

ViVo-vikio is both descriptive and prescriptive. It is descriptive insofar as its contributors aim to list all the lexical items found for a given meaning. It can be said to be prescriptive as well, as it offers the possibility of adding pragmatic parameters: Contributors can give their opinions on the status of each lexical item by tagging it with one of four usage labels: *rekomendita*, *uzebla*, *mapli bona*, or *mal-rekomendita*.²⁰¹ From the users’ point of view, ViVo-vikio can be considered mainly an active dictionary, as it is intended to find or create new lexical items for expressing notions that, for the most part, exist in another language (secondary neologisms).

Each dictionary article mainly consists of a definition, one or more foreign language equivalents, and proposed lexical items in Esperanto, which are regularly subject to comments in a discussion field. The source of the proposed lexical items is not systematically indicated, but contributors often mention in the discussion field whether they created the unit themselves or in what real situation they observed it being used.

201 Literally recommended, may be used, less good, and not recommended.

5.3 Linguistic knowledge and language proficiency of network contributors

Who are the contributors to such networks of practice? Are they professional lexicographers in their respective speech communities or ordinary speakers interested in language matters? What are their competencies in relation to the lexicon? To try to answer these questions, I conducted a survey to determine the degree of *linguistic knowledge* and *language proficiency* of network contributors.

5.3.1 Objective of the survey

The survey was designed to help answer the following questions:

- What **linguistic knowledge** do the individuals who contribute to Esperanto electronic networks of practice have? Can they be considered linguists, or rather *folk linguists* (“common” speakers)?
- What **language proficiency** do these individuals have; how competent are they in the language to which they contribute?

5.3.2 Methods

A survey-based methodology was chosen because it is well suited for group administration. Surveys allow confidentiality and an appropriate number of participants.²⁰² The general design principles applied during the creation of the survey instrument were based on best practices as described in reference works (De Ketele & Roegiers, 2009; de Singly, 2012). Most of the questions were close-ended questions.

The survey items were developed considering the results of previous sociological studies in the Esperanto speech community (Fiedler, 1998; Forster, 1982; Piron, 1989; Rašić, 1994; Stocker, 1996). As mentioned, two dimensions were explored: respondents’ **linguistic knowledge** and their **language proficiency**. For the purposes of the survey, these two concepts had to be translated into a

202 In hindsight, I would consider personal interviews, because it proved difficult to reach a satisfying number of responses. I suppose this is because many contributors participate only momentarily, and only a few contributors actively participate (on this aspect, see Abel & Meyer, 2016, p. 262). See the issues mentioned around Table 10 below (p. 168).

measurable form (de Vaus, 2014, p. 41), and indicators were developed. These indicators were spread throughout the survey: 4 indicators for linguistic knowledge (*ling*) and 4 indicators for language proficiency (*lang*). See Table 8 for details.

Dimensions (latent variables)	Indicator description	Indicator reference
Linguistic knowledge	Reading works on Esperanto grammar or terminology	2:1 ling
	Formal or informal training in fields related to linguistics	4:1 ling
	Linguistic knowledge (self-assessment)	4:2 ling
	Main occupation over the past 30 days	6:6 ling
Language proficiency	Esperanto language competence (self-assessment)	2:2 lang
	Esperanto language competence (diploma)	2:3 lang
	Frequency of use of the Esperanto language in the last three months (speaking and writing)	2:4 lang
	Language knowledge (self-assessment)	4:2 lang

Table 8. The eight indicators used in the survey for assessing the two dimensions *linguistic knowledge* and *language proficiency*.

Although well-established measures were used to assess language proficiency, indicators were developed ad hoc for linguistic knowledge. Reliability was evaluated during the pilot phase. Several questions served as measurement variables for each one of these eight indicators. The next section succinctly describes the survey structure and its questions. The full survey can be found in the appendices.

5.3.2.1 Survey design

After a welcome message and a brief explanation about the research, the survey instrument was divided into six main sections:

Section 1 addressed respondents' contributions to online language-related activities in the Esperanto speech community. It aimed to explore how actively respondents participate in the electronic networks of practice considered in the thesis, how long they have been participating, and how much they contribute.

Section 2 concerned respondents' knowledge and frequency of use of the Esperanto language. For the first item (indicator 2:1 ling),²⁰³ respondents were asked to mark, on a list of Esperanto works related to Esperanto grammar or terminology, which ones they had read partly or entirely. Among these works were the *Fundamento* (the foundation of Esperanto, usually considered the obligatory authority over the language), three grammar works, and six other publications on terminology in Esperanto. This item was used as an indicator of linguistic knowledge. The two subsequent items were indicators of language proficiency. The Esperanto language competence of respondents was measured through subjective and objective questions. The subjective question (indicator 2:2 lang) was based on the European Language Portfolio's self-assessment grid developed by the Council of Europe (Language Policy Unit, Council of Europe, n. d.) and concerned reading comprehension and writing skills.²⁰⁴ The objective question (item 2:3 lang) consisted of asking respondents whether they held a language diploma. Although this question may be a more objective and therefore more reliable indicator, it cannot be applied to all respondents, as one can have an excellent command of a language without having passed an examination. Finally, speakers were questioned about the frequency with which they had used Esperanto in speaking and writing during the past three months (indicator 2:4 lang).

Section 3 contained three questions concerning respondents' knowledge and use of languages in general. Firstly, respondents were asked to indicate their mother tongue(s). Secondly, they were asked to list any other language(s) they spoke well enough to have a smooth conversation. Lastly, a question based on a European Commission survey (TNS Opinion & Social, 2012, p. 49) investigated what activities respondents carried out using the languages they know.

Section 4 was composed of questions intended to explore respondents' language and linguistic-related competences. The first question (indicator 4:1 ling) was a closed yes/no question about whether participants had received formal (school, university) or informal (private seminars, self-teaching) training in fields

203 The indicators are coded, either ling (for linguistic knowledge) or lang (for language proficiency) and a number. See Figure 35 in the appendices for details (p. 379).

204 The self assessment grid presents 30 scales for evaluating listening, reading, spoken interaction, spoken production and writing activities. It was officially translated in Esperanto in 2007 (Konsilio de Eŭropo, 2007). The present survey focused on reading comprehension and writing skills. Listening, spoken interaction and spoken production were not assessed.

related to linguistics.²⁰⁵ The second question in this section was a mixed self-assessment question about respondents' language and linguistic knowledge (indicators 4:2 ling, lang). The six aspects covered were largely inspired by the RaDT's²⁰⁶ professional profile for terminologists, which lists prerequisites for terminology professionals (Rat für Deutschsprachige Terminologie, 2004, p. 3). The six aspects covered by the question were (a) language competence, (b) feel for the language and linguistic creativity, (c) competence in several languages, (d) knowledge of Esperanto word formation processes, (e) knowledge of lexicographical and terminological principles, and (f) knowledge of language change phenomena. This question was a self-assessment.

Section 5 comprised subjective and objective questions on the same topic: participants' specialized knowledge in a wide range of fields. The subjective question was a self-assessment,²⁰⁷ whereas the objective question was a closed yes-no question about actual work or formal learning experiences in the range of fields.

Section 6 was intended to gather baseline demographical data: age, country of birth, current country of residence, time spent in the current country of residence, highest level of education, main occupation over the past 30 days (indicator 6:6 ling), and employment status.²⁰⁸ The main goal was to portray the respondents and confirm the international character of language-related activities of the Esperanto speech community. It was also designed with a comparative purpose to be juxtaposed with other surveys in the Esperanto speech community.

5.3.2.2 Survey implementation

The survey was hosted on the University of Geneva's **LimeSurvey server** and potential participants were invited to take part online in the fall of 2015. The LimeSurvey application was chosen because it is free, open source,²⁰⁹ and offers numerous export and import options compatible with other free and open source

.....
205 Theoretical or prescriptive/applied linguistics, translation and interpreting, foreign languages, lexicology lexicography terminology terminography, and interlinguistics or Esperanto studies.

206 The RaDT is the Council for German-Language Terminology.

207 On the scale: not competent at all, somewhat competent, competent, very competent.

208 Employee, director, self-employed, unemployed (students, unemployed, jobseeker or retired individual, or other).

209 In my own experience, Esperanto speakers are fond of open source solutions.

software.²¹⁰ Working on university servers was a plus for data security. Lime-Survey was also a wise choice from a practical point of view, as it is hosted by the University of Geneva and suggested to the institution's research staff. Consequently, it was a turnkey solution that did not require any server installation.

As far as the **selection of respondents** is concerned, my first intention was to make a census—to obtain responses from all contributors of the electronic networks of ReVo (about 300 contributors), Esperanto-tradukistoj (about 280), and ViVo-vikio (about 30).²¹¹ Sampling was not an option because contact information such as names and e-mail addresses of contributors were not fully accessible; although the contents of ReVo and *Esperanto-tradukistoj* messages are publicly available on the web, no member list is available. Only group administrators can fully access members' email addresses. Administrators of ReVo, *Esperanto-tradukistoj*, and ViVo-vikio were contacted and asked whether the member lists could be made available for research purposes. They declined my requests, as this might violate data protection principles and/or the groups' terms and conditions, but they suggested that I (have them) send general messages to the entire group of contributors.

As an initial step, the survey was **pilot tested** on a small group of 7 contributors to ascertain the usefulness and clarity of the questions and reveal any weaknesses in its design (see Table 9). Browsing through individual messages I had collected in my mailbox for ReVo,²¹² I was able to identify a small sample of active contributors. I looked for contributors who had participated for several years (at least two) and appeared to be relatively active (at least 40 contributions in total²¹³). I identified 12 such contributors and sent invitations to participate directly to their

210 E.g. csv export option for further statistical analysis in the free and open source environment RStudio.

211 Data are from 2013. AEKO was excluded because the project is no longer active and Lingva Konsultejo was not included because data for this network were gathered at a later point in time (project constraints).

212 E-mail addresses are not visible in the groups' interface online, but when receiving e-mail notifications from the group one can see the original sender's address. Since I had been a member of the *ReVo* group since 2013 (the year when I collected the corpus data), I could see some other members' e-mail addresses in my mail client when I ran the survey in 2015.

213 The number of contributions for a particular individual cannot be assessed with absolute certainty as contributors can use several aliases in one and the same group. For instance, a contributor was identified that had been using a different e-mail address for every calendar year: username2005@xx.xx, username2006@xx.xx, username2007@xx.xx, etc.

individual e-mail addresses. The message sent to them can be found in the appendices.

After the pilot phase, a **slight adaptation** was required. A typo was corrected, and the question “*Kiom kompetenta vi mem taksas vin por la sekvaj fakoj*” (How competent do you think you are in the following domains) was changed to an array question with a dual scale instead of an array question by column so that people could differentiate between training and work experience in a specific field. Subsequently, invitation messages and reminders were sent to all members of the three groups (ReVo, *Esperanto-tradukistoj*, and ViVo) by myself or their administrators. Potential participants were not offered any incentive. As individuals taking part in online collaborative dictionaries donate their time on a regular basis, I supposed they would be willing to take part in the survey with no direct reward. The invitation message can be found in the appendices.

Estimated number of contributions	Estimated contribution timeframe	Participated to the pilot phase
560	2003–2007	Yes
1500+	Unknown	Yes
513	2000–2011	Yes
468	2006–2012	Yes
106	2005–ongoing	Yes
74	2001–ongoing	Yes
41	2009–ongoing	Yes
677	2001–2004	No ²¹⁴
170	2009–2010	No
646	2003–2011	No
551	2005–ongoing	No
284	2003–ongoing	No

Table 9. Profile of ReVo contributors selected to pilot test the survey.

Several **issues** were encountered during the implementation of the survey. As mentioned, a major problem was the impossibility of accessing member lists. A general invitation had to be sent from a list member—either the administrator or myself. Furthermore, members of Yahoo groups²¹⁵ are not automatically notified when a message is posted. They can choose to receive individual, daily digest, or special delivery e-mails, or to read group posts on the group’s website. Members

.....
214 The e-mail address of this contributor was no longer valid.

215 Astronomia Terminaro, ReVo and tradukado use a Yahoo group.

that read posts on the group’s website may only occasionally contribute and would not have seen the survey invitation in their mailboxes. For example, as of October 2015, only about half of the members of *Esperanto-tradukistoj* had received individual messages:

Type of notification chosen	Number of members
Individual messages	146
Daily digest	45
Special delivery e-mails	20
Read posts on the group’s website	72
<i>TOTAL MEMBERS</i>	283

Table 10. Types of notification chosen by the members of the Yahoo group tradukado. The figures are for October 2015 and were received by e-mail from the administrator.

A further **technical complication** in the case of Yahoo groups is the fact that an individual may use several e-mail addresses in the group, meaning the groups might be smaller than they seem and statistical analyses may be flawed. Finally, some contributors are no longer active in the groups. This is a severe practical difficulty for conducting research on contributors to mass collaboration systems. It is not rare that an individual’s participation is temporary and limited to a few contributions. As a researcher, I did not have access to the detailed statistics of members joining and leaving the groups, thus I cannot draw any objective conclusions about participation. Although the research was dampened by these technical limitations, insightful results were obtained, as presented in the next section.

5.3.3 Results

In this section, I present the data gathered from our respondents. **Forty-one full responses** were received; 26 respondents indicated having already contributed to ReVo, 19 to *Esperanto-tradukistoj*, and 11 to ViVo-vikio (some respondents contributed to several groups). These results should be interpreted as illustrative rather than representative;²¹⁶ as mentioned, ReVo has about 300 contributors,

.....
216 Surveys are usually designed for representativeness, but as mentioned above, in retrospect I would choose others methods if I were to do this research again.

Esperanto-tradukistoj about 280, and ViVo-vikio about 30. Thus, 41 full responses are not sufficient to draw definitive conclusions but provide initial insights.

The results are presented in two main blocks, demographic profile and language knowledge.²¹⁷ The two blocks give a summary of the information gathered through the survey and a descriptive analysis of the data. When appropriate, comparable results from other studies are discussed. A detail analysis of indicators and conclusions concerning the two dimensions (*linguistic knowledge* and *language proficiency*) are presented in Section 5.3.5.

5.3.3.1 Demographic profile

Respondents reported high **levels of formal education**. The majority of them (35) had either a diploma from a higher education institution, such as a university, (24) or a research diploma from such an institution (11). The rest (6) of them had finished high school or another school above middle school. These results seem to be in agreement with Fiedler's statement that Esperanto speakers usually have higher education levels than the average population (1998, p. 24).

In terms of **age distribution**, most of the respondents were 40 or older (36), and almost half of the respondents (17) were 54 or older. According to Fiedler (1998, p. 24), these results are not particularly surprising, as Esperanto is learned mostly by young people on one hand and people in the so-called third age (during retirement) on the other, and several previous studies have shown an overrepresentation of older persons.

When asked about their **work status**, most respondents reported working actively (27) as an employee, a director, or a freelancer. Concerning their **main occupations** (indicator 6:6 ling), about a third of them (13) declared a main occupation related to IT, and only very few of them (3) clearly mentioned a main occupation directly linked to language (text or book revising and/or editing and translation).²¹⁸ This constitutes an interesting result: most of the respondents (38) did not indicate a main occupation related to language. As far as the **country of**

.....
217 Results on the frequency of contributions were not used, because they were group-dependent. The corresponding question was changed for the last group to which the survey was sent, to cover contributions to all groups. Results therefore cannot be compared.

218 Here I must mention, however, that some main occupations mentioned by respondents did not exclude that the respondent had been a language professional in the past (sample answer: *retired*), or were too imprecise to determine if the main occupation was related to language (sample answer: *teacher*).

residence is concerned, respondents lived in 16 different countries,²¹⁹ mostly in Europe. Nine respondents did not live in the country in which they were born.

5.3.3.2 *Language knowledge*

Respondents' respective **mother tongues**²²⁰ included Russian (8), Dutch (7), English (6), German (6), French (5), Portuguese (2), Polish (2), Hungarian (2), Italian (1), Serbian (1), Czech (1), and Catalan (1). Further languages fluently spoken but not mentioned as mother tongues were Spanish, Ukrainian, and Belarusian. Except for the two Hungarian native speakers, respondents therefore all had Indo-European language backgrounds, predominantly from the Germanic, Romance, or Slavic branches. Respondents were multilingual; the **number of languages spoken fluently** varied from two to seven individually. Participants showed a high level of multilingualism in comparison with the average European: in a study published in 2012, just over half of Europeans (54%) were able to hold a conversation in at least one language in addition to their mother tongue (TNS Opinion & Social, 2012, p. 5). Respondents of the present survey spoke an average 1.58 languages fluently *in addition to* their mother tongue and Esperanto. This value is similar to that obtained by Fiedler when asking her respondents what language they spoke²²¹ in her survey of readers of the magazine *Esperanto*. A large majority of respondents reported being fluent in English (see Figure 8).

219 The countries represented were, in alphabetical order: Belgium, Brazil, Canada, Czech Republic, France, Germany, Hungary, Luxembourg, Moldova, Netherlands, Poland, Russia, Serbia, Switzerland, UK, USA.

220 One respondent indicated two mother tongues.

221 Fiedler asked "What other languages do you speak?" ("*Kiujn aliajn lingvojn vi parolas*"). She obtained a value of 1.6 foreign languages in addition to Esperanto.

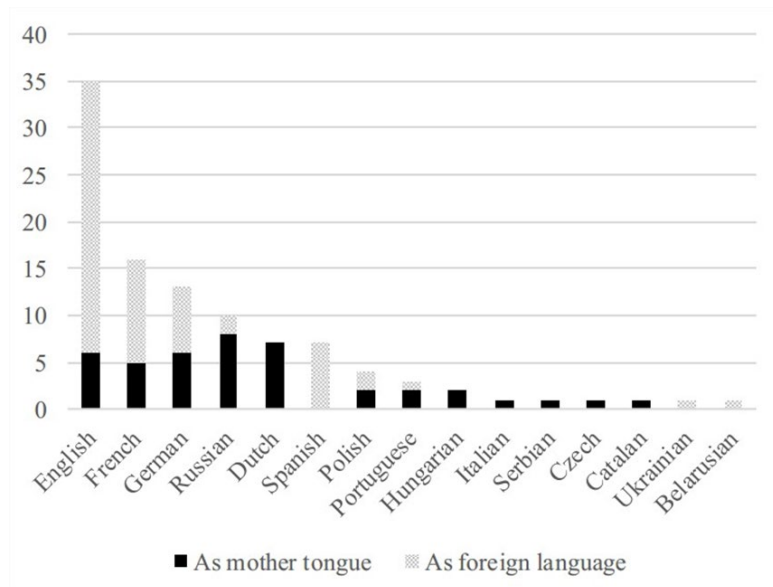


Figure 8. Languages in which the 41 respondents from electronic networks of practice consider they can have a smooth conversation (Esperanto is not included here).

In terms of **situations of language use**, most respondents²²² indicated they used Esperanto for reading books, newspapers, magazines, and communicating with friends and on the Internet; some²²³ responded that they used the language to communicate with family members, for watching movies, listening to the radio, studying languages, or during travel; finally, only a few²²⁴ indicated they use it in professional settings, such as workplace communication (e-mails, phone calls, etc.), reading professional literature, studying something other than languages, or during business trips.

As for their **Esperanto language level** (indicators 2:2 lang, 2:3 lang), almost all respondents self-assessed their reading and writing skills to be of C1 or C2 levels. Over half of respondents (26) claimed to hold at least one language diploma, whether UEA-KER for level B1, B2, or C1 (13); UEA-ILEI (3); or a national examination (17). Hence, most respondents showed high Esperanto language

222 More than 25 respondents out of 41.

223 Between 10–25 respondents.

224 Less than 10 respondents.

qualifications from a subjective (self-assessment) and, for those with diplomas, from a more objective (examination) point of view.

Concerning the **frequency of use of the Esperanto language** (indicator 2:4 lang), about a third of respondents (12) had been very active speaking and writing the language over the past three months,²²⁵ and another group of respondents (10) had been very active in writing but less in speaking.²²⁶ The rest had been moderately active. Finally, two outliers indicated they had not spoken the language over the past three months and only used it somewhat²²⁷ in writing. Only about half of the respondents (12+10), therefore, can be considered to have used the language actively over the previous three months. In comparison, Fiedler obtained a figure of 23.6% of respondents indicating they used Esperanto on a daily basis (Fiedler, 1998, p. 24). Such figures are not particularly surprising if one assumes that Esperanto shares many characteristics of diasporas (Piron, 1989, p. 171) and that it is spoken mainly as a second language (Fiedler, 1998, p. 27).

When asked to mark on a list **which Esperanto works on Esperanto grammar or terminology** they had read partly or entirely (indicator 2:1 ling), each respondent but one responded positively (with a mark) for at least one of the works listed. Everyone but this person had consulted at least one of the three reference grammar works listed. The *Fundamento* had been consulted by almost all respondents (37). Works on Esperanto terminology achieved much less popularity; the majority of respondents (25) had not read any of the works listed related to terminology.

225 They used the language on more than 60 days both in speaking and writing.

226 They used the language on more than 60 days in writing and at least once in speaking.

227 On 11 to 30 days.

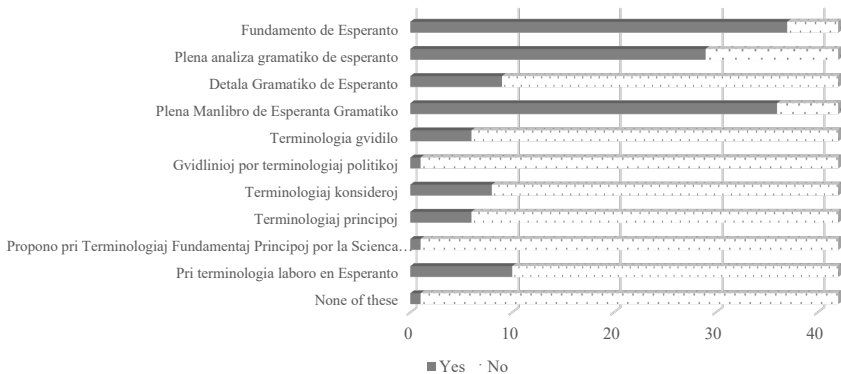


Figure 9. Number of respondents from electronic networks of practice who indicated to have read works on Esperanto grammar or terminology partly or entirely.

When asked to give a **subjective assesment of their language and linguistic knowledge** (indicators 4:2 lang, 4:2 ling; see Table 11), a vast majority of respondents²²⁸ responded positively; they indicated a high degree of competence in Esperanto (29+11) founded in a well-developed feel for the language, great language creativity (26+13), and extensive knowledge about word creation processes of Esperanto (22+16). Interestingly, although almost all respondents agreed they had good knowledge of word creation processes, about half of them²²⁹ responded negatively regarding the last two elements: they indicated they did not have comprehensive knowledge of language evolution phenomena (16+4) nor about lexicographical or terminological principles (13+8). In respondents' minds, therefore, it seems clear that incompetence in lexicography, terminology, or knowledge about language change does not constitute a barrier for contributing to language-related electronic networks of practice.

When asked to indicate **formal or informal training in fields related to linguistics** (indicator 4:1 ling, see Table 12), *all* respondents reported having formal training (school or university) or informal training (private seminars or self-teaching) for at least one of the five topics. If the field “foreign language” is not considered, however, the figures fall to about half of the respondents. As the survey question comprised both formal and informal training, these results must be

228 Those who responded “Extremely well” or “Somewhat” (“Extremely well” + “Somewhat”).

229 Those who responded “Not very well” or “Not at all” (“Not very well” + “Not at all”).

compared with the actual occupations of respondents, which give an indication of whether the individual was a professional working in the field or a layperson interested in the field but having another main occupation. This comparison will be drawn in the discussion section.

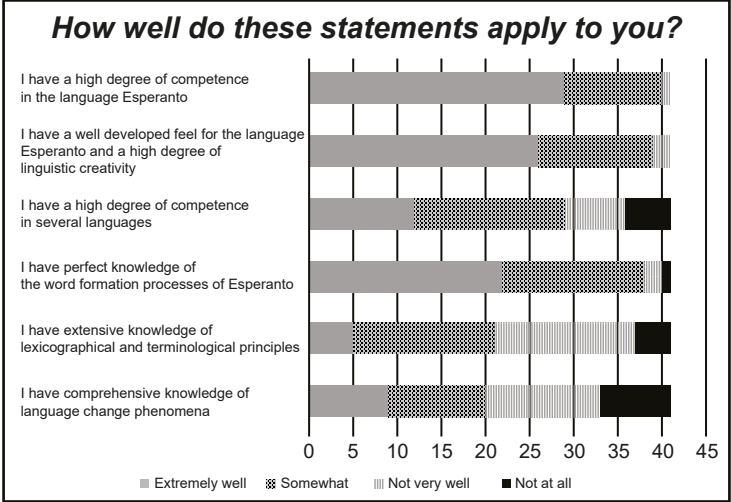


Table 11. Respondents’ self-assessment about their language and linguistic knowledge.

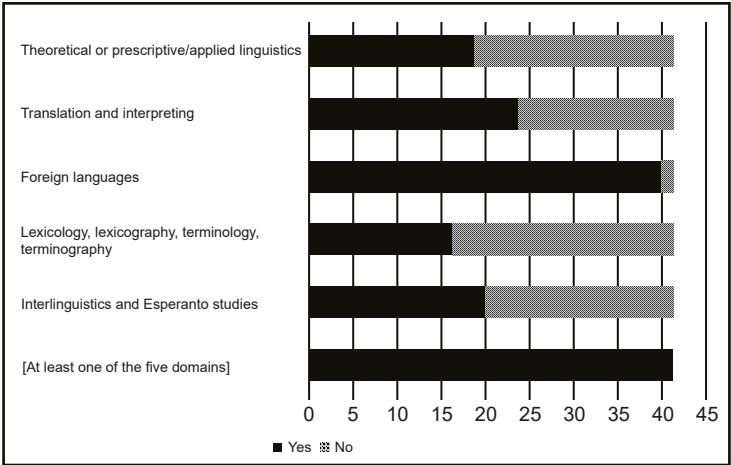


Table 12. Domains for which respondents reported having formal training (school or university) or informal training (private seminars, self-study).

5.3.4 Discussion

In this section, I present a discussion of the results in light of the survey's twofold objectives: determining respondents' (a) linguistic knowledge and (b) language proficiency.

5.3.4.1 *Analysis of results*

As mentioned in Section 5.3.3, **linguistic knowledge** was assessed using four indicators: (a) reading works on Esperanto grammar or terminology, (b) formal or informal training in fields related to linguistics, (c) self-assessment of linguistic knowledge (several criteria), and (d) main occupation over the past 30 days. **Language proficiency** was evaluated through four indicators: (a) self-assessment of Esperanto language competence, (b) diplomas for the Esperanto language (objective competence), (c) frequency of use of the Esperanto language in the last three months, and (d) self-assessment of language knowledge (several criteria). For better comprehension **on the respondent level**, a tentative **metric** was developed to assess each respondent's relative position in the two dimensions of linguistic knowledge and language proficiency (see the detailed metric in the appendices, p. 378). In the domain of folk linguistics, Preston uses the term "folk" to refer to all persons except academic linguists (2011, p. 15). In his definition, therefore, almost all respondents can be expected to be **folk linguists**, as they did not indicate a main occupation related to language in the survey. Instead of adopting a dichotomous approach to folkness, I suggest it might be a question of degree and thus developed a tentative scale. The indicators were given specific weights. One's main occupation had the greatest weight (1 point out of 3) for determining the degree of linguistic knowledge; however, other indicators, such as reading specific language-related works or receiving training, were also taken into account because an individual may become competent in some aspects of linguistics without being a linguist per se. A similar approach was adopted for language proficiency (language proficiency). In this model, respondents gained in language proficiency if, for instance, they used the language on a regular basis, even if they did not hold a diploma.

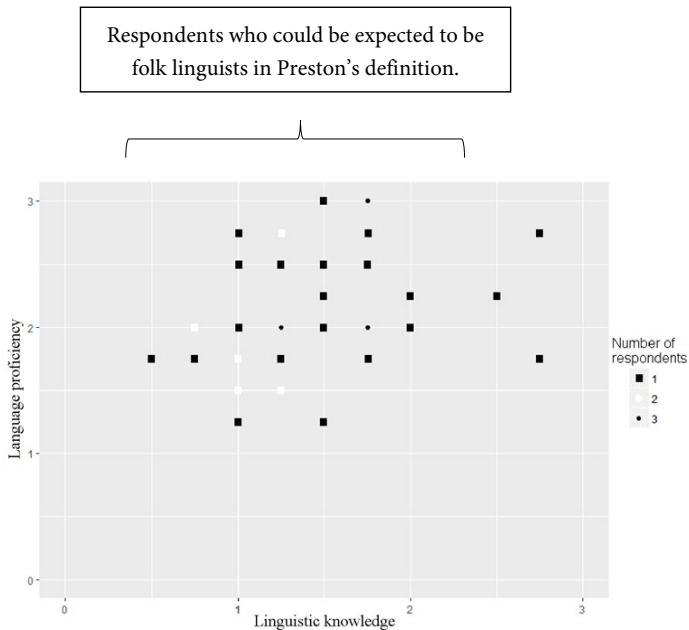


Figure 10. Respondents' relative position on the two dimensions of linguistic knowledge and language proficiency, according to my tentative metric for representing the two dimensions of linguistic knowledge and language proficiency.

The two-dimensional graphic representation of results data based on this metric shows that although most respondents would be folk linguists in Preston's definition, they did have some degree of linguistic knowledge in the model. Two respondents, in fact, received all the points for linguistic knowledge except those for having a language-related profession.²³⁰

Commenting further on linguistic knowledge, many respondents reported having had formal training (school or university) or informal training (private seminars, self-teaching) in domains related to languages, especially in foreign languages. Also, every respondent but one had consulted at least one of the three Esperanto works on grammar listed. Respondents, therefore, could be expected to have at least some basic knowledge of linguistic-related notions. However, more than half of them indicated not having had training in theoretical or prescriptive/applied linguistics nor in lexicology, lexicography, terminology, or termino-

²³⁰ In fact neither of them was professionally employed (one was a student and the other one was jobless and but active in the areas of IT and website development).

graphy, not even on an informal basis. In addition, most respondents had not read the six suggested works on terminology. Finally, about half of the respondents indicated having no comprehensive knowledge of language change phenomena regarding lexicographical or terminological principles. It is doubtful, therefore, that respondents would be competent in terminology/lexicography, neological creation, or language planning; but this does not prevent them from contributing to electronic networks of practice.

As for **language proficiency**, almost all respondents self-assessed their reading and writing skills to be of C1 or C2 levels, and this high self-assessment was often supported with an actual language diploma (though not necessarily at C1/C2 levels). The data thus speak for high Esperanto language qualifications for most respondents. Only half of respondents had been using the language relatively actively over the past months, but this may be explained by the characteristics of the speech community (diaspora, second language). Finally, *all* respondents subjectively indicated having a high degree of competence in Esperanto, and almost all of them declared they had a well-developed feel for the language, great language creativity, and extensive knowledge about the word formation processes of the language. According to our results, respondents appeared to be mostly nonlinguists with strong interests in language.

5.3.4.2 Summary

As reported, **restricted access to contributors** was a challenge, thus censoring contributors revealed a difficult task. This is because member lists were not available to the researcher for data protection reasons. Due to the limited dataset,²³¹ results could only be analyzed qualitatively and not quantitatively.

Indicators for measuring the degree of linguistic knowledge in particular would merit further consideration. The validity of the survey item concerning the current main occupation is particularly questionable. The pilot test did not reveal any problem, but later in the survey, responses were given that were hard to interpret (“retired” or “teacher”). In the future, the question linked to this indicator could be reworded or combined with another indicator.

In addition, some **common problems of survey design** were encountered, such as ambiguity of interpretation from respondents’ perspectives. Although the questionnaire had been piloted, a few respondents indicated in their comments that (a) some questions were not clear or specific enough, (b) the possible answers

.....
231 For instance for *ReVo* only 26 out of about 300 expected group members participated (about 9%).

for some questions did not cover every situation, and (c) sometimes choosing an answer seemed like an arbitrary decision. These issues are typical of the research methodology chosen. Opting for another methodology—for instance, conducting interviews with some of the respondents—may help to circumvent this issue and better understand contributors' profiles.

To conclude, this survey could be complemented by further studies also relating to newer collaborative groups in more recent technologies; for instance, the public page *Lingva Konsultejo* on Facebook. A qualitative approach may be better suited if contributors cannot be contacted easily and the sampling is nonrandom, or alternatively, an approach developed specifically for hard-to-reach populations by which the sampling frame is unknown (e.g., Marpsat & Razafindratsima, 2010).

The survey was designed to **explore the two dimensions of linguistic knowledge and language proficiency** of contributors in three electronic networks of practice examined in the present thesis (ReVo, ViVo-vikio, and *Esperanto-tradukistoj*). To my knowledge, no research to date has questioned the actual language-related competencies of speakers contributing to language-related electronic networks of practice.

Respondents to the survey mostly appeared to be folk linguists in Preston's definition. The statement that the majority of respondents are folk linguists is strongly supported by the fact that more than half of them indicated not having had training in theoretical or prescriptive/applied linguistics nor in lexicology, lexicography, terminology, or terminography, not even on an informal basis, and further by the clear indication that a great majority declared a main occupation that was not related to linguistics (in fact, about a third of them reported having a main occupation in the IT sector).

In contrast, respondents seemed to have an excellent command of the Esperanto language, both from objective and subjective points of view. They felt confident with the language and indicated having great language creativity and knowledge of word formation processes. These results lead to the conclusion that, at least for the speech community under the microscope, some individuals who are active on the Internet are highly competent in their language and choose to spend time engaging in language-related activities, although they are not language professionals. This seems to be a first step in empirically confirming Quirion's ideas that on the Internet a wealth of individuals are ready to invest time and energy for projects for which they care and that folk speakers could engage in neological creation activities (Quirion, 2012, p. 137).

Having presented the theoretical background (Chapters 2 and 3) and the context of my investigation (Chapters 4 and 5), in Part 3, I will continue with the empirical investigation and my proposal for language managers. In Chapter 6, I will explore speakers' lexical environments. In Chapter 7, I will explain how I extracted metalinguistic statements with an opinionated autonym for analysis before I present the results on this analysis in Chapter 8. In Chapter 9, I discuss the results obtained in Chapters 6 through 8 and highlight the types of data that are particularly relevant for language managers.

Part 3

Empirical investigation and proposal

6 Speakers' lexical environments: insights from focus-groups

6.1 Introduction

In Section 3.3.1, we saw that the lack of lexical knowledge is one of the major issues language managers encounter in their endeavors to implant their target lexical items. In lexicography, scholars long took for granted that individuals who need lexical information consult a dictionary (Bergenholtz, Nielsen, & Tarp, 2009). This might be one of the reasons language managers have repeatedly tried to disseminate their target lexical items using dictionaries. For instance, the paper dictionary about fencing, *Diccionari d'esgrima*, was offered to individuals working for the Olympic Games in Barcelona in 1992 (see Vila i Moreno & Vila i Moreno, 2007).

In both studies cited here, the dissemination of target lexical items through a dictionary seems to have been largely unsuccessful. In the first study (Vila i Moreno & Vila i Moreno, 2007, p. 76), only 36% of surveyed individuals even knew that the dictionary existed and less than 10% owned a copy. The second study fared even worse; the target group members interviewed were unaware of the existence of the dictionary, nor did they know that Catalan terms existed for the notions considered (Gresa Barbero, 2016, p. 40).

Using a dictionary to disseminate target lexical items among target speakers amounts to assuming that the subsequent concepts are true: (a) Target speakers conduct external searches for lexical information, (b) they do so by using a dictionary, (c) among all dictionaries available, they use those of language managers, and (d) they actually use the lexical items found in this dictionary. Such issues call for a better understanding of the lexical environment in which target speakers find themselves, but studies investigating target speakers in their environments (e.g., Ballarin, 2009; Ní Ghearáin, 2011) are rather the exception. Let us examine the four assumptions.

(a) Do speakers conduct external searches for lexical information at all? Investigating lexical sources that speakers use could have fallen into the scope of lexicographical research, but to date, lexicography has mostly focused on specific dictionaries, usually institutional or otherwise prestigious ones (e.g., Nesi, 2012b), and has largely ignored the extralexicographical situations in which speakers find themselves. According to Tarp (2009), "No known user research has produced real information on the objective user needs, i.e. the needs that may occur in the

extra-lexicographical situation preceding the dictionary consultation” (pp. 292–293).

(b) Do speakers undertake external searches for lexical information using dictionaries? Lexicography scholars now claim that it is evident that individuals can satisfy their need for lexical information from sources other than traditional ones (dictionaries), such as newspapers, books, and online texts (Fuertes-Olivera & Tarp, 2008, p. 77; García Llamas, 2015, p. 63). A new, specific research area named *accessology* has been developing (see Bergenholtz & Gouws, 2010), focusing specifically on access to lexical information.²³² Only very recently (e.g., Kunkel, 2015) have scholars started to ask ordinary speakers to find language-related information.²³³

(c) Among all dictionaries available, would speakers choose those of language managers? Quirion (2012, p. 137) suggested using an online collaborative dictionary in which language managers would encourage Internet users to write their lexical ideas and comments on a lexical platform (website). Language managers would validate a certain number of the proposals, and the users themselves would choose the lexical item to be retained among all the proposals. Quirion suggested that such a dictionary platform might be better accepted by speakers than a traditional target lexical item dictionary. Whether speakers might prefer such a platform has, to my knowledge, not been investigated.

(d) Finally, do speakers actually use the lexical items they find in dictionaries? Being in possession of or consulting a dictionary does not necessarily equate to actively using these dictionary contents in speech.

To approach these fundamental questions not from language managers’ but directly from speakers’ perspectives, I explored speakers’ lexical environments through a focus group study. In Section 6.2, I present the methodology applied: why focus groups were chosen as a methodology, how participants were recruited, how groups were implemented, and what analytical framework was used for analysis. In Section 6.3, I report and discuss the findings in detail. I provide a summary of the main findings in Section 6.4.

232 In a sense, lexicography has always dealt with access (Bergenholtz, Bothma, & Gouws, 2015), but now some studies are taking lexical information access as their core research question.

233 Translation studies have been interested in the way language professionals search for lexical information (e.g. Künzli, 2001; Nord, 2002), but the general public has been largely ignored.

6.2 Methods

6.2.1 Focus group rationale

In this chapter, I opted for a focus group methodology. A **focus group** is “a group interview—centred on a specific topic (‘focus’) and facilitated and coordinated by a moderator or facilitator—which seeks to generate primarily qualitative data, by capitalising on the interaction that occurs within the group setting” (Sim, 1998, p. 346; on focus groups, see also Liamputtong, 2011; Morgan, 1998b; Reid & Reid, 2005; Stewart et al., 2007).

Traditionally, focus groups have been conducted face-to-face. With the advance of new technologies, however, they have increasingly shifted to taking place online.²³⁴ In the present investigation, the focus group interviews were conducted online for three main reasons. First, it has been demonstrated that more ideas are generated through computer-mediated communication (Reid & Reid, 2005, p. 131). Second, but not less importantly, online focus groups represent a viable solution for hard-to-include populations (Tates et al., 2009), of which the Esperanto speech community can be considered an instance (see Section 4.2.5). Finally, online focus groups are likely to be a more appropriate method for researching a topic involving online issues (Hughes & Lang, 2004, p. 107). As lexicographical resources (e.g., e-dictionaries) and other resources used by ordinary speakers (e.g., search engines) are increasingly shifting from paper to online-only, it seemed relevant to use an online research method.

In contrast to surveys,²³⁵ this methodological design allowed me to gain a wide view and gather profound information, as I could ask respondents to explain their thoughts. As Flick (1998) mentioned, “opinions which are presented to the interviewer in interviews and surveys are detached from everyday forms of communication and relations. Group discussions on the other hand correspond to the way in which opinions are produced, expressed and exchanged in everyday life” (p. 116). Focus groups are “particularly useful for exploratory research when rather little is known about the phenomenon of interest” (D. W. Stewart et al., 2007,

234 Focus groups conducted online are typically called online focus groups, computer-mediated focus groups, Internet-based focus groups, electronic focus groups, chat-based focus groups or virtual panel discussions (Tates et al., 2009, p. 2).

235 Lexicographical research has been keen on using survey methodologies. According to Flinz (2014, p. 214), surveys are the most frequent methodology used for research on dictionary usage situations (both for second language and native speakers), on specialized dictionaries and on online dictionaries.

p. 41). This is because focus group settings allow participants to stimulate one another and clarify their points of view. Focus groups emphasize discovering unanticipated findings and are indispensable for assessing a range of opinions about an issue (Chambliss & Schutt, 2013, p. 199).

In this investigation, I opted for **synchronous focus groups**²³⁶ in which participants all took part simultaneously in a prearranged live session using a chat room or an online conferencing tool. They immediately reacted to each other's responses.²³⁷ I had three primary reasons for choosing this variant. First, I wanted to be able to clearly limit participation time and reassure participants that they would spend, at the most, 1.5 hr. on my study. This seemed important because participation was voluntary and not compensated by any incentive. Second, conducting a synchronous group helped me make sure every participant would spend roughly the same amount of time on the study.²³⁸ Third and finally, I was interested in the immediacy of responses because I did not want participants to look up any information or discuss the issues with other speakers outside the groups. In synchronous groups, any thought is posted immediately (Hughes & Lang, 2004, p. 101).

6.2.2 Discussion guideline

Following established methodological practices, I developed a questioning route—a sequence of questions in complete, conversational sentences (Krueger, 1998b, p. 9). This allowed me to ensure comparability across groups and therefore the quality of the subsequent analysis (Krueger, 1998b, p. 12). The discussion guideline consisted of an introduction, opening, transition, and questions.²³⁹ The complete guideline can be found in the appendices.

.....
236 Online focus groups can be conducted either synchronously or asynchronously (Rezabek, 2000).

237 Asynchronous focus groups use email, a listserv or mailing lists and participants can read each other's contributions and post comments at any time over a set period, whenever it is convenient for them.

238 Asynchronous groups, for their part, encourage behaviors of dominant talkers (Krueger, 1998c, p. 58), that is individuals who consider themselves to be experts and dominate the talk and also behaviors such as monologuing (Hughes & Lang, 2004, p. 99), i.e. typing a series of comments on a solitary thread.

239 An essential feature of focus groups is that not all questions are equal (Krueger, 1998b, p. 21): some questions—opening, introductory, and transition questions—are phrased with the sole purpose of preparing the participants for the pertinent, key questions. My guide hence comprised four types of questions: opening, introductory, transition and key questions.

The guideline was divided into two main sections.²⁴⁰ The first section was built around situations in which speakers were missing lexical items in Esperanto. It served to investigate whether speakers would report conducting external searches for information in such situations and, if applicable, where they conducted such searches (traditional sources, Web, etc.).

The second section concentrated on Quirion's idea of a collaborative dictionary. Quirion (2012) suggested implanting new lexical items by a collaborative process; language managers should work together with the masses, with ordinary speakers who would give lexical input and vote on the final lexical item to be retained for a specific notion. This approach could fall into what is called "collaborative lexicography," or "a bottom-up approach (Carr, 1997) which encourages lexicon readers to contribute to the writing of lexicon entries" (Meyer & Gurevych, 2012, p. 259), or more precisely, semicollaborative lexicography, a platform on which "users can collaborate by proposing something, but they don't have access to the backend databases of the project, which therefore cannot be modified by them" (Melchior, 2012, p. 337).

Typologies of collaborative dictionaries can vary from one scholar to the other, as they vary between dictionaries at large. As presented in Section 5.2 (p. 149), Wiegand et al. (2010), for instance, distinguish between four types of dictionary typologies (pp. 82–93). However, in a focus group where ordinary speakers are present, the question cannot be asked in a linguist's terms. Therefore, I chose to approach the question of collaborative dictionaries with participants without using a specific term but rather a paraphrase, speaking of dictionaries "compiled partly or completely by non-specialists" and letting participants tell the story. As Krueger (1998a) mentioned, in qualitative research, "We seek to tell someone else's story ... we must listen before we can understand." (p. 3)

6.2.3 Choice and recruitment of participants

Careful recruitment of participants is a key for the success of focus groups (Morgan, 1998a, p. 85). The goal of such groups is to hear from participants in depth, which implies a purposeful selection for generating the most productive discussions (Morgan, 1998a, p. 56). Therefore, purposive sampling is the method I chose to apply for recruitment. It is a nonrandom sampling technique by which a smaller

240 To avoid the common mistake of cramming the interview guide with too many questions and not having time to ask the participants about the reasons of their responses (D. W. Stewart et al., 2007, p. 115).

group of key individuals are targeted to represent the views and attitudes of a larger group. In purposive sampling, a specific population is identified, and only its members are included in the research (Kelley, Clark, Brown, & Sitzia, 2003, p. 264).

My purposive sampling strategy served two objectives: (a) to avoid overrepresentation of language professionals or language aficionados²⁴¹ and (b) to recruit participants with various backgrounds: ordinary speakers, language professionals, and individuals who contribute to what I henceforth will call **alternative resources**,²⁴² or any nontraditional resource a speaker uses for responding to a lexicographically relevant need. Such resources include not only online collaborative dictionaries but also online discussion groups or lexicon-related collaborative platforms that are not dictionaries (e.g., Tatoeba). Although the language competence of my participants was meant to be fairly homogeneous, their linguistic background was purposely heterogeneous.²⁴³ I brought together people with training/experience in language-related domains and ordinary speakers, and involved a few speakers who actively contribute to alternative resources. The advantage of heterogeneous linguistic knowledge was that participants would elicit a wide range of opinions and attitudes and would need to explain their views to each other. Additional minor criteria were:

- Age and country of residence: Because Esperanto is an international L2 language, I wanted to have participants from several countries and continents; I also wanted opinions from both younger and older people
- Linguistic knowledge: The objective was to include both speakers who had training or work experience in at least one language-related domain, ordinary speakers, and individuals who contributed to alternative resources
- Content creation on the Web: I chose participants that were active on the Web to ensure a minimal computer literacy and Internet access (a prerequisite for online focus groups)
- Sufficient command of the language: I wanted to recruit fluent participants in order for the focus groups to run smoothly

.....
241 This would be likely to happen if I made an open call to take part in a language-related study.

242 The term 'alternative resources' here is inspired by Nesi's concept of 'alternative e-dictionaries'. Since speakers use sources other than dictionaries for responding to lexicographically relevant needs, I prefer to speak of 'alternative resources'. (2012b)

243 Researchers have suggested in some situations it is desirable to have focus groups that are made up of a particular mix of people such as users and nonusers of a product or service (D. W. Stewart et al., 2007, p. 54).

To recruit participants, I relied on individual canvassing for listing participants with a good command of the language without necessarily being language professionals and on existing online groups for contacting groups of individuals who contribute to alternative resources. My canvassing strategy consisted of contacting individuals who owned blogs and/or websites in Esperanto.²⁴⁴ Existing group recruitment was based on individuals contacted in a previous survey I conducted on alternative resource contributors who had explicitly agreed to be contacted again.

6.2.3.1 *Blog authors sample*

Contacting blog authors was not as easy as I expected, as many Esperanto-speaking bloggers used a platform such as Ipernity, which requires signing up and does not allow one to contact a large quantity of fellow bloggers. Also, some blogs did not allow contact other than posting public comments. In such cases, for ethical reasons I chose not to contact these authors, as posting a public comment about a research project may have been seen as intrusive.

The blog authors sample comprises two subsamples. My first sample was created on the basis of a list of 2753 Esperanto blogs²⁴⁵ compiled by an Esperanto speaker in Japan. I suspect that Japanese blogs were over-represented. All blogs were manually consulted, and only a very modest share could be included in the final sample:

- In 55% of cases, blogs were excluded because their author could not be contacted
- 14% of cases were excluded because the blog URL was a duplicate, either within this sample itself or because it already appeared in another one of my samples
- About 6% of further cases were excluded because the URL link to the blog was dead or the blog had disappeared and content from the provider was displayed

244 Contacting blog and website authors presented a large range of advantages: Firstly, I could conclude from their posts or webpages that they had a sufficient level of the language to take part in an in-depth group interview. Secondly, contacting them would not raise ethical issues because if their e-mail address was visible on the blog or website or if first contact was enabled through a form, the individuals to whom I was about to write would expect to be contacted and should not perceive my message as an intrusion. Thirdly, since they authored blogs and/or websites, one could suppose that they use the language actively, at least in writing. Finally, I knew they were computer-literate and could participate in an online chat, a stringent criterion since I opted for an online methodology.

245 See <http://www.eonet.ne.jp/~skrg/esperantajblogoj.html> (last accessed 2016-02-01).

- About 20% of other blogs were considered irrelevant for my investigation (and therefore excluded): blogs that displayed (almost) no contents (about 6%); blogs that were mainly or exclusively in another language²⁴⁶ (about 6%); blogs created not by an individual but rather by an Esperanto group for a specific event or project (about 6%); blogs whose author I knew personally (about 2%)²⁴⁷; and blogs authored by children or containing adult content (only isolated cases)

Thus, from the original list of 2753 blogs, there remained a total of 126 blogs that could be listed for my research (about 5%). Noticing that the first sample of blog authors was rather small, I tried adopting another approach and created a second subsample. Using search engines,²⁴⁸ I was able to add additional relevant bloggers to my list, thus the blog sample consists of 132 units.

6.2.3.2 *Website authors sample*

My second sample was based on a list of 129 personal websites compiled by the Esperanto speaker Frank Merla, and available on his website.²⁴⁹ As the website is hosted in Germany, I suspect that European and especially German websites were over-represented. This was welcome, as it could counterbalance the large quantity of Japanese blogs sampled. I visited all 129 websites and found that 54 of them (about 45%) were relevant²⁵⁰ and offered a contact possibility (either e-mail or contact form).

246 These languages were English (44), Japanese (41), Portuguese (21), Spanish (20), Russian (15), Chinese (11), French (7) Indonesian (6), Hungarian (5), Korean (5), Polish (4), Catalan (3), Italian (3), Farsi (2), Greek (2), Norwegian (2), Swedish (2), Arabic (1), Czech (1), Finnish (1), German (1), Glosa (1), Hindi (1), Interlingua (1), Irish (1), Romanian (1), Sjal (1) and Urdu (1).

247 I feared the fact that we knew each other would have an impact on the research results.

248 I used Google search engine from my Mozilla Firefox browser with Geneva's university IP address in Switzerland, on February 2, 2016. I launched a search using the key words "esperanto blogo" and search criteria 'Past year' and 'verbatim'. I listed the first 100 results. This approach was not particularly useful, as about a third of the search results did not lead to a blog per se. Also, in many cases there was no contact possibility, the contents were that of a group rather than an individual or were duplicates from my previous sample. Ultimately, this second sample contained 6 additional bloggers.

249 http://maklerejo.de/?Ligiloj___Personaj_retejoj, last accessed 2016-02-17.

250 Since in the past I personally have participated actively in events of the Esperanto-speaking community in Germany, I knew quite a few of these website authors (39), which I excluded from the sample.

6.2.3.3 E-network contributors sample

Finally, I included individuals who had contributed to at least one of the five electronic networks presented in Chapter 5. I sampled contributors who had participated in an earlier research survey and had agreed to be contacted again. This sample amounts to 21 Esperanto speakers.²⁵¹

As illustrated in **Table 13**, the final frame population comprised 207 individuals, 186 of whom were blog or website authors and 21 of whom were previous survey respondents.

<i>Samples:</i>	<i>Blog1</i>	<i>Blog2</i>	<i>Web</i>	<i>E-net.</i>	<i>Total</i>
<i>CONSIDERED FOR SAMPLE</i>	2753	100	129	26	3008
<i>EXCLUDED FROM SAMPLE</i>	2627	94	75	5	2801
<i>No contact possibility</i>	1563	11	14	0	1588
<i>Dead link or empty page</i>	361	2	2	0	365
<i>Group content</i>	205	16	6	0	227
<i>Other languages</i>	204	1	4	0	209
<i>(Almost) no contents</i>	180	6	2	0	188
<i>Known author</i>	72	8	39	5	124
<i>Duplicate</i>	40	15	7	0	62
<i>Not a blog or personal webpage</i>	0	35	1	0	36
<i>Child or adult content</i>	2	0	0	0	2
SAMPLE	126	6	54	21	207

Table 13. Summary of items respectively considered for the samples, excluded from the samples and included into the final samples.

.....
251 I excluded speakers whom I knew personally (5 out of 26).

6.2.4 From samples to focus groups

The 207 individuals of my frame population were all contacted using a comparable contact script (see under Contact scripts in the appendices, starting on p. 387). This script informed participants of the broad goal of the study and led to an online prescreening survey.

6.2.4.1 Prescreening survey

Of the 207 individuals who were contacted, 48 (23%) filled out the prescreening survey (see **Table 14**).

<i>Samples:</i>	<i>Blog1</i>	<i>Blog2</i>	<i>Web</i>	<i>E-net</i>	<i>Total</i>
<i>Individuals in the sample</i>	126	6	54	21	207
<i>Individuals who completed the prescreening survey</i>		18	16	14	48

Table 14. Number of individuals who completed the prescreening survey vs. number of individuals in the four subsamples.

This survey, which is fully reproduced in the appendices, was initially designed specifically for filtering out unwanted participants. Because the focus groups concerned a language-related topic, I supposed language specialists would be eager to participate. I wanted their contribution to be kept to a reasonable figure because linguists probably only represent a tiny proportion of the Esperanto-speaking population.²⁵² I further wanted to avoid imbalance of domain knowledge, for I feared the focus group could turn into an agonal dialogue if a large number of language specialists were conversing with a small number of ordinary speakers. As observed by the sociologist Amey (1996, p. 79), it may occur that in a debate on a specific topic, experts on the matter take advantage of their higher status to disqualify less informed speakers who do not share their views. Hence, previous training or work experience in a field related to linguistics²⁵³ was assessed through

252 In a survey-based study conducted by Rašić (1994, p. 105), only 5 out of 156 respondents (i.e. about 3%) indicated “lingvisto ks.” (linguist or similar) as their profession (There were in total 19 categories of occupation). Also from personal experience there are many Esperantists who are not at all involved with linguistics.

253 Linguistics or applied linguistics, translation and interpreting, foreign languages, lexicography or terminology, esperantology or interlinguistics.

the prescreening survey, as illustrated by Figure 11. Further objectives of the screening were to ensure potential participants agreed with participation conditions (informed consent) and to collect their different time zones and availabilities (practical information).

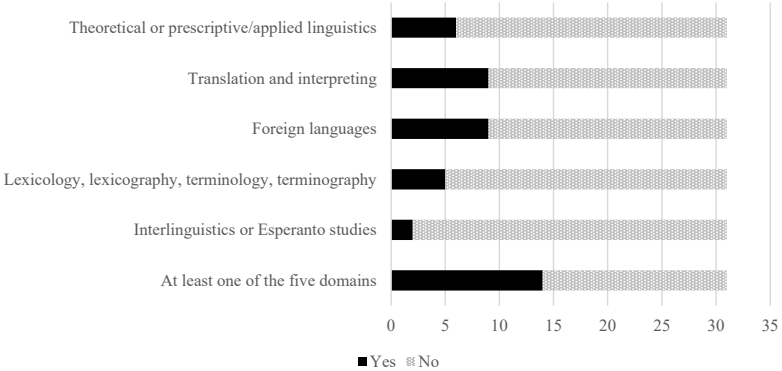


Figure 11. Domains for which focus group participants (there were 31 participants, see Section 6.2.4.2) reported having a diploma or work experience. Data were gathered through the prescreening survey.

At this stage, I took steps to help recruited participants understand why and how data was collected and for what purpose it would be used (Convery, 2012, p. 53). Invited participants had the right to opt out at any time of the data collection process and were informed about this possibility. They also had the opportunity to contact me as a researcher at any time of the process. Once potential participants had fulfilled the prescreening, they were invited to a group based on the time availabilities they had listed in the survey and on the group composition rationale explained in the following section.

6.2.4.2 *Group composition*

I chose to construct heterogeneous groups, as participants with different backgrounds lead to intensified dynamics in discussion, which reveal more aspects and perspectives of the phenomenon under study. Heterogeneous groups are more suitable to investigate a diverse range of responses and can prove useful when assessing the attitudes and beliefs of a community (Liamputtong, 2011, p. 35). Furthermore, as Stewart et al. mentioned (2007), “if a group of technical specialists is brought together to discuss a complex problem, it is likely that the discussion will take on a

very different character than if the group were composed of a few technical people, a few nontechnical but knowledgeable lay persons, and a few novices” (p. 51). Also, I made sure to mix participants from various age groups and geographic locations. Although it may seem evident to proceed so for the international language of Esperanto, this posed some practical difficulties due to participants’ disparate time zones and personal availabilities.²⁵⁴ Although several participants showed considerable flexibility, this practical aspect was a decisive factor.

No participant was excluded from the focus groups based on answers to the prescreening survey. However, out of the 48 individuals who showed interested and agreement in participating, only 31 were eventually part of a group.²⁵⁵ The 17 other either explicitly dropped out, did not further respond, or did not have schedules compatible with those of other participants.

I chose to begin with relatively small groups and formed groups of about 6 participants.²⁵⁶ According to Morgan (1998a), the adequate size for focus groups lies between 6 and 10 people.²⁵⁷ Online groups cannot be as large as face-to-face ones, the maximum number of participants probably lying between 6 and 8, and even in smaller groups with only 5 participants, some elements might be missed if participants are multi-threading and all posting in parallel, making the chat screen scroll too quickly (Hughes & Lang, 2004, p. 107). Selected participants received an invitation to a specific group (see under Confirmation scripts in the appendices, p. 401).

In Figure 12, an overview of group participants is provided in the form of a Venn diagram. The appendices contain additional data about the participants (country of residence, formal education, age, and use of Esperanto).

254 15 time zones were represented. See Stewart & Williams (2005, p. 406) on practical issues of international online focus groups.

255 $n_1 = 8, n_2 = 5, n_3 = 5, n_4 = 6, n_5 = 7$.

256 Participants were over-invited at first in order to compensate for an expected no-show rate. It is indeed advisable to recruit more individuals than required and to assume that at least 2 participants will cancel, sometimes at the last minute (D. W. Stewart et al., 2007, p. 58).

257 Sometimes, however, smaller groups (3–5 participants) have been reported to run more smoothly than larger ones (Peek & Fothergill, 2009, p. 37) and in the literature opinions vary concerning the most effective size for a group. Moreover, according to Krueger “[...] there is greater benefit in conducting two groups of six participants instead of one group of twelve, This gives the researcher the power to compare the results of the two groups.” (Krueger, 1998a, p. 18).

6.2.4.3 Online implementation

The 5 focus groups were conducted in an anonymized environment. Participants each received a pseudonym and a password to access an online chat interface (see under Technical script in the appendices, p. 403).²⁵⁸ Several reminders were sent to participants before the group took place.

For the chat interface, I chose AJAX Chat.²⁵⁹ It is a free, open source, fully customizable web-based chat implemented in Javascript, PhP, and MySQL. The standalone instance (release 0.8.7) was installed on a private web server. The language files were localized into Esperanto. Data from all focus group conversations were saved as separate MySQL tables on a private server and then downloaded to a local hard drive immediately after each focus group session. More information about the choice of this interface can be found in the appendices.

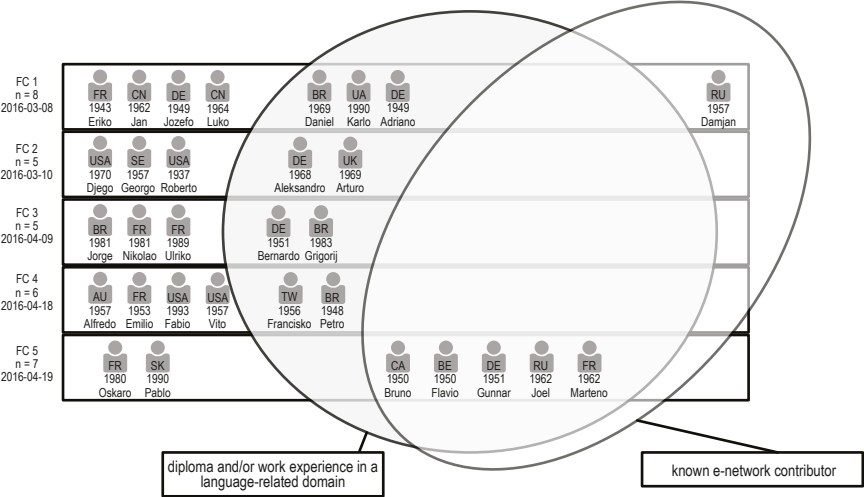


Figure 12. Distribution, represented as a Venn diagram, of focus group participants according to the following criteria: professional language training and/or experience, known contributors to electronic networks of practice, current country of residence (ISO Alpha-2 country codes), and year of birth. The names are pseudonyms.

²⁵⁸ They were asked to test the chat interface before the group interview in order to avoid any technical problems during the real interview.

²⁵⁹ It was developed by Sebastian Tschan and is maintained by Philip Nicolcev. <https://frug.github.io/AJAX-Chat/> (last accessed 2016-01-20).

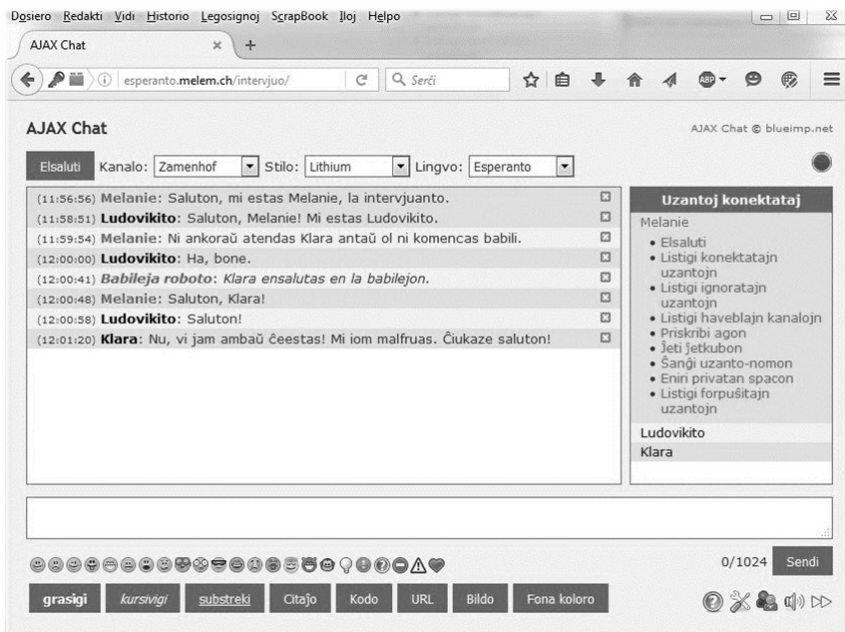


Figure 13. AJAX Chat (Esperanto localization) in a Web browser with sample conversation involving one moderator (Melanie, light font) and two participants (Ludovikito and Klara, dark font).

During the focus group interviews, the discussion guideline was systematically chronologically followed. As a moderator, I adopted a conversational approach, and tried to let participants speak as much as possible. Sometimes intervention on my part was necessary for the discussion to remain focused or to avoid that a particular individual dominate the discussion. Sometimes questions triggered off immediate reactions whereas at times no answer could be offered by participants before I as a moderator gave guidelines about what was being asked.²⁶⁰ Question 8 was asked only in the first group and then abandoned because time was scarce and I decided to make sure I kept enough time for both key questions. A figure in the appendices shows how time was divided between the questions.

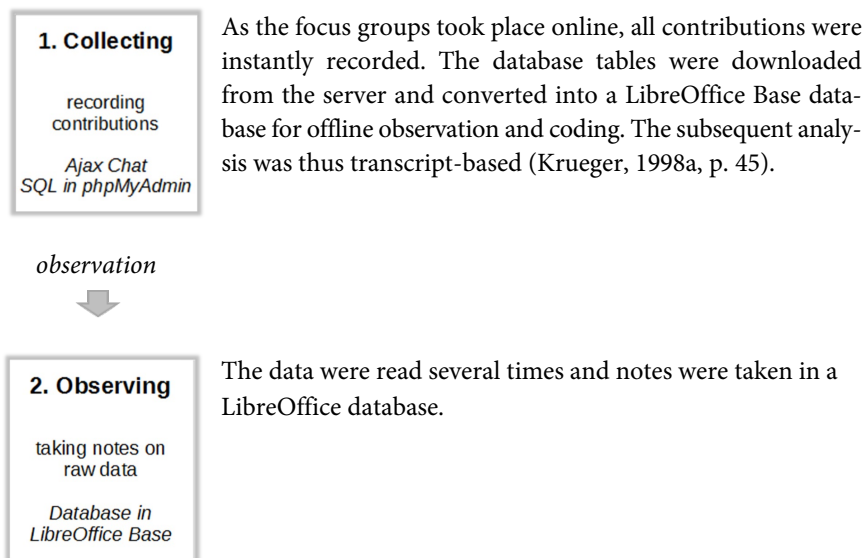
Participants all contributed actively, each sending an average of about 30 messages to the group. Across focus groups, individual participation was therefore relatively balanced, with the exception of a few dominant talkers. Not only did

²⁶⁰ Focus group 5, for instance, seemed to find question 6 too vague and could barely give replies other than “It depends on the context”.

some speakers seem to have more stories to share, but some seemed to type faster, which gave them more room in an online setting. Age was perhaps a factor as well; for instance, Ulriko was a fairly young participant who may have had better typing skills than participants of an older generation. A figure in the appendices contains an overview of the contribution distribution across participants.

6.2.5 Analytical framework

The way focus group data should be analyzed is subject to debate (Liamputtong, 2011, p. 172). In the present investigation, I used thematic analysis, a method used to identify, analyze and report patterns within a data set (Liamputtong, 2011, pp. 173–174).²⁶¹ From the thematic perspective, my analysis was guided by Flick's basic open coding questions (1998, p. 183). My analytical framework consisted of 5 subsequent steps.



²⁶¹ Other approaches are e.g. an impressionistic summary, assertions analysis, pragmatical analysis and analysis of associative structures (see D. W. Stewart et al., 2007, p. 133).

abstraction



3. Coding

grouping and
categorizing data

*Database in
LibreOffice Base*

The coding was guided by three predefined families of codes:

1. Lexical resource
2. Lexical search behavior
3. Lexical resource criteria

For each family, open codes were inductively identified in the data and affixed to sets of notes. This interpretation process was iterative, as I went several times over the data, completed my notes and revised my coding scheme as I progressed. Mutually exclusive and exhaustive codes were abstracted to reduce the data. The full coding scheme can be found in the appendices.

interpretation



4. Examining

exploring
categorized data

*Database in
LibreOffice Base*

Once coded, the categorized data were examined again. In particular, statements that had received the same code were compared.

selection



5. Presenting

using continuous text
and text matrices

Finally, in Section 6.3, results are presented in continuous text, using direct quotes where appropriate. They are discussed in light of existing research.

6.3 Results

6.3.1 Is lexical information needed?

The focus groups showed, first and foremost, the absence of an external search for lexical information in a host of reported cases in which speakers were missing lexical items in Esperanto. Instead of looking for external information, participants reported using what translation studies might call functional equivalences (e.g., Nida, 2001). Several of them mentioned that at times they, quite simply, circumvent the lack of a lexical item in Esperanto. They do so using another word, a metaphor, a paraphrase, by defining the notion, by quoting the word in a foreign language, or even by switching languages altogether if the conversation partner speaks another language. These circumventing strategies were reported to be used both in spoken and written situations. In some reported cases, no external search was conducted because the situation did not allow for it:

Situations of spoken language remain a problem, because then it's not possible to simply consult Wikipedia, encyclopedia, the Internet, etc.

Gunnar (my translation)^{lxxv}

However, not conducting external searches was at times also a deliberate choice in situations where speakers could have had the opportunity to consult lexicographically relevant sources:

In my experience, sometimes with people I correspond with, we do not know the appropriate expression in a situation, so we simply write the meaning of this expression and the word in our language, or in another language.

Jorge (my translation)^{lxxvi}

Another commonly reported behavior that does not involve external searches for information on the part of speakers is lexical creation. Participants recounted creating missing lexical items. According to several of them, it is even easier to create new lexical items in Esperanto than it is in other languages. This is not a particularly surprising statement, as Esperanto was designed for active ease of use with an explicit mechanism of word formation (see Schubert, 2015). The observed phenomenon that speaking several languages can trigger word formation processes may also play a role. As Pruvost and Sablayrolles (2003) pointed out regarding French, “One noticed that natives of a language other than French who were learning French ...

do not hesitate to create new designations in French by applying word formation rules they have assimilated. Mastering more than one language probably has an influence on the intellectual mechanisms that come into play in language-related activities, and the mental gymnastics underlying the transfer from one lexicon to the other is likely to facilitate word formation processes in any language” (p. 78, my translation).^{lxxvii} According to their descriptions, participants use standard designation formation processes (i.e., the combination of existing Esperanto roots), as well as structural calques or false loans, (i.e., loans with graphic/phonetic and morphological change; following Humbley’s terminology, 2015, p. 38). The ever-lasting debate of the Esperanto speech community regarding the two main possibilities for lexical creation in Esperanto, (i.e., using internal resources of the language versus borrowing roots from other languages)²⁶² was very perceptible during the focus group interviews. It is often said that Esperanto speakers tend to prefer using internal resources of the language by combining existing Esperanto roots (e.g., Gobbo, 2017, pp. 6–7), and this was reflected in participants’ comments:

You should think twice about inventing new roots. But new compounds are always fun and easy to understand.

Adriano (my translation)^{lxxviii}

Adriano was by far not an isolated case: Several participants reported, as a first step, checking for existing items in some sources of their choice and, as a second step—if they had not found anything appropriate—creating new lexical items. Nikolao, for instance, also reported that he creates compounds whenever he can, but he uses foreign roots if he cannot do otherwise. Karlo reported calquing foreign lexical items if he did not find anything in his external sources:

Yes, sometimes [modern designations for new types of clothing] are missing, but in the Esperanto Wikipedia or in the dictionary *Hejma Vortaro* or even in the *PIV* dictionary, you can find interesting versions. If I don’t find anything, I search in other European languages and, possibly, I caulk the expression I liked. Sorry, “caulk” is not the right word. I calque the expression.

Karlo (my translation)^{lxxix}

.....
262 With a few exception for proper names (e.g. Facebook), Esperanto ‘domesticates’ the lexical material it takes from other languages: foreign roots are combined with an Esperanto ending. Thus, complete borrowings rarely occur, but calques and hybrids are frequent.

This once again underlines the need for language managers to act quickly—as soon as a new concept is forming—as several participants mentioned that they create new lexical items if they do not have one at hand. Language managers are expected to face much more resistance from target speakers if they try to modify the existing lexicon than if they offer to fill a lexical gap.

The point is even stronger as, for several participants, lexical creation seemed to be the last-resort strategy, to be used only if no lexical items exist. A conversation excerpt between Emilio and Vito clearly illustrates this attitude:

Until now I searched examples in Esperanto related to art (a very limited set, from my own experience) and in case I did not find anything I created my own term and added an explanation at the end of the article.

Alfredo (my translation)^{lxxx}

Alfredo, it's better to use good existing words than creating new ones and make the dictionary even bigger.

Vito (my translation)^{lxxxi}

You're right. We should not introduce new words if some already exist and are appropriate. But when they don't exist or when they're not appropriate?

Emilio (my translation)^{lxxxii}

This short conversation between three participants illustrates two focal points: 1) the need to uncover whether a lexical item already exists in the language for a specific concept and 2) the need to know whether this existing item is “appropriate.” To this end, participants reported conducting external searches for information, as I show in the next section.

6.3.2 External search for information: Diversity on the rise

Adapting Wiegand's (1998) work, three types of situations in which Esperanto speakers look for external information can be considered (pp. 551–552):

1. Speakers look for a lexical item that does not exist in Esperanto
2. Speakers look for a lexical item that exists in Esperanto but which they do not know

3. Speakers know the lexical item in Esperanto but are missing one of its characteristics in their individual lexicon or grammar

I am speaking of external search for information, borrowing a concept from marketing:

Nearly every introductory marketing and consumer behavior textbook depicts the consumer purchase decision process as a series of steps progressing from problem recognition, to information search, to evaluation of alternatives, to purchase decision, and finally to postpurchase behavior. In the information search stage, consumers actively collect information to make potentially better purchase decisions. (Schmidt & Spreng, 1996, p. 246)

By **external search for information**, I mean the situation in which a speaker decides to start a lexicographical consultation because they are experiencing a lexicographically relevant need (see Bergenholtz, Bothma, & Gouws, 2015, p. 4, on the extralexical preconsultation phase).

As previous empirical studies have shown,²⁶³ the focus groups revealed that every individual uses their own set of external resources, and they do so in a unique way, thus presenting further empirical evidence of Bergenholtz et al.'s statement regarding access to lexicographically relevant material:

It is ... evident that access routes and the different steps followed by an individual will be unique and that there is not only one possible route or set of steps. (Bergenholtz et al., 2015)

Here, the primary aim is to gather insights about the sources participants report using.

6.3.2.1 Dictionaries

Participants did report employing dictionaries. As Jan notes,

When I come across this problem, I can look up in a paper dictionary or simply on an online dictionary.

Jan (my translation)^{lxxxiii}

²⁶³ For instance Künzli in relation to sources used during the translation process (2001, p. 515).

Types of dictionaries cited included first and foremost general dictionaries (monolingual or multilingual)²⁶⁴ but also specialized dictionaries covering specific spheres of human activity (e.g., *Hejma Vortaro* for use at home, or *Komputada Leksikono* or other terminological dictionaries) and to a lesser extent, encyclopedic dictionaries.

It is essential for language managers to know if speakers among the target speech community use a dictionary and who are the ones to do so: If target group speakers generally do not use dictionaries or use them only very rarely, any efforts on the part of language managers to disseminate their target lexical items using dictionaries are condemned to remain fruitless. In the focus groups, there was absolutely no agreement on whether to use dictionaries. Participants who did not have training or experience in language-related domains (Roberto, Georgo) reported that they rarely used a dictionary at all.

As several scholars have pointed out, individuals can satisfy their need for information from sources other than a dictionary, such as newspapers, books, and online texts (Fuentes-Olivera & Tarp, 2008, p. 77; García Llamas, 2015, p. 63). In a street poll of German native speakers, Ripfel (1990) found, first, that a majority of respondents did not possess a German monolingual dictionary, and second, that a large part of this majority *did* have lexically relevant needs to which they responded using other sources of information. From what was revealed during the focus group interviews, speakers who reported rarely using a dictionary also had lexical needs, which they often solved by means other than a dictionary. Every individual and every case is different, but this still constitutes an urgent appeal for language managers to segment their targets and tailor their dissemination strategies according to the actual lexical environment of their target group segments (in which dictionaries are not necessarily present). As Drame mentioned regarding the implementation of terminology policies:

*Constant campaigning using **diverse** channels, methods, and media may be necessary. It will serve to **address the diversity of the users** with their different learning styles, physical abilities and other preferences, and the reinforcement of the information through repetition.* (Drame, 2009, p. 126, my emphasis)

264 Specific titles were sometimes mentioned according to the native ethnic language of the participant.

For those who did report using one or more dictionaries, there was no agreement as to precisely *which* types of dictionaries or which specific dictionaries. This is not surprising, as the groups were intended to be homogeneous and participants were expected to have different needs and reported behaviors, but interestingly sometimes answers regarding types of dictionaries were even complete opposites, for instance regarding the use of paper versus electronic dictionaries:²⁶⁵

I never use an online dictionary. But why not?

Emilio (my translation)^{lxxxiv}

I admit that I only use online dictionaries.

Ulriko (my translation)^{lxxxv}

This further highlights the diversity of speakers' environments and thus the necessity for language managers to reflect on disseminating strategies: target lexical items cannot be disseminated by online means for target speakers that live offline, and vice-versa. This may sound like a truism, but the case studied by Gresa Barbero (2016) is an eloquent example of language managers largely missing their target.

6.3.2.2 *Alternative resources*

When speakers conduct an external search for information, they may use traditional lexical resources such as dictionaries, as mentioned. However, external information can also include other types of sources that are not intrinsically lexicographical: fellow speakers (see Section 6.3.2.3) or what I call alternative resources (i.e., any nontraditional resource a speaker uses for responding to a lexicographically relevant need). Alternative resources can be websites, online encyclopedia (Wikipedia), search engines, etc.

If speakers do search for external information, do they do so using dictionaries? According to Lew and De Schryver (2014), "General internet search engines [are] encroaching on the grounds traditionally reserved for lexicographic

265 Here, I suspect that age might be the decisive factor, the younger generation being more prone to the use of electronic resources. At the time of the group interview, Emilio was over 60 and Ulriko under 30 years old. Fabio, also from the younger generation, mentioned that he specifically uses the ReVo dictionary simply because it has a mobile application. He implied that if the Akademio launched a mobile application, he might be inclined to use it.

queries” (p. 341), a statement that was confirmed by several participants in my focus groups. Speakers can find what they need on the Internet, as Roberto mentioned:

For modern or strange words, or strange usage, I go to Google search. Usually I find what I’m looking for on the first page.

Roberto (my translation)^{lxxxvi}

In Roberto’s case, he uses a search engine for reasons of speed—he can find what he needs quickly, but the Internet (e.g., Google or Wikipedia) is also a useful source because it can provide lexical information that *cannot* be found in dictionaries:

I often talk about computers, cars, etc. For example, I did not find a translation of “smartphone” in the *PIV* dictionary ... If I look for modern words I do it first in the German or English Wikipedia, and then I switch to the Esperanto Wikipedia to see texts on the same topic—I often found words that were not present in the *PIV* dictionary.

Adriano (my translation)^{lxxxvii}

It is not surprising that speakers look for modern words in places other than dictionaries, for traditional linguistic resources that follow theoretical models are quickly outdated, whereas other resources are frequently updated (e.g., Zesch, Müller, & Gurevych, 2008).

Wikipedia in particular was repeatedly mentioned as a useful tool to find Esperanto equivalents and was even elevated to the status of a dictionary:

Paradoxically, Wikipedia is a relatively good multilingual dictionary and a relatively bad encyclopedia in my opinion. In other words, it offers relatively satisfactory lemmas (except for proper nouns, which it doesn’t translate much) even though article contents are limping along.

Marteno (my translation)^{lxxxviii}

Participants reported using the Internet not only to find lexical items but to find information *about* lexical items. One piece of information for which they used the Internet as opposed to dictionaries was language usage:

On the Internet, you can find examples of real usage, not only those quoted in *PIV*.

Karlo (my translation)^{lxxxix}

They searched examples but also wanted some statistics (e.g., what lexical items are most used by fellow speakers). Several of them (Jan, Pablo, Bernardo), both with and without experience in language-related domains, reported using Google as a concordancer to see whether and how often a specific lexical item they have in mind was being used by other speakers. Some participants (Alfredo, Joel, Jan), also both with and without experience in language-related domains, used a concordancer proper as well, the online Tekstaro.²⁶⁶

Some external search behaviors would not necessarily come to the mind of a linguist but were mentioned relatively often by participants, such as the “Google method,” which consists of coining a new designation or finding a new designation in a resource the speaker does not consider to be fully reliable and searching it in Google to see whether it is the “right” one (the frequency of use often being the criterion for validating the new designation). This once again underlines the influence of current language usage on future language usage and further suggests that it may prove difficult for language managers to modify the existing lexicon if a lexical item other than the target lexical item is already in use.

6.3.2.3 *Fellow speakers*

Finally, participants reported using the Internet to find lexical sources and to get in touch with fellow speakers, whom they ask for advice. As mentioned in Section 2.4.3, nowadays people use technologies to get what they need from each other. This is well illustrated by a quotation from Joel:

Generally you shouldn't hesitate to ask other speakers. With the Internet this is really easy!

Joel (my translation)^{xc}

Mailing lists and social networks (specific Facebook groups or blogging platforms) were mentioned as technical means to reach other speakers for solving lexical issues, and from what participants reported, it seems that some speakers do get the help they need from each other:

.....
²⁶⁶ <http://tekstaro.com>.

Sometimes I actively search on the Internet and on the chat and sometimes I receive very good help.

Georgo (my translation)^{xci}

6.3.3 Collaborative dictionaries: Not trustworthy but useful

We have seen that participants reported resorting to traditional dictionary sources and to fellow speakers and alternative resources to respond to their lexicographically relevant needs. Would they be inclined to use dictionaries compiled partly or entirely by nonspecialists? In the heterogeneous focus groups, opinions expressed on this matter ranged from completely pessimistic to partly optimistic.

6.3.3.1 Caution with collaborative dictionaries

There were negative spontaneous reactions regarding collaborative dictionaries.²⁶⁷ The idea that professional dictionaries were generally of better quality than dictionaries that nonspecialists edited in part or in whole met the approval of several participants across groups. For instance, the idea of such a dictionary was met with strong opposition from Aleksandro, according to whom dictionaries that nonspecialists make either partly or entirely are inept (*fuŝaj*) and contain errors (*erarigaj*). According to Aleksandro, because nonspecialists edit entries—and despite their good intentions—the resulting dictionary product is full of mistakes (*plenerara*). Roberto confirmed he had seen a large number of inappropriate language equivalents (*tradukaĵoj*) in such multilingual dictionaries because “lots of things in the Esperanto community are made by volunteers who have better will than abilities.” Ulriko explained that the lack of domain knowledge could represent a major issue for specialized topics. Daniel explained that some dictionaries are faulty with regard to what a linguist would expect of them. An example he gave is that folk speakers could include words they had themselves coined in the dictionary and that these words would represent only their personal views. In this sense, the dictionary would be linguistically biased because the domain would not be approached objectively but rather from a specific person’s standpoint.

.....
267 As appears in the discussion guideline in the appendices, I did not use the word “collaborative dictionary” with participants, but rather a vague paraphrase: “dictionaries compiled partly or completely by non-specialists”. In some groups, the concept had to be set out in detail.

I think that collaboration between a lot of professionals and a few ignorant people can be a mess. A dictionary is good only if it's reliable.

Adriano (my translation)^{xcii}

6.3.3.2 *The (un)importance of dictionary authors*

The idea that dictionary making should remain a task that professionals undertake is something that both a language professional and another participant with no language-related training or work experience mentioned, and it gained the approval of additional participants. However, during the focus group interviews, it became clear that participants often had not thought about dictionary editors' competence or that they did not know who actually write the dictionaries they use. This is in accordance with Jackson's (2013) remark that participants in surveys and interviews are often unable to provide precise information about the publishers and titles of their dictionaries (p. 68).

You do not know that the editor is a layperson. But you can find that out if there's a word you understand better than the editor does. That is if you can be sure it's not a typo or a misprint.

Georgo (my translation)^{xciii}

In fact we do not know how knowledgeable volunteers who contribute to these online dictionaries really are: maybe they are very competent and well informed.

Bruno (my translation)^{xciv}

But, in fact, I use only two dictionaries ... I never thought about it! Now that I am starting to think about it, maybe the dictionary in Esperanto, my native language, would be more professional because it's a paper dictionary. But I never thought about the abilities of the person(s) who wrote it.

Grigorij (my translation)^{xcv}

6.3.3.3 *Use of resources that meet lexical needs*

Not knowing the editor, or knowing that the editor is not a professional did not refrain some participants from using a collaborative dictionary. Some participants

even reported that this piece of information was not important to them as long as the dictionary met their needs.²⁶⁸ Reactions included the following:

I admit that I do not pay attention to this either: I check whether a particular word is appropriate for me.

Ulriko (my translation)^{xcvi}

When I choose a dictionary, I am more concerned about seeing whether the dictionary can meet my needs ... But not every dictionary can meet my needs, which is why I own several dictionaries.

Jan (my translation)^{xcvii}

It thus seems that some of the participants use such dictionaries exclusively for a practical purpose and not as a reference. This is in line with the evolution that Lew and De Schryver described:

For many centuries, dictionaries were viewed with authority, often admired and revered with awe, and the status of ‘the dictionary’ in some countries could be likened to that of the lay Bible ... As dictionaries moved from the bookshelves gradually onto floppy disks, optical disks, internet servers, and now mobile devices, they found themselves as it were in the same league as utility and productivity software, which in turn encouraged a more pragmatic and less ideological or dogmatic view of dictionaries. This trend was only strengthened as users themselves started getting involved in bottom-up dictionary-making. As a result of these developments, dictionaries—which have always been inherently practical—have now come to be recognized as even more practical. (Lew & De Schryver, 2014, pp. 341–342, my emphasis)

Some participants mentioned that they would not be reluctant to use such a dictionary if, for instance, professionals checked nonspecialist contributions and/or if their contributions were clearly indicated as such. Several participants reported that a dictionary does not necessarily have to be completely reliable. In addition, it was mentioned that reliability may depend on how the dictionary editing process is organized: for instance, if nonspecialists make contributions that more

.....
268 There was no agreement on this aspect and other participants mentioned they would ideally like to know who edits their dictionaries.

experienced people check, dictionaries could be reliable. Opinions ranged from the need for a dictionary to be reliable, to the more moderate view that a dictionary can be read with a critical eye and used as a reservoir of ideas rather than as a resource that one should trust beyond doubt:

One needs to approach these dictionaries with a particularly critical mind (or assuming that they cannot be trusted, like someone just said). It is important to know the language well so that you don't pick up a mistake from the dictionary. "To ask a good question, one needs to know most of the answer."

Joel (my translation)^{xcviii}

The Wiktionary can be a good thing; I often use it (for languages other than Esperanto), but, of course, it must be used with a critical mind.

Pablo (my translation)^{xcix}

For several speakers, it seemed possible to use collaborative dictionaries in some given situations, and they mentioned that a collaborative dictionary can even prove to be useful in restricted contexts. It can be used, for instance, if the topic is not controversial, if one needs only a general idea about something, or if it is not about a (highly) specialized topic. Moreover, again, several participants were expecting to find in such nontraditional resources useful information that they would not have found in a traditional dictionary:

The plus with such "dictionaries" (or searching a word with Google) is the modernity and timeliness of the words.

Oskaro (my translation)^c

6.3.4 When speakers turn their filters on

As discussed earlier, participants did report using, at times, external sources. Several of them mentioned that they adopt comparison strategies (especially in the context of written communication). Some also reported a hierarchized search behavior; for instance, Alfredo mentioned first trying to check resources and creating a new lexical item only if he does not find anything relevant, and Adriano mentioned first checking paper dictionaries and then consulting online resources afterward.

More importantly, not only did participants mention that they cross-compare external resources but also several of them reported that once they have gained information from sources, they want to decide for themselves which lexical item is “the best one” or “the right one.” Thus, the sources that the participants reported using are not necessarily used as reference resources. Some speakers use their dictionaries as reservoirs of ideas, and they judge whether the contents are good enough to be used. If they feel that the dictionary is mistaken or that it does not propose an adequate solution (a solution they like or approve of), they feel entitled not to use what they have found in the resource.

Sometimes you notice that an Esperanto translation in a dictionary is just something contrived from the editor. In such cases, I feel that I have the right to think whether another solution would be more appropriate; that is a solution that I myself think up.

Bernardo (my translation)^{ci}

Bernardo was not alone. A host of participants reported filtering the external information they find, even in major dictionaries, such as the *PIV* dictionary:

PIV is not an absolute dogma. You should use it to check, but also consider other sources (not only dictionaries but also real usage).

Damjan (my translation)^{cii}

I don't hesitate not to use a word from the dictionary if I think it's not appropriate ... So, I don't “trust” dictionaries at all.

Ulriko (my translation)^{ciii}

Speakers reported a large range of behaviors for coping with situations in which they need lexical items in Esperanto. Consulting a dictionary is only one of many solutions, and it is by far not always the solution chosen. The path taken seems to largely depend on the speaker and on the communicative context, but more importantly, many speakers reported that they keep filtering the external information they gain, whatever the source. This suggests that even if language managers succeed in disseminating their target lexical items to target speakers (through a dictionary or another resource), the target speakers might still filter them. How does this filtering happen? This is what Chapter 8 investigates, after I develop a proof of concept for observing target speakers' opinions in Chapter 7.

6.4 Summary

In summary, the focus groups showed the following:

- In a host of cases, an external search for lexical information does not occur, which leads to such strategies as the use of paraphrases or lexical creation on the part of speakers.
- Participants may check for existing lexical items before they coin new ones (last-resort strategy), but they may create new ones if they feel the need for lexical items.
- Each participant has his or her own approach to the use of external resources (unique access routes): approaches may be contradictory (e.g., a participant who never uses an online dictionary versus a participant who uses only online dictionaries), and they are organized according to a specific hierarchy.
- Participants use both print and online dictionaries, although two participants who do not have training or experience in a language-related domain mentioned that they rarely use a dictionary.
- Participants who do not use dictionaries use other strategies for responding to their lexical needs.
- Participants use alternative resources: for example, search engines or Wikipedia, for reasons of speed but also because traditional resources, such as dictionaries, may not contain the types of information they are looking for (for example, neologisms, language usage, frequency of use).
- Some participants use the “Google method” to validate lexical items: if the frequency of the item is deemed satisfactory, the lexical item is thought to be the correct one.
- Some participants do not hesitate to ask fellow speakers for information or opinions.
- The range of opinions among participants about collaborative dictionaries was from completely pessimistic to partly optimistic.
- Participants had not systematically thought about who had authored their dictionaries, but they were critical of the content.
- The dictionary does not necessarily need to be reliable: for some participants, the only criteria that seem relevant when they are deciding to use a dictionary or another lexical resource is whether this resource meets their needs.

- Some participants use lexical resources as reservoirs of ideas rather than as reference resources.

Chapter 6 was devoted to speakers' lexical environments. The next two chapters explore speakers' lexical opinions. Chapter 7 is a methodological chapter that describes a proof of concept for observing speakers' lexical opinions in context. It explains the aspects related to corpus compilation, the detection of metalanguage and autonymy, the filtering of relevant units, and the analysis of metalinguistic and autonymical units. Thereafter, Chapter 8 presents the actual results obtained via analyzing the metalinguistic statements with an opinionated autonym found in the corpus.

7 **Methods: A proof of concept for detecting opinionated autonyms**

7.1 **Introduction**

In Chapter 3 (Section 3.4.2), I explained that language managers could take advantage of the clues that individuals leave in speech about their lexical opinions, especially on the Web. Just as marketers and clinicians use sentiment analysis (opinion mining) to systematically extract and study subjective information from individuals, I suggest that it should be possible to systematically monitor ordinary speakers' subjective opinions regarding lexical items. The present chapter is a proof of concept for validating this idea.

In my review of previous studies on lexical opinions in Section 3.3.2, I mentioned that very few scholars have collected real language data concerning lexical criteria but that such data could prove useful in better planning lexical interventions. Thus, the goal of my proof of concept is to lay the foundation for and demonstrate the feasibility of monitoring lexical criteria in context based on natural language processing, combining the two dimensions of opinion and autonomy. Because the present investigation is a first step in this direction, some of its tasks are performed manually, but with the quick evolution of technologies, they could be partly or fully automated in the future.

In this chapter, I present the proof of concept and apply it to a corpus collected within five networks of practice (Chapter 5) of the Esperanto speech community (for information on the community, see Chapter 4) to extract opinionated autonym candidates for a qualitative analysis conducted in Chapter 8. I start by exploring the existing theoretical foundations for the detection of opinions and autonyms in the corpora (7.2). I then explain the choices made in corpus compilation (7.3). Next, I observe (7.4) contexts containing lexical criteria in my corpus to determine which features represent opinion and autonomy in Esperanto. Based on the features identified, I suggest and evaluate indicators (Section 7.5) for detecting opinionated autonyms in Esperanto. Subsequently (Section 7.6), I combine indicators for extracting opinionated autonym candidates from the corpus, extracting from the entire corpus (69,792 contributions) a set of 4,090 opinionated autonym candidates that can be used for qualitative analysis in the next chapter. Because this is a first try for such a framework, I present in the last section (7.7) the limitations of my approach as well as ideas for improving the detection results.

7.2 Theoretical foundations for the detection of opinions and autonyms based on natural language processing

7.2.1 Opinionated autonyms in corpora

In Section 3.4, I called explicit metalinguistic statements “sentences where discourse reflects upon itself, where language itself is the subject, and where language is creating and manipulating the elements and rules that make it possible,” borrowing the notion and its definition from Penagos (2004b, p. I-4–I-5).

Both a terminological clarification and a conceptual restriction are needed here, as explicit metalinguistic statements can serve multiple functions in language. They can, for instance, serve *to agree on the code*, fulfilling an explanatory function “whenever the addresser and/or the addressee need to check up whether they use the same code” (Jakobson, 1980, p. 86) as illustrated here:

The characteristic syndrome associated with lumbar stenosis is termed neurogenic intermittent claudication. (Rodríguez Penagos, 2004b, p. V–138)

Here we report that activation of Rap by forskolin and cAMP occurs independently of protein kinase A (also known as cAMP-activated protein kinase). (Rodríguez Penagos, 2004b, p. 137)

Explicit metalinguistic statements can also serve *to teach the code* (let us think, for instance, of our foreign language teachers in school), or to negotiate the code, serving a didactic purpose²⁶⁹ in Delavigne’s terms (2000).

In Section 3.4, I also talked about the idea that speakers have views about the lexicon and that they may, under certain conditions, express these views using metalanguage. In this situation, the views become observable for the linguist in real-language data. Among their explicit metalinguistic statements, speakers may notably express judgements about the adequacy of given lexical items in the autonymical condition (Sitri, 2003, p. 205). In this context, their explicit metalinguistic statements serve *to evaluate the code*, as was illustrated in Saint’s paper (2016) based on Twitter data:

269 French: visée didactique.

Heard on Radio-Canada as a replacement for “cuisinomane”: “gastro-naute.” It’s a bit long, but it’s an interesting option! #foodie [my translation, the bold emphasis is mine]^{civ}

Here, the speaker expresses his or her views on the lexical item “gastro-naute.” In the present investigation, I call such evaluative explicit metalinguistic statements **metalinguistic statements with an opinionated autonym**.

The opinion²⁷⁰ in an metalinguistic statement with an opinionated autonym can be directed toward a general trait of (a) language or how a speaker uses (a) language. In the present investigation, I am particularly interested in a metalinguistic statement in which the opinion is directed toward a lexical item in the autonymical condition, as is the case in the example presented earlier. I call such statements **metalinguistic statements with an opinionated autonym**, and the autonym about which an opinion is being expressed is an opinionated autonym. This is necessary to mark the distinction among the types of metalinguistic statements—those containing or not containing opinions, and those containing or not containing autonoms—as illustrated via sample contexts of my corpus in Table 15.

Metalinguistic statements with an opinionated autonym are characterized by the presence of both an opinion and a lexical item in the autonymical condition. It follows that detecting contexts with opinionated autonoms can be decomposed into a two-step binary classification task—for example, assessing (a) whether a given context contains an opinion and (b) whether this same context contains an autonym. As Schapire and Singer mentioned, such a decomposition is common in natural language processing classification tasks:

While numerous categorization algorithms [...] can be adapted to multi-label categorization problems, when machine-learning and other approaches are applied to text-categorization problems, a common technique has been to decompose the multi-class, multi-label problem into multiple, independent binary classification problems (one per category). (Schapire & Singer, 2000, p. 136)

This is the approach adopted here. Indicators are constructed and evaluated separately for opinion and autonymy (Section 7.5), and then, they are recombined (Section 7.6) for extracting opinionated autonym candidates.

.....
270 An operational definition of opinion is proposed in Section 6.4.1.

By **classification**, I mean the “action of arranging a whole set into ... existing classes” (Nakache, Metais, & Timsit, 2005, p. 418) as illustrated in Figure 14.

Subset	Opinion	Metalinguage	Autonymy	Sample contexts from my corpus	Types of metalinguistic statement (MS)
a.	yes	no	no	<i>I really regret this useless discussion proposed by Mr. [Name]...^{cv}</i>	- (not an MS)
b.	yes	yes	no	<i>I don't like roots with double vowels, either, and I think they shouldn't exist in Esperanto.^{cv}</i>	Opinionated MS without autonym
c.	yes	yes	yes	<i>I never used the word “brodkasti” because, to me, it sounds like an Anglicism that should be avoided.^{cvi}</i>	Opinionated MS with autonym
d.	no	yes	no	<i>In Russian and in Mongolian the middle consonant is voiced^{cvi}</i>	Non-opinionated MS without autonym
e.	no	yes	yes	<i>In Portuguese there are no separate words for NERD and GEEK.^{cix}</i>	Non-opinionated MS with autonym
f.	no	no	no	<i>[Name], who lived a long time in Tashkent, could most likely get a close look of less exotic varans there.^{cx}</i>	- (not an MS)

Table 15. Sample contexts from the corpus illustrating types of metalinguistic statements in relation to opinion and autonymy and delimiting the notion of metalinguistic statement with an opinionated autonym in Subset C.

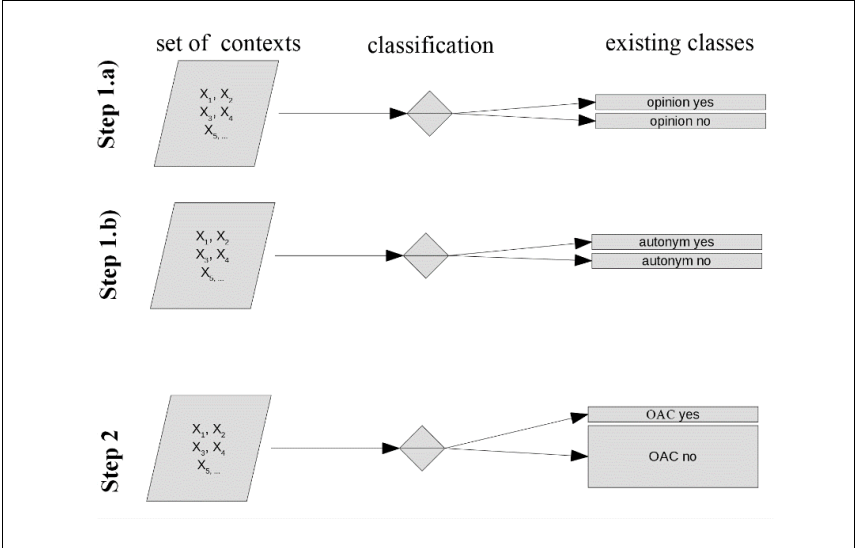


Figure 14. Classification of contexts with opinions and autonyms using a two-step binary classification for identifying opinionated autonym candidates (OAC).

A classification can be done manually or automatically. Here, the goal is to automate, to the extent possible, the classification of contexts using natural language processing.

To evaluate the performance of a natural language processing method of classification, one “measur[es] the difference between result and requirement” (Nakache et al., 2005, p. 418). To represent requirement in the present investigation, a test set (Section 7.3.3) was built with manually annotated contexts. No metric is intrinsically associated with evaluation, but several standard measures exist (Nakache et al., 2005, p. 418). Sokolova and Lapalme (2009) mentioned that

The correctness of a classification can be evaluated by computing the number of correctly recognized class examples (true positives), the number of correctly recognized examples that do not belong to the class (true negatives), and examples that either were incorrectly assigned to the class (false positives) or that were not recognized as class examples (false negatives). (p. 429)

The performance of binary classification tasks is typically evaluated using a table of counts or a contingency table (Manning & Schütze, 1999, p. 577) such as the following one:

	YES is correct	NO is correct
YES was assigned	(a) true positive	(b) false positive
NO was assigned	(c) false negative	(d) true negative

Table 16. Contingency table for evaluating the performance of binary classification tasks.

From such a table, several evaluation measures can be calculated, notably:

- The proportion of correctly classified objects, called **accuracy** (Manning & Schütze, 1999, p. 577):

$$\frac{a + d}{a + b + c + d}$$

- The proportion of relevant items within the positive set, called **precision** (Manning & Schütze, 1999, p. 534):

$$\frac{a}{a + b}$$

- The proportion of all relevant contexts in the entire corpus that has been included in the positive set, called **recall** (Manning & Schütze, 1999, p. 534):

$$\frac{a}{a + c}$$

- The harmonic mean of precision (P) and recall (R), called **F-measure** or **F1-measure**:²⁷¹

$$\frac{2PR}{P + R}$$

From these fundamental evaluation measures, in Section 6.5, I used specifically precision and recall for evaluating indicators and classification tasks.

271 The F-measure is a standard measure for capturing both precision and recall in one figure only (Castillo, Donato, Gionis, Murdock, & Silvestri, 2007).

7.2.2 Building and manually annotating a test set

To evaluate the performance of indicators and classification tasks, I constructed a test set representing the requirement. This test set was built from all five sub-corpora with 1566 randomly selected contexts.²⁷² These contexts were manually classified according to the types of metalinguistic statements presented above in Section 7.2.1 (p. 211).²⁷³ The following figures resulted from the manual classification task:

Type of metalinguistic statement (MS)	Opinion	Metalinguage	Autonymy	Astronomic Terminaro	Esperanto-tradukistoj	Lingva konsultejo	Retposhta rondo...	VIVO-vikio	TOTAL
a. (not MS)	yes	no	no	6	21	42	20	0	89
b. <i>Opinionated MS without autonym</i>	yes	yes	no	11	38	27	47	0	123
c. <i>Opinionated MS with autonym</i>	yes	yes	yes	51	142	83	129	45	450
e. <i>MS without autonym</i>	no	yes	no	9	41	49	20	12	131
d. <i>MS with autonym</i>	no	yes	yes	30	96	83	57	16	282
f. (not MS)	no	no	no	147	31	97	105	11	491
<i>TOTAL</i>				254	369	381	378	184	1566

Table 17. Results of manual classification for the test set.

In the test set, about a third of contexts (450 out of 1566) are metalinguistic statements with an opinionated autonym. The goal here is to correctly identify contexts fitting into this category using natural language processing.²⁷⁴

272 Where n , the number of contributions selected, is the required sample size for binary data with a 5% margin of error (z-score of 5%) and a 95% confidence interval. Randomness was obtained using the random function, or RAND, in Microsoft Excel.

273 This task was undertaken by myself using Knowtator.

274 Needless to say it is unrealistic to fully achieve it in the present investigation: this chapter tries an approximation towards this goal and serves as a proof of concept for further investigations.

7.2.3 Adapting the Metalinguistic Operation Processor

To automate the classification of contexts, I adapted an existing program: the metalinguistic operation processor developed by Penagos (2004b).²⁷⁵ The metalinguistic operation processor is an application programmed in Python for extracting metalinguistic information from natural language corpora. It identifies explicit metalinguistic statements in English-language scientific papers and technical documents. I used its extraction module (extract.py):

The candidate extraction module performs a search and filtering process to locate and select from the normalized file the metalinguistic sentences that will merit further processing. (Rodríguez Penagos, 2004b, p. IV–97)

This module uses a pattern list for extracting candidate sentences. The patterns can be formulated using regular expressions, which is a definite advantage for an agglutinative language, such as Esperanto (see example in the appendices). The pattern list can also contain multi-word expressions.

Two major adaptations were made to Penagos’s candidate extraction module. First, the pattern lists were adapted to my context; for example, patterns were developed for Esperanto starting from zero and were elaborated with a new type of text in mind (folk language on the Internet rather than scientific and technical documents). Penagos “selected lexical patterns that could be indicators of metalinguistic activity in text, and expanded the list to 116 different patterns using other plausible verbal forms, as well as lexical items and nominal modifiers, such as *term*, *word*, *phrase*, *vocabulary*, *terminology*, etc., that could indicate that the sentence was metalinguistic in nature” (Penagos, 2004, pp. I–17). Comparable patterns had to be found for Esperanto (Section 7.5). In addition, folk speakers may use metalinguistic language that would not be necessarily expected in grammar works; for instance, a dictionary’s name (in the following example, *PIV*) can be turned into an adjective describing an autonym in folk language:²⁷⁶

Se oni kontentas pri la PIV-a "volano" kaj "volanludo", tiam la nomo estu "badmintono", kiu estas internacia.

275 The author kindly provided us with the full program he wrote in Python.

276 Here and in other parts of the investigation, I quote the Esperanto original corpus, to which I add my English translation. Sometimes, the original corpus contains mistakes (e.g. typos), which are repeated here to accurately reflect the original corpus.

If you're satisfied with the “volano” and “volanludo” in PIV, then the international name must be “badmintono”.

Second, programmatic modifications were made: the module was adapted for classification rather than extraction. The resulting classification module takes single contexts as inputs and operates a binary classification based on a given pattern list. It generates a list of positively and negatively classified contexts (see Figure 47 in the appendices for an illustration of this process). In addition, a short evaluation module was written in Python to assess the performance of classification tasks. It takes as inputs the list of positively and negatively classified contexts from the automated classification as well as a similar list elaborated during manual classification (for an illustration, see Figure 48 in the appendices).

With such tools, it becomes possible to use indicators in the form of regular expressions to automatically classify a context into one of the two categories, for example, “autonym yes” or “autonym no” (and “opinion yes” or “opinion no”). An example is provided in the appendices. In running the classification module on the test set and the subsequent evaluation module, one can determine the percentage of relevant contexts that a particular indicator obtains (see the appendices for an example).

7.3 Corpus compilation

7.3.1 Guiding principles

A **corpus** is a collection of text data that are “empirical, analyzing the actual patterns of use in natural texts” (Biber, Conrad, & Reppen, 1998a, p. 4). In his seminal monograph, Biber explained that compiling a state-of-the-art corpus implies that various aspects have been reflected on beforehand:

Some of the first considerations in constructing a corpus concern the overall design: for example, the kinds of texts included, the number of texts, the selection of particular texts, the selection of text samples from within texts, and the length of text samples. (Biber, 1993, p. 242)

In effect, no single way of compiling and using corpora exists, and the design of a corpus must be guided by the research direction:

Corpus linguistics is not a monolithic, consensually agreed set of methods and procedures for the exploration of language. [...] The importance of our findings from a corpus, whether quantitative or qualitative, depends on another general factor which applies to all types of corpus linguistics: the corpus data we select to explore a research question must be well matched to that research question. (Macenery & Hardie, 2012, pp. 1–2)

Here, I examine a typical applied research problem: finding the (most efficient) way for language managers to reach (partial) implantation under the condition of structural uncertainty, in which they must make decisions quickly (see problem statement in Section 3.2). In this context, the corpus must be designed to serve as a tool for supporting swift decision-making, wherefore speed is a major design factor.²⁷⁷ The first choice I made in this investigation was to restrict it to data available on the Internet, as these data can be quickly and automatically retrieved. Currently, collecting offline language data (conversations, written data in print or PDF format, etc.) is time consuming and does not meet the design priority of speed.²⁷⁸ In addition, my corpus is intended to be a *monitor corpus*. A **monitor corpus** is not a static collection of language data but rather is a dynamic, open-ended entity (McEnery & Wilson, 1997, p. 22). Monitor corpora have been used, for instance, among lexicographers for trawling streams of new texts in search of neologisms (McEnery & Wilson, 1997, p. 22), and they follow the evolution of specific language phenomena (Habert, 2000, p. 13). I adopted the idea of monitoring the opinions of individuals on specific lexical items in the autonomous condition with the aim of allowing for faster feedback on (new) lexical items. Thus, the number of texts is not intended to be limited: for the present investigation, corpus collection started from the first contribution available in each network and lasted until the date of data collection.

As far as the selection of texts is concerned, Biber (1998b) and his colleagues further told us that

277 In a comparable approach, opinion mining applications (e.g. Hu & Liu, 2004), for instance, often focus on online data that can be quickly obtained in large quantities.

278 To a certain extent, this constitutes a bias, because off-line language data (such as everyday conversations, debates during Esperanto meetings, etc.) are being completely discarded.

a corpus is not simply a collection of texts. Rather, a corpus seeks to represent a language or some part of a language. The appropriate design for a corpus therefore depends upon what it is meant to represent. (p. 246)

In the present investigation, the aim of the corpus is to shed light on the large variety of explicit metalinguistic statements²⁷⁹ that speakers may make about lexical items. The types of texts included are therefore ones that can be expected to have a high density of metalinguistic statements with an opinionated autonomy, with the general aim of seeking to represent the spectrum of possible cases from a qualitative perspective. As highlighted in Chapter 4, the Esperanto speech community and its electronic networks of practice constitute a useful starting point. Five online networks were selected (for details, see Chapter 5).

7.3.2 Corpus collection

First and foremost, a remark regarding confidentiality and ethical principles is necessary when it comes to the processing of information obtained from the Internet. Generally, the boundaries between public and private spaces in Internet-based research are hard to define. For this investigation, one of the main difficulties lies in the fact that members of electronic networks of practice usually consider their contributions to be private, although almost anyone connected to the Internet can access them.²⁸⁰ As a consequence, these individuals do not expect to be objects of study and are not aware that their comments are being analyzed. The present investigation circumvents this issue by being exclusively centered on lexical items and by anonymizing all of the data used or quoted, as has been done in comparable research (Saint, 2016). To protect the privacy of contributors, I do not construct any individual profiles, and I do not link to any of their contributions.

Although the corpus de facto discards offline data, it does seek to be representative of the metalinguistic statements with an opinionated autonomy that Esperanto speakers make online.²⁸¹ According to Biber (1993), **representativeness**

279 Refer to 3.4 for a definition (p. 114).

280 Three of the networks in the present investigation require registration and can thus be considered to be semi-private rather than public.

281 One should investigate whether statements made online are generally representative of statements made in the target speech community (i.e. in offline interactions), but this is out of the scope of the present investigation.

“refers to the extent to which a sample includes the full range of variability in a population” (p. 242). The corpus addresses this variability by combining subcorpora from groups concerned directly with dictionary compilation (both for general and specialized languages) and more general, punctual problem-solving groups. Five subcorpora were built from data collected from the five electronic networks of practice presented in Chapter 4. The sizes of the corpora varied sharply as shown in the following table:

	Astronomia Terminaro	Esperanto-tradukistoj	Lingva konsultejo	Retposhta rondo...	ViVo-vikio
Source type	Yahoo group	Yahoo group	Facebook group	Yahoo group	dedicated web interface
Retrieval tool	PGOffline 4	PGOffline 4	Python SDK for Facebook's Graph API	PGOffline 4	webmaster
Time frame	2000–2012 ²⁸²	2003–2013	2011–2017	1999–2013	2010–2013
Format	db3	db3	txt	db3	xml
Contributions ²⁸³	747	9043	37,019 ²⁸⁴	22,110	873
Tokens	~332,000	~832,000	~897,000	~167,000	~12,000

Table 18. Figures for the five subcorpora Astronomia Terminaro, Esperanto-tradukistoj, Lingva konsultejo, Retpoŝta rondo ... and ViVo-Vikio.

Because they came from different sources, the five subcorpora had three formats (db3 database entries, xml file, and text files), which were all converted to raw text format (.txt) for processing. This was and will remain a challenge for language managers: the online technologies that people use to connect with one another are ever shifting and ever growing. What is today a dominant trend could be outdated tomorrow:

Online, people can switch behaviors as soon as they see something better. It's the force of these millions of people, combined with the rapid evolution of new technologies by trial and error, that makes the groundswell so protean

-
- 282** The discussion list was active mainly within the timeframe 2000–2002, although a few messages were sent up to the year 2012.
 - 283** The figures provided are here for informative purposes only, as the study is qualitative in nature.
 - 284** This subcorpus contains only text: Elements such as Facebook polls or images were not collected.

in form and so tough for traditional businesses to deal with. (Li & Bernoff, 2011, p. 12)

This is clearly visible in our subcorpora. For instance, the Yahoo group *Esperanto-tradukistoj* had as many as 1426 contributions in 2008 but became less active over the years. In 2016, it gathered only 145 contributions in total. Conversely, the Facebook group *Lingva Konsultejo*, which was completely inexistent in 2008, is booming at the moment, with 6000 contributions per year on average. For monitoring purposes, language managers must therefore be ready to quickly shift from one technology/format to another and follow speakers where they are active. Here, I employed a free client for downloading messages from Yahoo groups (PGOffline 4), I used a Python module for connecting to Facebook's graph API, and I also worked directly with a webmaster for one of the subcorpora.

7.3.3 Corpus cleaning

As Habert et al. (1997) mentioned, for a monitor corpus, it would be ideal to develop filters that easily clean any new text (p. 162). In the present investigation, these cleaning "filters" were partly automated. The initial corpus cleaning phase is crucial, although it is often underestimated (Habert et al., 1997, p. 161). Here, accented characters and redundant passages especially required my attention.

Over the years, Esperanto has used various writing systems for accented letters (for a detailed description, see Haszpra, 2001). Although Esperanto has been included from the start in the Unicode standard,²⁸⁵ the use of accented characters was and still is not systematic. Typing such characters may have been difficult because a standard keyboard layout was long missing. Moreover, despite Unicode standards, the Yahoo groups from which three of the subcorpora originated did not allow for the use of accented Esperanto characters. Speakers employed various strategies for replacing accented characters. Although the popular variant of using an "x" after the letter that needs a diacritic could easily be handled programmatically,²⁸⁶ the official use of "h," which is ambiguous, and a variety of other strategies

.....
285 The 5 circumflex small and capital characters *ĉ, ĝ, ĥ, ĵ* and *ŝ* (U+0108, U+0109, U+011C, U+011D, U+0124, U+0125, U+0134, U+0135, U+015C, U+015D) and the breve one *ŭ* (U+016C, U+016D) were included in the very first Unicode 1.0.0 version (see The Unicode Consortium, n. d.).

286 Because 'x' is not part of the Esperanto alphabet, a straightforward search and replace function can be programmed. Other techniques suffer ambiguity and are therefore more difficult to programmatically deal with.

that speakers employed (e.g., apostrophe, circumflex, asterisk, the letter “w” [both before and after the character that needed to be accented], or another accent on the letter)²⁸⁷ required some manual work, not to mention that speakers at times simply used a nonaccented character instead of the accented one.²⁸⁸ Fortunately, with advances in technology, Esperanto keyboards are becoming paramount (on computers, smartphones, etc.). This difficulty can be expected to significantly decrease if not disappear in the future. The corpus was consequently pre-edited to be Unicode compliant and to display the appropriate accented characters. This is crucial for the subsequent extraction of opinionated autonym candidates based on lexical patterns (Section 7.6).

Another issue was the removal of redundant passages. In electronic networks of practice based on email technology, speakers often quoted chunks of previous emails, for example, for responding to a specific idea or comment, or they used the email reply function that included the messages from the senders. Fortunately, because quoted passages were marked with a specific html tag in the subcorpora collected for the present investigation, a programmatic filter could be applied using regular expressions. This problem could also be expected to disappear over time thanks to the use of new technologies: on Facebook, for instance, speakers write short messages (24 lexical items on average in the *Lingva Konsultejo* subcorpus) and generally do not quote previous contributions verbatim.²⁸⁹

7.3.4 Morphological tagging and pre-editing

Once cleaned, the corpus was analyzed morphologically, tagged with morpheme boundaries, and pre-edited. This was useful given the nature of the Esperanto language and the tasks I wished to accomplish with my corpus. To this end, I used ESPSOFT, an application that Witkam (2010) programmed in Visual Basic, which is open source and freely available.

Despite the unchangeability of Esperanto morphemes, morphological analysis raises ambiguity issues. I do not discuss the details here,²⁹⁰ but I mention that due to these ambiguities, some manual work was needed after the automated morphological analysis. The following manual editing was done on the corpus:

.....

287 E.g. ‘law mi’, ‘laù mi’ ‘anta*u’, ‘dan*gera’, ‘^gentila’, ‘amba^u’.

288 E.g. ‘antau’ for ‘antaŭ’, ‘pseuda’ for ‘pseŭda’, ‘tauga’ for ‘taŭga’, etc.

289 Rather, they use Facebook’s reply function.

290 E.g. Witkam (2007), Bick (2016, p. 1075–1076) and Guinard (2016) discuss these ambiguities and suggest solutions.

- Tag correction: Incorrectly tagged lexical items were manually assigned the correct tags; for example, “korelativo” was tagged “kor-e-lat-iv-o” (five morphemes) via the ESPSOOF analyzer but was manually corrected to “korelativ-o” (two morphemes).
- Additional tagging: Lexical items that the ESPSOOF analyzer could not handle were tagged manually; for example, “zamenhofajn” as an adjective based on a proper noun was not recognized and thus was tagged manually as “zamenhof-ajn.”
- Error editing: Obvious typos that the analyzer identified were corrected—for instance, “publigikita” instead of “publikigita,” “atntigi” instead of “atentigi,” etc.

7.3.5 Choosing units of analysis

The analysis of electronic networks poses two main methodological challenges (Akrich, 2012, p. 4). The first is the great size of the data collected. This was solved in the present investigation by filtering relevant contexts using a semiautomated methodology as presented later (Section 7.6). Second, a message in an electronic network can be seen as one of two things: either a whole of contents, or a unit in relation to other units. In other words, one can study either the contents of messages or the networks of relationships that result from these messages. In the present investigation, I considered the electronic networks to be a medium of communication and adopted a discourse-centered approach. My analysis focused on the exchanged contents within the networks, not on the networks themselves, nor on their functions, their characteristics or processes, or the relations between members or sent messages. This is because I was not specifically interested in the dynamics of the exchanges but rather in their contents. Therefore, contributions were the units of analysis in my investigation. By **contribution**, I mean an email in a group mailing list, a post in a social network, or a comment on an interface: in concrete terms, a single post, whether an original post or a reply to a post, in *Lingva konsultejo* (Facebook), a comment in the discussion section of *ViVo-Vikio*, or an email in *Astronomia Terminaro*, *Esperanto-tradukistoj*, or *Retpoŝta rondo* (Yahoo groups).

7.4 Grammatical features of opinionated autonyms in Esperanto

The present investigation concentrates on the contexts of Type c., that is metalinguistic statements with an opinionated autonym (Section 7.2.1). Previous works²⁹¹ have shown that both opinions and lexical autonyms may be flagged in speech. The next two sections, respectively, deal with the flagging of opinion (Section 7.4.1) and that of autonymy (Section 7.4.2) in written language.

7.4.1 Opinion indicators in written language

First and foremost, it is appropriate to provide an operational definition of opinion. In Section 3.3.2, I defined **opinion** as an umbrella term used to aggregate attitudes, sentiments, feelings, emotions, and preferences. Here, I am interested in opinions that are explicitly expressed in language.

Speakers use language for various purposes, including for expressing emotions,²⁹² opinions, sentiments, evaluations, and attitudes, for example, toward “products, services, organizations, individuals, issues, events, topics” (Liu, 2012, p. 6). Natural language processing (under various terms, including *opinion mining* and *sentiment analysis*) has shown interest in opinionated text contents in corpora for more than a decade:²⁹³

Sophisticated language processing in recent years has made possible increasingly complex challenges for text analysis. One such challenge is recognizing, classifying, and understanding opinionated text. (Kim & Hovy, 2004, p. 61)

The main tasks of opinion mining are “(1) to find product features that have been commented upon by reviewers and (2) to decide whether the comments are positive or negative” (Ding, Liu, & Yu, 2008, p. 231). The goal in particular is to discover all opinion quintuples (entity or object, opinionated aspect of this entity,

.....
291 Appropriate bibliographical references are indicated in the following sections, respectively for opinion and autonymy.

292 See e.g. Jakobson’s emotive function: “The so-called EMOTIVE or “expressive” function, focused on the ADDRESSER, aims a direct expression of the speaker’s attitude toward what he is speaking about.” (Jakobson, 1960, p. 354)

293 The growth of the opinion mining field is linked with the rapid growth of social media and networks on the web (B. Liu, 2012, p. 5).

orientation of the opinion, opinion holder, and time of opinion) in a given text document (Liu & Zhang, 2012, p. 418). In my framework, I simplify this quintuple and propose detecting contexts in which an opinion is expressed, regardless of the opinion orientation.

According to Liu (2012; see also Liu & Zhang, 2012, p. 424), the most important indicators of sentiments are opinion-bearing lexical items, for example, lexical items “that are commonly used to express positive or negative sentiments” (p. 12).

For example, beautiful, wonderful, good, and amazing are positive opinion words, and bad, poor, and terrible are negative opinion words. [...] Apart from individual words, there are also opinion phrases and idioms, e.g. cost someone an arm and a leg. (Liu & Zhang, 2012, p. 423)

From a grammatical viewpoint, research has shown that in particular, adjectives are central indicators of opinions (Liu & Zhang, 2012, p. 423). However, “verbs and noun can be used to express opinions as well, e.g., verbs such as ‘like’ and ‘hate,’ and nouns such as ‘junk’ and ‘rubbish’” (Ding et al., 2008, p. 234). Articles on opinion mining and sentiment lexicons abound for English, less so for other languages,²⁹⁴ and to the best of my knowledge, no one has worked so far on opinion-bearing lexical elements in Esperanto. Consequently, I propose providing an overview of types of lexical elements that are flagging opinions, based on observations from my corpus.²⁹⁵

7.4.1.1 *Opinion-bearing content morphemes*

To begin with, I must mention that Esperanto should take an approach largely different from that of English, as the language is completely agglutinative in nature (Schubert, 2015, p. 2216), and its content lexical items are generally complex.

In Schubert’s classification (1993, p. 222), Esperanto has seven types of morphemes:²⁹⁶ roots, prefixes, prefixoids, suffixes, suffixoids, endings, and declension

294 See e.g. the challenge of creating a sentiment lexicon for Portuguese (Souza, Vieira, Buseti, Chishman, & Alves, 2011).

295 Within the framework of the present investigation, opinion indicators are restricted to lexical elements, although I observed other indicators in my corpus, e.g. the use of the subjunctive mood to express an opinion.

296 I will use Schubert’s classification here, although other classifications are possible. Blanke (1981, p. 27), for instance, proposed the three main categories basic morphemes (German: *Grundmorpheme*), word formation morphemes (German:

morphemes. In Esperanto, content lexical items *must* end with an ending (Wüster, 1931, p. 296). The direct consequence is that all Esperanto content lexical items are complex (Schubert, 1989, p. 259).^{297,298} A lexical item as short as “domo” (house), for instance, is complex: it is made up of the content morpheme “dom” (a root meaning “house”) and the function morpheme “o,” indicating a noun. Lexical items composed of more than two morphemes are not uncommon in Esperanto (see frequency distributions in Bick, 2016, pp. 1076–1077).

Using Schubert’s classification, for the purposes of this investigation, I propose distinguishing between two types of Esperanto morphemes:

- Content morphemes, for example, morphemes that express concrete meanings: roots, prefixes, prefixoids, suffixes, and suffixoids²⁹⁹
- Function morphemes, for example, morphemes that play grammatical roles without carrying concrete meanings: endings and declension morphemes³⁰⁰

In Esperanto, not only roots but also any content morphemes may carry opinions from speakers, and through mechanisms of word formation, one content morpheme may result in various lexical items belonging to distinct categories of parts of speech. The table below (Table 19) provides characteristic examples of this phenomenon for two types of content morphemes: a root (*taŭg*) and a suffix (*ind*).

Wortbildungsmorpheme) and grammatical inflectional morphemes (German: *grammatische Flexionsmorpheme*).

- 297** A limited set of function morphemes are, on their part, unbound (*la, unu, du tri*, etc.).
- 298** Readers particularly interested in the details of Esperanto grammar can refer e.g. to (D. Blanke, 1981; Schubert, 1989, 2015).
- 299** Affixoids are distinguished from affixes by Schubert because they behave differently in word formation processes (Schubert, 1989, p. 260–261, 2015, p. 2218)
- 300** In my view, there are a few borderline cases: for instance in “O-finaĵo” (noun ending), the first ‘O’ could perhaps be considered to be a content morpheme, although it is usually a function morpheme. At the very least, such an occurrence seems to stand in direct contradiction with some existing classifications such as that of Brosch (2009), who in my comprehension mentions that the endings *-o, -a, -e* cannot be used in word formation (German: *nicht wortfähig*).

Resulting part of speech	Examples from the corpus with lexical items constructed from the opinion-bearing root <i>taŭg</i> ³⁰¹	Examples from the corpus with lexical items constructed from the opinion-bearing suffix <i>ind</i>
adjective	<i>Sed eble proporcio estas pli <u>taŭga</u> vorto, kaj mi konsentas kun vi, ke “ĥemia konsisto” temas pri proporcioj inter ĥemielementoj. But maybe proporcio is a more appropriate word, and I agree with you that ĥemia konsisto is about proportions between chemical elements.</i>	<i>“Tagnokto” estas <u>uzinda</u> sinonimo de “diurno”. Tagnokto is a synonym of diurno that is <u>worth using</u>.</i>
noun	<i>Tial mi pensis pri KLANO kaj POPOLO. Sed eble estas pli ĝustaj vortoj. Aŭ eble iu konvinkos min pri la <u>taŭgeco</u> de POPULACIO... This is why I thought about klano and popolo, but maybe there are words that are more correct. Or maybe someone will convince me about the <u>appropriateness</u> of populacio.</i>	<i>La radiko “agr/” ja estas neoficiala, kaj ĝia <u>uzindeco</u> apud la oficialaj “kamp/” kaj “agrikultur/” ŝajnas al mi dubinda. The root agr/ is certainly not official and it looks doubtful to me that it would be <u>worth using</u> in addition to the official kamp/ and agrikultur/.</i>
adverb	<i>En la tavolo literatura, oni libere uzu, ekz-e, la vorton “trista”; sed estas <u>maltaŭge</u> uzi tian vorton en la parolo kaj la skribo ĉiutagaj.</i>	<i>Se estas vortoj inter la mova ĉefverbo kaj la cetera 1-verbo, estas <u>inde</u> uzi “por”.</i>

301 ■ **taŭga** in “**taŭga vorto**” is an adjective constructed from the root *taŭg-* (content morpheme meaning *appropriate*) and the ending *-a* (adjective function morpheme), and the excerpt can be translated as “appropriate word”

■ **taŭgeco** in “**taŭgeco de POPULATION**” is a noun constructed from the root *taŭg-* (content morpheme meaning *appropriate*), the suffix *-ec* (showing a quality or a characteristic) and the ending *-o* (noun function morpheme), and the excerpt can be translated as “appropriateness of [the word] POPULATION”

■ **maltaŭge** in “**estas maltaŭge uzi tian vorton**” is an adverb constructed from the prefix *mal-* (turning an associated or associated morpheme(s) to their opposite), the root *taŭg-* (content morpheme meaning *appropriate*) and the ending *-e* (adverb function morpheme), and the excerpt can be translated as “it is inappropriate to use such a word”

■ **taŭgas** in “**povigi’ ofte taŭgas**” is a conjugated verb constructed from the root *taŭg-* (content morpheme meaning *appropriate*) and the ending *-as* (function morpheme indicating a verb in present tense), and the excerpt can be translated as “[the word] ‘povigi’ is often appropriate”

Resulting part of speech	Examples from the corpus with lexical items constructed from the opinion-bearing root <i>taŭg</i> ³⁰¹	Examples from the corpus with lexical items constructed from the opinion-bearing suffix <i>ind</i>
	<i>In the literary register, you can freely use for instance the word <i>trista</i>, but it's <u>inappropriate</u> to use such a word in all speech and writing.</i>	<i>If there are words between the main movement verb and the target I-verb, it's <u>worth</u> using <i>por</i>.</i>
verb	<i>Des pli, ĉar por la verbo 'empower', 'povigi' ofte <u>taŭgas</u>. Even more so, because for the verb empower, [the verb] <i>povigi</i> is often <u>appropriate</u>.</i>	<i>Tio estas tino, kaj tute <u>uzindas</u> la vorto "tino" por ĝi. This is a tino [a tub], and it's really worth using the word tino for it.</i>

Table 19. Examples of opinion-bearing lexical items for various parts of speech (POS) respectively formed from two opinion-bearing morphemes: the root *taŭg-* and the suffix *-ind*.

Esperanto is very productive in terms of word formation.³⁰² The examples presented above are only but a few possibilities among all of the lexical items found in my corpus based on the root morpheme *taŭg*³⁰³ and the suffix *ind*, and this is true for most opinion-bearing morphemes observed in my corpus.

Because Esperanto morphemes remain unchanged throughout the language, in some cases, it seems more appropriate to choose a morpheme-based approach to the language, as the morpheme is the smallest unit of meaning that is visible in writing. The elements carrying opinions (or any semantic contents) do *not* start at the level of lexical items but rather at the smaller level of content morphemes. Using morphemes instead of lexical items as basic units allows one to reduce the proportion of out-of-vocabulary units.

For simplification purposes, I therefore propose generally assuming that if a lexical item comprises among its morphemes at least one³⁰⁴ content morpheme carrying an opinion, the resulting lexical item is likely to carry an opinion as well.

302 Due to its schematic design, see again Schubert for word formation details (2015).

303 Other possibilities are e.g. *maltaŭga* (adjective), *maltaŭgas* (verb), *maltaŭgeco* (noun), *metaŭga* (adjective), *metaŭgeco* (noun)...

304 A lexical item can obviously contain more than one opinion-bearing content morpheme. This is e.g. the case for the adjective **prefer-ind-a** in a sentence such as: "*Matenruĝo*" estas eble preferinda. ([The word] *Matenruĝo* is maybe preferable.)

This should be true regardless of whether the morpheme in question is the head morpheme, as the following examples illustrate:³⁰⁵

*Nome mi jam renkontis la **fuŝformon** *ĥoruso*!
I've namely encountered the awful form *ĥoruso*!*

*Intertempe vi ĉiuj rajtas jam redakti, aldonante tradukojn, korigante **tajpfuŝojn**, kaj pri ĉio ajn.
In the meantime, you may all start editing, adding translations, correcting typos, and whatever else.*

Evidently, this generalization remains an approximation: the semantics of complex Esperanto lexical items³⁰⁶ is not as straightforward as summing up the meanings of each morpheme (Schubert, 2015, p. 2216), and many morphemes are polysemous. Furthermore, in a limited number of cases, it seems more appropriate to reason at a higher level than the morpheme:

- When a complex lexical item allows for the disambiguation of a morpheme
- When a complex lexical item (or a larger unit) has a noncompositional meaning³⁰⁷

In my corpus, I identified three types of opinion-bearing content morphemes: (a) opinion-bearing roots proper, (b) opinion-bearing affixes, and (c) opinion-bearing roots behaving as affixes.

The number of opinion-bearing roots is open ended and theoretically unlimited. In the present context, it seems important to distinguish between monosemous roots and polysemous roots. Monosemous roots identified as opinion-bearing can be expected to carry opinions regardless of the context. This is, for

305 Here, in the first example *fuŝformo* (erroneous form), the morpheme *fuŝ* is not the head, while in the second example *tajpfuŝojn* (typos) this same morpheme *fuŝ* is the head morpheme of the lexical item. In both cases, the morpheme is opinion-bearing.

306 That is, again, almost all Esperanto lexical items, and all content Esperanto lexical items.

307 Some Esperanto complex lexical items are not transparent, but they are rather the exception. Complex lexical items show a tendency towards idiomatization, but generally complex lexical items in Esperanto are compositional to a very high degree (Dasgupta, 1993, p. 367).

example, the case for the root *taŭg* introduced earlier, as it is for several other roots observed in my corpus (*ĉagren, elegant, kompetent, ted*, etc.). Polysemous roots, on the other hand, will point to opinion only in a restricted set of cases. “signif” is one of these roots. Especially in the form of an adjective, it means “significant” or “meaningful” and *is* opinion-bearing, but specifically as a verb, it almost always carries the meaning of “to signify,” “to mean,” and “to imply,” in which case it is *not* opinion-bearing.

Opinion-bearing?	Example from my corpus
YES	La diferenco tamen estas sensignifa, kaj mi apenaŭ farus tian gramatikan analizon legante la frazon. <i>However, the difference is <u>insignificant</u>, and I would barely make such a grammatical analysis while reading the sentence.</i>
YES	La fina dokumento estas konsiderata tre signifa dokumento. <i>The final document is considered to be very <u>important</u>.</i>
NO	Ĉu eble ekzistas konfuzo kion signifas ‘propra nomo’? <i>Is there maybe confusion about what “proper noun” <u>means</u>?</i>
NO	Tial necesas distingi, ĉu oni priskribas la signifaron de vorto [...] <i>Therefore, it is important to determine whether we describe all of the <u>meanings</u> of a word.</i>

Table 20. Examples of lexical items (both opinion-bearing and nonopinion-bearing) from the corpus based on the polysemous root “signif.”

The number of opinion-bearing affixes and affixoids is finite. In my corpus, I identified five cases that led to the formation of opinion-bearing lexical items (see Table 36 in the appendices).

Along typical Esperanto affixes, I also identified five “pseudo-affixes,” for example, five roots that tend to behave as affixes in word formation and potentially carry opinions (see Table 37 in the appendices).

7.4.1.2 Larger opinion-bearing expressions

Opinion-bearing lexical items (or in the present case, morphemes) are insufficient for identifying opinionated contexts: “There are restrictions imposed to such methods since multi-word expressions, slang and social attributed connotations not contemplated in the thesaurus or dictionary are not accessible” (Souza, Vieira,

Buseti, Chishman, & Alves, 2011, p. 60). Therefore, I also identified in my corpus larger expressions: (a) expressions indicating opinions (see Table 38 in the appendices), (b) expressions indicating attitude³⁰⁸ (see Table 39 in the appendices) and (c) expressions indicating consent (see Table 40 in the appendices).

7.4.1.3 *Conclusions for an opinion lexicon in Esperanto*

From the observations made in my corpus and presented earlier, an opinion lexicon for Esperanto should include at least the following types of lexical elements:

- Monosemous content morphemes—both roots and affixes (*aĉ, fuŝ, mis, taŭg*)
- Polysemous content morphemes (*signif*)
- Larger expressions (*laŭ mia opinio*)

These categories should be distinguished for facilitating natural language processing tasks.

7.4.2 *Autonymy indicators in written language*

In her doctoral dissertation concerning the French language, Delavigne (2001) concluded that a “grammar” of autonymy indicators remains to be established (p. 498). This applies to Esperanto as well and is even truer in the present context given the large diversity of writing styles and thus the indicators that speakers use in electronic networks.

For clarification purposes, I’ll start by introducing the concept of “**autonym**.” From a linguistic viewpoint, it is necessary to distinguish between a **sign in use** or an **ordinary sign** on the one hand, and an autonym or a **metalinguistic sign** on the other.³⁰⁹ The sign in use points to a referent that is a real-world concept (an

308 Statements in which speakers express their opinion in the form of a conative attitude statement regarding language, i.e. they state their readiness for example for using a specific lexical item. The conative component of attitude as defined by Baker (1992, p. 13), that is “a behavioural intention or plan of action under defined contexts and circumstances.”

309 Rey-Debove in French uses respectively “*signe ordinaire*” and “*signe métalinguistique*”.

object, a thought, etc.), whereas the sign being mentioned refers to another sign as illustrated in the table below.

	Ordinary sign	Metalinguistic sign (autonym)
Linguistic description (Authier-Revuz, 2003, p. 71–72)	$\text{sign} = \frac{\text{signified}}{\text{signifier}}$	$\text{sign} = \frac{\frac{\text{signified}}{\text{signifier}}}{\text{signifier}}$
Notes	Here the sign is divided into two components: the signified (concept) and the signifier (sound-image).	Here, the signified of the sign is itself a sign, composed of a signified and a signifier.
Examples	<i>Kaj la virino venis antaŭ la mateniĝo, kaj falis antaŭ la pordo de la domo de tiu homo, kie estis ŝia sinjoro; kaj ŝi kuŝis, ĝis fariĝis lume.</i> <i>Then came the woman in the dawning of the day, and fell down at the door of the man's house where her lord was, till it was light.</i> [King James Bible, 1987]	<i>Mi tamen havis la ideon, ke ĝuste en Esperanto oni nuntempe uzas la vorton "aŭroro" kiel sinonimon por "mateniĝo".</i> <i>But I had this idea that precisely in Esperanto nowadays the word "aŭroro" is used as a synonym of "mateniĝo".</i> [my translation]

Table 21. Ordinary signs and metalinguistic signs.

Any linguistic sign, and even units smaller than phonemes, graphemes, syllables, etc., or bigger than linguistic signs (e.g., a long quote) can stand in the autonymical condition (Authier-Revuz, 2003, p. 76; see also Rey-Debove, 1971, p. 48), as the following example from my corpus illustrates with the single letter “x” as an autonym:

Litero “x” ekzistas en pola alfabeto, sed nuntempe ne estas uzata en vortoj.
[my translation]^{cx1}

Language uses several operations to mark a distinction between ordinary signs and autonyms in written speech. However, no convention exists for marking autonymy in speech (Tamba, 2003, p. 64). In the example above, for instance, the autonym is clearly flagged with quotation marks, but in other cases, it may appear in the same way that an ordinary sign would:

mi preferas vertikalon, ĉar bone kongruas kun horizontalo^{cxii}

There is a series of indicators that point towards the presence of autonoms in speech, which have already been discussed in the scholarly literature:³¹⁰

- **Syntactic features** (e.g. Authier-Revuz, 2003; Bosredon & Tamba, 1998; Rodríguez Penagos, 2004b, p. II–71; Sitri, 2003, p. 212)
- **Lexical features** (e.g. Delavigne, 2001, pp. 491–498; Rodríguez Penagos, 2004, pp. II–71; Sitri, 2003, p. 212)
- **Paralinguistic features** (e.g. Rodríguez Penagos, 2004, pp. II–71; see also Delavigne, 2001, p. 520)

I introduce these three types of indicators in the following sections, using examples from my corpus.

7.4.2.1 *Autonymical speech: Breaking the rules of a regular language*

Autonyms are signs being mentioned. When inserted into a sentence, they do not behave as they would if they were ordinary signs, from a grammatical viewpoint. When used as such, an autonym tends to lose the grammatical properties it has as an ordinary sign and takes on the grammatical role of a noun in a sentence (Authier-Revuz, 2003, p. 76; Bosredon & Tamba, 1998, p. 180), even if it belongs to another grammatical category. In French, for instance, an autonym that would be a feminine noun when used as an ordinary sign could behave as a masculine noun in the autonymical condition (see Tamba, 2003, p. 64), or an autonym that would be plural as an ordinary sign could behave as a singular noun in a sentence (Bosredon & Tamba, 1998, p. 180):³¹¹

ce “journaux” est bien écrit

A sentence containing an autonym may thus appear to be “agrammatical” (Rey-Debove, 1971, p. 46), if analyzed in terms of ordinary signs.

The grammatical features flagging autonymy are evidently language dependent. To the best of my knowledge, the existing scientific literature has not dealt

.....
310 Rodríguez Penagos also mentions a pragmatic dimension (2004b, p. II–41).

311 The verb in this example sentence is in the singular although the noun “journaux” is in the plural.

with autonymy in Esperanto. Thus, I provide an overview here of the phenomena I observed in my corpus.

At the sentence level, it becomes clear that lexical items may shift their parts of speech when inserted as autonyms into a sentence, as the following example illustrates. “egala” and “ekvivalenta” are both adjectives when used as ordinary signs, but here, they shift their parts of speech and occupy a noun function, serving as the object of the transitive verb “uzi” (“uzus” in the conditional):

Eble mi uzus “egala” anstataŭ “ekvivalenta”

[adverb]+[pronoun]+[transitive verb]+[noun]+[adverb]+[noun]

If analyzed in terms of ordinary signs, the sentence would appear to be grammatically incorrect, with a transitive verb with no noun-object:

[adverb]+[pronoun]+[transitive verb]+[*adjective]+[adverb]+[*adjective]

In my Esperanto corpus, I observed four types of “broken” rules: (a) the absence or separation of the accusative, (b) the absence of noun-adjective agreement (singular-plural), (c) the absence of agreement with determiners, and (d) letters losing their case. Descriptions and examples of these cases are provided in the appendices.

7.3.2.2 *A limited set of metalinguistic indicators*

Lexical items may stand alone in the autonymical condition within a sentence, but they may also be accompanied by an anaphoric determiner or a metalinguistic lexical item serving as a metalinguistic operator (Tamba, 2003, p. 64). A metalinguistic lexical item is, by definition, a lexical item that encloses the concept of metalanguage without being an autonym (Rey-Debove, 1978, p. 32). Following Rey-Debove (1978, p. 27), a language’s lexicon contains lexical items of the object-language, neutral lexical items (high-frequency nonthematical lexical items, grammatical lexical items, etc.), and metalinguistic lexical items. Any lexical items (object-language, neutral, or metalinguistic) can stand in the autonymical condition.

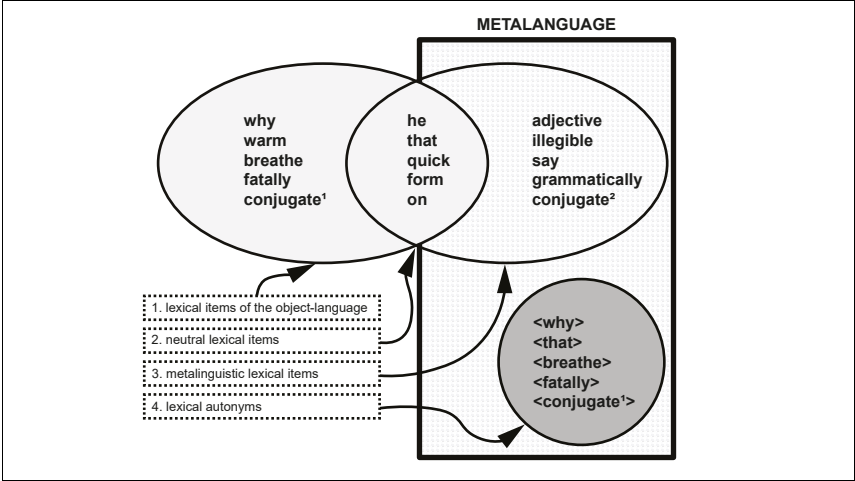


Figure 15. Representation of a language’s lexicon, containing three sets of lexical items: lexical items of the object-language, neutral lexical items, and metalinguistic lexical items. Metalanguage is formed by metalinguistic lexical items and by lexical items from any one of the three lexical sets standing in autonymical condition. Adapted from Rey-Debove (1978).

Thus, from a purely lexical point of view, metalanguage has two components: metalinguistic lexical items and autonyms (Rey-Debove, 1985, p. 27). Here, I am interested only in metalinguistic lexical items flagging autonyms, which I call metalinguistic indicators.

As Delavigne noted for the French language (2000), autonymy is flagged differently among different speakers and can be marked by lexical items of various parts of speech, including verbs (*appeler, signifier, designer*, etc.), connectives (*c’est-à-dire*), nouns (*nom, mot*), adjectives (*fameux, dit*), and adverbs (*plus exactement*). Similar grammatical categories can be found in my Esperanto corpus: verbs (*tradukeblas, rimigas, parafrazeblas*, etc.), connectives (*nome, tiel nomata, t.n.*), nouns (*prepozicio, neologismo, lingvouzo*, etc.), adjectives (*litova, tipografiaj, netransitiva*, etc.), and adverbs (*laŭvorte, anglalingve, plursence*, etc.)

However, as explained for opinion-bearing lexical elements, for Esperanto, it seems reasonable to adopt a morpheme-based approach in most cases and to limit the larger approach to a restricted set of lexical items or larger expressions (e.g., for the connective *tiel nomata* or *t.n.*).

In my corpus, I observed three types of metalinguistic indicators: (a) metalinguistic roots, (b) proper nouns pointing to metalinguistic contents, and (c) metalinguistic lexical items. Examples are provided in the appendices.

7.3.2.3 (Dis)ambiguation? A creative set of paralinguistic indicators

In the existing literature concerning autonymy, quotes are often cited as paralinguistic disambiguators of autonoms (e.g., Delavigne, 2001, p. 491; Rey-Debove, 1985, p. 28). However, what is striking from our corpus is the considerable diversity of autonym indicators to which speakers resort, with quotes being the only one in many indicators used.

In my corpus, I identified eight paralinguistic indicators flagging autonymy: (a) quotes, (b) underscores, (c) hyphens, (d) asterisks, (e) colons, (f) capitalization, (g) parentheses, and (h) line breaks. In addition, these indicators may be combined. Examples for each category are presented in the appendices.

7.3.2.4 Degree of explicitness of autonoms

Indicators of autonymy are largely ambiguous in isolation (e.g., Rodríguez Penagos, 2004, pp. II–72). For instance, capitalization may be used to mark emphasis rather than autonymy in a sentence:

La dua senco de “sankcio” estas tute MALA kaj KONTRAŬA al la unua senco, kio ja estas sennecese konfuza
The second meaning of “sankcio” is completely the OPPOSITE and the CONTRARY to the first meaning, which is unnecessarily confusing

Therefore, it is the combination of indicators that points toward the presence of an autonym. Counting the number of contexts that respectively present at least one metalinguistic and one paralinguistic indicator in the metalinguistic statement with autonym contexts of the test set (see Section 6.3.2), it seems that most of these contexts contain *both* types of indicators (325 out of 425 or about 76%).

Degree of explicitness of metalinguistic statements with an opinionated autonym	Opinion operator?	Metalang. indicator?	Paraling. indicator?	Astronomia Terminaro	Esperanto-tradukistoj	Lingva konsultejo	Retpoŝta rondo...	Vivo-vikio	TOTAL
Rather explicit	yes	yes	yes	38	103	44	102	38	325
	yes	no	yes	3	18	26	10	2	59
	yes	yes	no	5	2	9	9	3	28
Rather implicit	yes	no	no	1	5	4	1	2	13
TOTAL				47	128	83	122	45	425

Table 22. Degrees of explicitness of a metalinguistic statement with an opinionated autonym according to the presence or absence indicators (opinion, metalanguage, and autonymy) in a metalinguistic statement with an opinionated autonym. The figures result from a manual classification of the training set.

However, what also becomes apparent in the above table is that the proportion *might* be dependent on the type of medium. For instance, for the Facebook group *Lingva Konsultejo*, this proportion falls to about 53% of metalinguistic statements with an opinionated autonym contributions presenting both a metalinguistic and a paralinguistic indicator. My intuition suggests that the number of indicators decrease depending on the length of contributions (Facebook contributions being on average shorter than those based on the medium e-mail), but a quantitative study would be needed to confirm this impression.³¹²

Be that as it may, what is certain is that the explicitness of autonyms varies in corpora. Depending on the number of visible indicators in the corpus, I propose to speak of degrees of explicitness for autonyms from relatively explicit metalinguistic statements with an opinionated autonym contexts that contain the three types of indicators (metalanguage and autonymy indicators as well as opinion indicators) ...

Min iom nervozigas la vorto “dubiĝo”.
The word “dubiĝo” makes me a bit nervous.

.....
312 There could also be other reasons: for instance, the profile of participants.

... to contexts containing none of these indicators, but that are still metalinguistic statements with an opinionated autonym contexts:

*mi preferas **vertikalon**, ĉar bone kongruas kun **horizontalo***
I prefer vertikalo[accusative], because it fits well with horizontalo

7.3.2.5 Conclusions regarding autonymy indicators in Esperanto

From the observations made in my corpus and presented above, autonym indicators for Esperanto should include, at least, the following types of elements:

- A set of syntactic norm deviation phenomena; that is, syntactic phenomena that would not be expected in nonmetalinguistic sentences (e.g., absence of the accusative, absence of noun-adjective agreement)
- A set of metalinguistic lexical elements
 - monosemous roots with a metalinguistic meaning (e.g., *transitiv*)
 - polysemous roots with at least one metalinguistic meaning (e.g., *pasiv*)
 - proper nouns that can have a metalinguistic function (e.g., *PIV*)
 - metalinguistic lexical items
 - metalinguistic lexical items disambiguating polysemous roots (e.g., *supersigno*)
 - abbreviations (e.g., *t.n.* for *tiel nomata*)
- Information on paralinguistic indicators
 - A set of indicators (e.g., quotes, asterisks)
 - Information about their insertion into sentences (e.g., before, after, within the autonym, combinations)

These pieces of information and this categorical division should aid natural language processing tasks, which I will discuss in the next two sections.

7.5 Building and evaluating indicators for opinionated autonyms in Esperanto

7.5.1 Developing sentiment and metalinguistic lexicons for Esperanto

In the previous sections, I presented grammatical features of metalinguistic statements with an opinionated autonym. I now propose to operationalize these categories for detecting opinionated autonym candidates in Esperanto by means of natural language processing methods. Lexical elements (opinion-bearing morphemes, metalinguistic indicators, etc.) are of paramount importance³¹³ for detecting the presence of opinionated autonyms in speech. To my knowledge, extensive lexicons exist in Esperanto neither for opinion nor for metalanguage. Therefore, in the following sections, I explain how I constructed such lexicons, starting from seeds and then specializing the vocabularies according to the needs of my research.

7.5.1.1 *Seeds for a sentiment lexicon*

Lexical elements are one of the indicators of sentiment in Esperanto speech (see Privat, 2001 [1930]). The elaboration of my sentiment lexicon for Esperanto was largely influenced by techniques of opinion mining (on opinion mining, see B. Liu & Zhang, 2012). The main tasks of opinion mining or sentiment analysis are “(1) to find product features that have been commented upon by reviewers and (2) to decide whether the comments are positive or negative” (Ding et al., 2008, p. 231). In the present chapter, I was only interested in the first task (i.e., detecting in which contexts speakers express an opinion). It is evident that this approach does not mean that I am not interested in the opinion contents; rather, it suggests that these contents will be the object of a nonautomated qualitative content analysis in the next chapter. Automatically detecting the orientation of an opinion in Esperanto would be a challenge in itself³¹⁴ and falls outside of the scope of the present investigation.

313 According to Liu (2012, p. 12), the most important indicators of opinions are opinion-bearing lexical items (“sentiment words” in Liu’s terms).

314 For instance because opinion-bearing words may be context-dependent, e.g. “long” in “The battery of this camera lasts very long” is positive while it is negative in “This program takes a long time to run” (examples taken from Ding, Liu, & Yu, 2008, p. 234)

Dictionaries for lexicon-based opinion mining can be created manually or automatically (Taboada, Brooke, Tailskid, Voll, & Stede, 2011, p. 268). There already exists a large range of freely available lexical resources,³¹⁵ but most of them are in English. Some authors developed methods for obtaining lexicons in other languages from English lexicons, such as for machine translation, cross-language projections or bootstrapping (details see Balahur Dobrescu, 2011, p. 26). In the present investigation, I opted for such an approach based on an existing English lexicon: I used Liu's lexicon entries (see Hu & Liu, 2004) as seeds (6789 seeds), semiautomatically searching for equivalents in English-Esperanto dictionaries. I obtained a "seed list" of opinion-bearing lexical items.

7.5.1.2 *Seeds for a metalinguistic lexicon*

A **metalinguistic lexical item** is a word that, to a certain degree, comprises the concept of language without being an autonym (Rey-Debove, 1978, p. 32). The metalinguistic lexicon is a restricted, small, and relatively closed set of units (Authier-Revuz, 2003, p. 75; Delavigne, 2001, p. 489; Harris, 1991, p. 275). Metalinguistic lexical items are mostly grammatical indicators (e.g., "adjective" and "declension" for talking about the language, as well as other words referring to linguistic aspects of a language, such as "illegible"). Unlike for sentiments, there exist several metalinguistic lexicons in Esperanto in the form of glossaries in grammar or language teaching materials.

The seeds for my metalinguistic lexicon were obtained from two types of sources: (a) reference works and (b) corpora. The reference works were:

A1. The Esperanto grammar PMEG (Wennergren, 2005), which contains a "gramatika vortareto" (small grammar dictionary). This work was chosen because it is often considered to be the current reference grammar for Esperanto.

A2. Brosch's (2008) interlinguistics diploma thesis, which contains a "malgranda terminaro vortfarada" (short-term list of word formation in German-Esperanto). This work was chosen because it tries to reconcile the terminologies of general linguistics with that of interlinguistics.

.....
or because opinion shifters (negation, sarcasm, etc.) can change the valence of an opinion (B. Liu & Zhang, 2012, p. 433).

315 WordNet Affect, SentiWordNet, Emotion triggers, etc., see e.g. the list in Balahur Dobrescu (2011).

The corpora used were:

B1. Zamenhof's *Lingvaj Respondoj*. Keywords were generated³¹⁶ using other Zamenhof works as a reference corpus (texts available on tekstaro.com, see Esperanto Studies Foundation, n. d.). These texts were chosen because they were written by the initiator of the language, often considered as the ultimate language reference by Esperanto speakers.

B2. Speakers' comments on the language blog *Lingva Kritiko* (Wennergren, n. d.). A word list was generated,³¹⁷ from which metalinguistic lexical items were manually extracted. This blog was chosen because the style and the format are less formal than standard grammar works or dictionaries and therefore can be expected to come closer to the way folk speakers express their views in electronic networks.

7.5.1.3 *Extending and specializing the lexicons*

The two seed lexicons constitute a good starting point, but are not sufficient. In sentiment analysis, "it has been argued that the usage of domain-independent lexicons is unsatisfactory and domain-specific lexicons should be constructed" (Souza et al., 2011, p. 61). Furthermore, as far as my metalinguistic lexicon is concerned, the four sources used for establishing a list of metalinguistic seeds predominantly originated from reference authors. However, given the context of my investigation, this lexicon must also include a larger spectrum (i.e., elements and synonyms folk speakers would use in informal communication, such as in networks where anyone can participate).

As the overall aim of my approach is to continuously "monitor ordinary speakers' subjective opinions about lexical items" (see Section 7.1), I wanted to opt for a technique that allows for the quick addition of new entries to the lexicons when new data are added to the monitor corpora. Thus, I chose to rely on a technique that allows for learning from unstructured text data (i.e., an unsupervised learning technique). As Firth (1957) once stated, one can "know a word by the company it keeps" (p. 11). Based on this latter idea, a statistical model of language can indicate the conditional probability of a word, given all the previous ones (Bengio, Ducharme, Vincent, & Jauvin, 2003, p. 1138). In the present investigation, I used Google's Word2Vec model, a skip gram neural network

.....
316 For this task, I used AntConc, a freeware concordancer software program developed by Laurence Anthony.

317 I also used AntConc for this task.

architecture (see Mikolov, Chen, Corrado, & Dean, 2013; Mikolov, Sutskever, Chen, Corrado, & Dean, 2013).³¹⁸ Representing lexical items as vectors³¹⁹ allows for the identification of similar lexical items within a corpus. For instance, from the vectors for “king,” “man,” and “woman,” it becomes possible to automatically find the close vector for “queen” (Mikolov, Chen, et al., 2013, p. 2).

Used on my Esperanto corpus, Word2Vec produced results that could be used for lexicon extension.³²⁰ From a seed such as “absurda” (absurd), for example, the algorithm generated the output “ampleksa” (extensive), “amuza” (entertaining), “belega” (magnificent), “eleganta” (elegant), and so on. Additional examples are provided in the appendices.

For extending and specializing the opinion lexicon, Word2Vec was used exclusively on my corpus because, to my knowledge, there is no corpus or readily available sources of texts specifically containing opinion-related contents in Esperanto (such as a collection of customer reviews). In fact, most Esperanto corpora are limited in size and representativeness (Normantas, 2013).

For extending the metalinguistic lexicon, I used the Esperanto version of Wikipedia, a large body of texts that includes articles on a wide range of topics (e.g., linguistics). Using the Word2Vec algorithm on the Esperanto Wikipedia produced relevant results. For instance, from the seed “mallongigo” (abbreviation), the algorithm retrieved “akronimo” (acronym), “alemana” (alemannic), “araba” (arabic), “bretona” (breton), “cirila” (cyrillic), “dialekto” (dialect), “esperantlingva” (in Esperanto language), “fonetika” (phonetic), “hebrea” (Hebrew), and so on. To specialize the metalinguistic lexicon, I used the Word2Vec algorithm on my own corpus. A few examples of results are provided in the table below. Examples for extension and specialization are provided in the appendices.

Equipped with the grammatical observations from Sections 7.4.1 and 7.4.2, as well as the two extended Esperanto lexicons for opinion and metalanguage here presented, I built indicators in the form of regular expressions that I tested on the test set. The results are provided in the two following sections.

318 I used the Gensim implementation of Word2Vec in Python.

319 Each lexical item is represented as a vector of floating point numbers, and semantically similar lexical items find themselves close to each other in space.

320 The list of results also comprised irrelevant elements that had to be filtered out manually (where “irrelevant” relies on my subjective evaluation of whether or not the element is likely to point towards an opinion). Also, the results are neither quantified nor compared with other unsupervised learning techniques. This would fall outside of the scope of the present thesis. Further investigations could, however, serve to identify the most efficient technique to extend the lexicons and thus optimize the process.

7.5.2 Indicators for detecting opinions in Esperanto corpora

In the previous sections, I concluded that an opinion lexicon for Esperanto should contain at least monosemous content morphemes (both root and affixes), polysemous content morphemes, and larger expressions. Accordingly, the indicators are divided into these three categories and monosemous and polysemous content morphemes are evaluated separately.

I divided the opinion lexicon obtained in the previous section into monosemous and polysemous opinion-bearing roots.³²¹ The list of monosemous roots is comprised of 345 items and that of polysemous roots includes 305 items (see under Opinion indicators in the appendices, starting on p. 433). During the test, the indicators allowed for the roots to be used anywhere within a lexical item (i.e., also in combination with other content morphemes). The results and examples are provided in the appendices.

For assessing indicators, I proceeded as follows: On the test set (see Section 7.3.3), I compared, for each context, the results obtained by the indicators (automated classification) with the results obtained through a manual classification. If a context was classified as containing an opinion (at least one), then the test was considered successful. This evaluation is a rough approximation, as it does not consider whether the opinion identified by an indicator is identical to the one identified through manual classification.

The limits—especially of polysemous roots as indicators—were clearly visible in the results: A root such as “klar,” for instance, allowed for the retrieval of relevant cases in which “klara” meant “clear” in the sense of easily intelligible. Such roots also created noticeable noise, especially in the astronomy group, in which one of the members liked to end their messages with the sentence “Klaran ĉielon!” to wish a clear sky for observation to fellow group members. Another example is the root “pez” (heavy), which was sometimes relevant when used to mean “burdensome” or “cumbersome,” but also at times irrelevant (e.g., when used in relation to weight). Further methods would thus be required to disambiguate polysemy.

The five opinion-bearing affixes observed in the previous section were tested on the test set (position of the affix anywhere within a lexical item). Detailed results are provided in the appendices (p. 437).

321 I automated this task using Python to compare the list of items against the number of meanings registered in the ReVo dictionary.

The results for *ind* and *end* were generally satisfactory. *Aĉ* was absent of the test set and could therefore not be tested. *Mis* and especially *ebli* obtained relatively unsatisfactory results:

- *Mis* resulted in 10 false positives. Some verbs and nouns constructed with this affix (e.g., “mistraktado de bestoj”, “miskompreni”) had not been manually classified as opinion-bearing.
- *Ebl* generated 155 false positives. The adverb “eble” and the verb “ebli” (eblas, eblus...) were the most problematic.

For *mis*, as the figures are relatively small, no action was taken. For *ebli*, however, I adapted the indicator to exclusively include adjectives formed with the root *ebli*, which improved the results (less items identified but increased precision).

The five opinion-bearing roots behaving as identified affixes were also evaluated on the test set. The results for *fuŝ* were relatively satisfactory. *Pseŭdo* could not be tested as there were no occurrences in the test set. The three remaining roots, *ŝajn*, *pov*, and *kapabl*, obtained relatively poor results:

- *Ŝajn* yielded 37 false positives. The adverbs “ŝajne” and “verŝajne,” the verb “ŝajni” (*ŝajnus*, *ŝajnas*...), and the adjective “ŝajna” were problematic.
- *Pov* resulted in 106 false positives. The verb “povi” (*povus*, *povas*...) was primarily problematic. Some nonopinion-bearing nouns were also observed (*disigpovo*, *distingpovo*).
- *Kapabl* generated 13 false positives. The verb “kapabli” (*kapablas*, *kapablos*...) and a few nouns (*tenkapablo*, *mensa kapablo*...) were problematic.

In my observations, these three roots had been typically noted in combination with a metalinguistic root within a lexical item (e.g., *ŝajn-anglismo*, *esprimpova*, *rimkapabla*), but such combinations were not present in the test set. I subsequently tested these three indicators with adjective endings only, but these indicators yielded a very small recall for a relatively moderate precision.

Larger opinion-bearing expressions that were observed were tested on the test set and obtained better precision than most other opinion indicators.

Each indicator in isolation can only classify a limited amount of contexts into the opinion category (limited recall). A solution to enhance recall is to add up several indicators.³²² However, as each indicator tended to produce a different

.....
322 In mathematical terms, this is a union of sets here (and not an intersection).

type of noise, precision may be negatively affected by a union. Table 23 (p. 246) illustrates this issue with the indicators that fared best on the test set. For example, the combined precision of monosemous and polysemous opinion-bearing roots appeared to be smaller than that of each indicator when used separately.

Unsurprisingly, recall seemed to be generally inversely proportional to precision. For instance, in isolation the affix “end” reaches a relatively high precision (88% of contexts classified as belonging into the opinion category have been correctly classified as such), but a minor recall of only 3% (i.e., 97% of contexts that contain an opinion have not been classified into the opinion category by this indicator). In contrast, the opinion-bearing affix “mis” has a much higher recall of about 40%, but about a fourth (26%) of the contexts classified into the opinion category were not manually classified as containing an opinion.

Indicator	Precision	Recall
Monosemous opinion-bearing roots	0.76	0.29
Polysemous opinion-bearing roots (without <i>long</i> and <i>grand</i>)	0.73	0.53
SUBTOTAL Opinion-bearing roots	0.71	0.55
Opinion-bearing affix <i>mis</i>	0.73	0.4
Opinion-bearing affix <i>aç</i>	–	–
Opinion-bearing affix <i>end</i>	0.88	0.03
Opinion-bearing affix <i>ind</i>	0.8	0.13
Opinion-bearing pseudo affix <i>fuş</i>	0.81	0.20
Opinion-bearing combination <i>abl-aj?n?\b</i>	0.75	0.16
SUBTOTAL Opinion-bearing (pseudo) affixes	0.74	0.33
Larger opinion-bearing expressions	0.83	0.29
SUBTOTAL Opinion-bearing (pseudo) affixes and larger opinion-bearing expressions	0.77	0.43
TOTAL All opinion indicators	0.69	0.76

Table 23. Results (precision and recall) for opinion indicators on the test set, in isolation (by type of indicator) and in combination (union of indicators).

7.5.3 Indicators for detecting autononyms in Esperanto corpora

Based on my observations, I constructed nine indicators in the form of regular expressions for detecting syntactic norm deviation phenomena. Details and results are provided in the appendices, but I give an example again here to explain the principle:

Mi ankaŭ ne ŝatas viskeco-n. Tial ĝi estas variaĵo.
I don't like viskeco[accusative] either. This is why it's a variant.

If the above sentence was nonmetalinguistic, one would expect the accusative case to be marked with the -n ending:

Mi ankaŭ ne ŝatas viskecon.

However, here the speaker used a hyphen to clearly separate the accusative case from the noun:

Mi ankaŭ ne ŝatas viskeco-n.

This indicates that the noun “viskeco” stands in autonymical condition (see Section 7.4.2.1, for examples of syntactic patterns that would not be expected in non-metalinguistic language). Rodríguez Penagos (2004) also used syntactic markers (he mentions, for example, apposition and copulative clauses (p. II–71): Here my contribution is to work with syntactic patterns that are specific to Esperanto (needless to say, there is no accusative in English, for instance).

For assessing indicators, I proceeded as I did with opinion indicators: On the test set (see Section 7.3.3), I compared, for each context, the results obtained by the indicators (automated classification) with the results obtained through a manual classification. If a context was classified as containing an autonym (at least one), the test was considered successful. This evaluation is a rough approximation, as it does not consider whether the autonym identified by an indicator is identical to the one identified through manual classification.

The majority of indicators (SA1, SA3, SA4, SA6, SA7, SA9)³²³ obtained relatively good results.³²⁴ Two indicators (SA5, SA8) did not yield any results on the

323 Here, I am using personal codes to refer to groups of indicators (see e.g. Table 24).

324 No definitive conclusions should be drawn, however, as the number of detected cases was extremely limited.

test set. The indicator SA2 was problematic because of the noise it generated: It not only detected autonyms, but also proper noun (see example in the appendices).

In addition to the norm deviation phenomena, I used a metalinguistic lexicon that included the following elements (see lists, examples and results in the appendices):

- 105 proper metalinguistic roots and 281 root combinations
- Two sets of metalinguistic lexical items formed on the basis of 373 language roots

The precision of proper metalinguistic roots was relatively poor. Apart from false positives, this can be explained by the fact that metalanguage does not equate to autonymy. For instance, in the test set of 986 metalinguistic statements, 732 (74%) are metalinguistic statements with autonyms and 254 (26%) are metalinguistic statements without autonyms. Metalinguistic lexical elements alone may point toward autonymy, but are not sufficient to detect contexts with autonyms. This finding was highlighted by Rodríguez Penagos (2004b), who suggested that “in general the autonimical nature of terms [lexical items] is done redundantly by two or more markers/operators” (p. II–70). An effective way to circumvent the issue is thus to combine metalinguistic roots with another type of indicator. On the test set, this strategy that combines metalinguistic roots with a paralinguistic indicator (here quotes) proved to be a success, increasing precision by 26%.

Apart from metalinguistic roots, I also tested two sets of metalinguistic lexical items formed on the basis of language roots. The first set includes two types of lexical items: substantive items formed on the basis of a language root and the root “ism”—francismoj” Gallicisms, “anglismoj” Anglicisms, “germanismoj” Germanisms, and so on—and verbal items formed on a language root and the suffix *ig* (causing something to be)—“francigi” make French (translate into French), “angligi,” “germanigi,” and so on. This first set reached a fairly high precision for classifying autonymy. The second set represents language adjectives—“angla lingva” English, “francilingva,” “germanlingva,” and so on—and language speakers—“anglaparolanto” speaker of English, “francilingvano,” “germanparolanto,” and so on. It rather poorly classified contexts with regard to autonymy.

Finally, I tested a set of paralinguistic indicators. The first two indicators (PO1, PO2) were designed to detect foreign language items between quotes or parentheses. The corpus is morphologically tagged and, as was already mentioned, apart from a few exceptions (“ankaŭ,” “mi,” etc.), Esperanto lexical items are all composed of several morphemes. Thus, apart from the few monomorphemic Esperanto lexical

items, a one-word lexical item that has not been tagged as plurimorphemic and is accompanied by a paralinguistic indicator (quotes, parentheses...) can be expected to be a foreign-language lexical item standing in autonymical condition. These two indicators functioned relatively well, but produced limited recall on the test set.

PO3 served to identify cases in which speakers propose several lexical solutions, linking them with the conjunction “*au*” (or). The lexical solutions are marked with paralinguistic clues (here quotes). This indicator reached a high precision on the test set.

PO4 was designed to recognize multiword lexical items standing in autonymical condition that are separated by the underscore character. This indicator proved to be counterproductive because it detected lexical items that are being emphasized, but that are not necessarily autonyms and a plethora of links containing the underscore character.³²⁵

Finally, PO6 was elaborated with the aim of detecting cases where speakers mention a lexical item that is considered to be incorrect (or unattested in natural language) and mark this incorrectness with an asterisk (*), as is common within linguistics. This indicator worked on the test set, but was so infrequent that it would be unwise to draw any conclusion as to its performance.

As was the case for opinion indicators, each indicator in isolation can only classify a limited number of contexts into the autonymy category (limited recall), but here again a solution is to add up the results of several indicators (union), as illustrated by Table 24, which presents isolated results and possible unions:

Indicator	Precision	Recall
Syntactic norm deviation phenomena SA1	0.85	0.07
Syntactic norm deviation phenomena SA3	0.82	0.02
Syntactic norm deviation phenomena SA4	1	0.01
Syntactic norm deviation phenomena SA5	–	–
Syntactic norm deviation phenomena SA6	1	0.01
Syntactic norm deviation phenomena SA7	1	0.00
Syntactic norm deviation phenomena SA8	–	–
Syntactic norm deviation phenomena SA9	1	0.01
SUBTOTAL Norm deviation phenomena	0.71	0.10

.....
325 For this latter issue, an adapted regular expression or a subsequent filter could be developed.

Indicator	Precision	Recall
Metalinguistic roots MLI-LR1 combined with a paralinguistic indicator	0.84	0.20
Metalinguistic roots MLI-LANG1	0.83	0.03
SUBTOTAL Metalinguistic indicators	0.83	0.22
Paralinguistic indicator PO1	0.83	0.01
Paralinguistic indicator PO3	0.91	0.06
Paralinguistic indicator PO6	1	0.00
SUBTOTAL Paralinguistic indicators	0.90	0.07
SUBTOTAL Metalinguistic indicators and Paralinguistic indicators	0.82	0.24
TOTAL All autonomy indicators	0.77	0.36

Table 24. Results (precision and recall) for autonomy indicators on the test set, in isolation (by type of indicator) and in combination (union of indicators).

Although these results are promising (e.g., precision of 90% for paralinguistic indicators), they also show the limitation of my small test set: Several indicators may have obtained good results because the number of cases was limited (and not because they were statistically relevant). Much larger corpora would be needed to confirm the potential of these indicators, but this is out of the scope of the present investigation that I undertook alone.

7.6 Extracting opinionated autonomy candidates from the corpus

At the beginning of this chapter (7.1), I proposed to decompose the task of classifying metalinguistic statements with an opinionated autonomy into a two-step binary classification task (i.e., determining (a) whether a given context contains an opinion and (b) whether this same context contains an autonomy). As these two tasks have now been undertaken, the results can be combined. Contexts containing an opinion and an autonomy find themselves at the intersection of the opinion set (Set A) and the autonomy set (Set C); that is, in Set C (see Figure 16). Combining both thus means determining which contexts have been identified as positive by *both* natural language processing classification tasks (intersection).

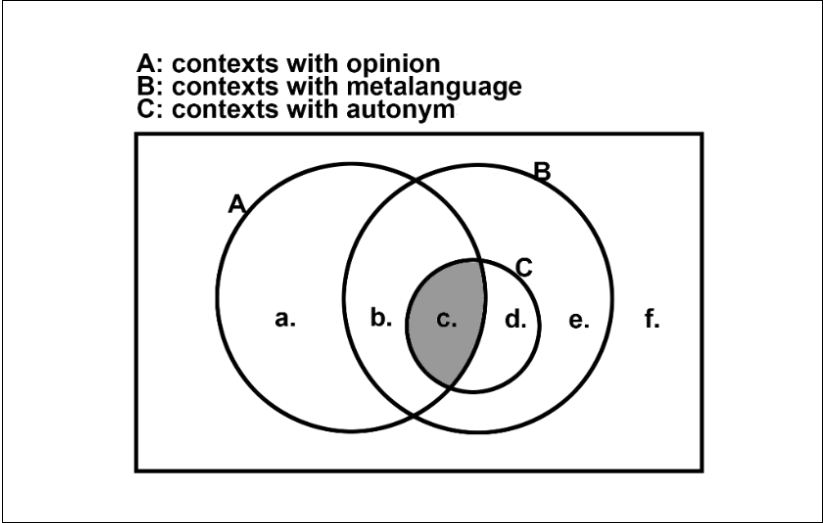


Figure 16. Venn representation of contexts containing both opinion and autonymy (intersection, Set C.).

For evaluation, as for the previous classification tasks, the list of contexts *automatically* classified into Set C can be compared against the list of contexts *manually* classified into Set C.

The classification results obtained in the previous sections have shown that I have obtained a maximal recall of 36% for autonymy and of 76% for opinion. Thus, I can, in the best case scenario, hope to extract at the intersection (Set C) a maximum of 36% of all contexts containing opinions and autonyms (i.e., about 162 out of 450 for the test set). This value is almost reached when combining all autonymy indicators with all opinion indicators (35%), but the resulting precision is relatively low (68%).³²⁶

For applications needing more precision, it seems more reasonable to try to combine autonymy and opinion indicators of a higher precision. Two combinations gave more satisfying results:

326 This precision is in fact even slightly lower than the respective precisions of all opinion indicators (69%) and that of all autonymy indicators (77%). This is because more false positives than true positives may add up at the intersection.

- Combination 2: Intersection between metalinguistic roots MLI-LR1 combined with a paralinguistic indicator on the one hand and larger opinion-bearing expressions on the other (precision 0.82, recall 0.09)
- Combination 3: Intersection between metalinguistic indicators and paralinguistic indicators on the one hand and opinion-bearing pseudo-(affixes) and larger opinion-bearing expressions on the other (precision 0.83, recall 0.17)

At the intersection between opinion and autonymy, one can thus automatically obtain a set of relatively relevant opinionated autonym candidates by means of natural language processing. With a precision above 0.8 (Combinations 2 and 3), this means that more than 8 out of 10 contexts are expected to contain both an opinion and an autonym.

7.7 Units of analysis extracted from the corpus

Now that I have developed a method to extract opinionated autonym candidates (opinionated autonym candidates) from large bodies of text with a precision that I deem to be acceptable ($> 80\%$), I can use it on my entire corpus (69,792 contributions) to find relevant contributions (i.e., contributions containing opinionated autonyms). As the recall is higher for Combination 3, the expected number of opinionated autonym candidates is also higher, as illustrated in Table 25. Assuming that the test set is representative of the entire corpus, using Combination 3 I can expect to extract 3394 opinionated autonym candidates, of which about 2817 should be relevant.

Indicator combination	Expected number of opinionated autonym candidates (OAC) ($OAC = n * \frac{450}{1566} * \text{recall}$)	Expected number of relevant contexts (RC) ($RC = OAC * \text{recall}$)
<i>Combination 1</i>	6987	4751
<i>Combination 2</i>	1797	1473
<i>Combination 3</i>	3394	2817

Table 25. Expected number of opinionated autonym candidates and expected number of relevant contexts for each indicator combinations above assuming the test set is representative of the entire corpus.

After running Combination 3 on the entire corpus, the actual number was higher than the expectation (+21%), yielding 4,090 opinionated autonym candidates.

	<i>autonym yes, opinion no</i>	<i>autonym yes, opinion yes (OAC)</i>	<i>autonym no, opinion yes</i>	<i>autonym no, opinion no</i>	<i>TOTAL</i>
<i>Astronomia Ter- minaro</i>	23	53	134	537	747
<i>Lingva konsultejo</i>	2,452	1,017	4,947	28,603	37,019
<i>Retposhta rondo...</i>	1,409	2,200	5,833	12,668	22,110
<i>Esperanto- tradukistoj</i>	860	779	2,205	5,199	9,043
<i>ViVo-vikio</i>	83	41	151	598	873
<i>TOTAL</i>	4,827	4,090	13,270	47,605	69,792

Table 26. Opinionated autonym candidates (OAC) extracted from the corpus.

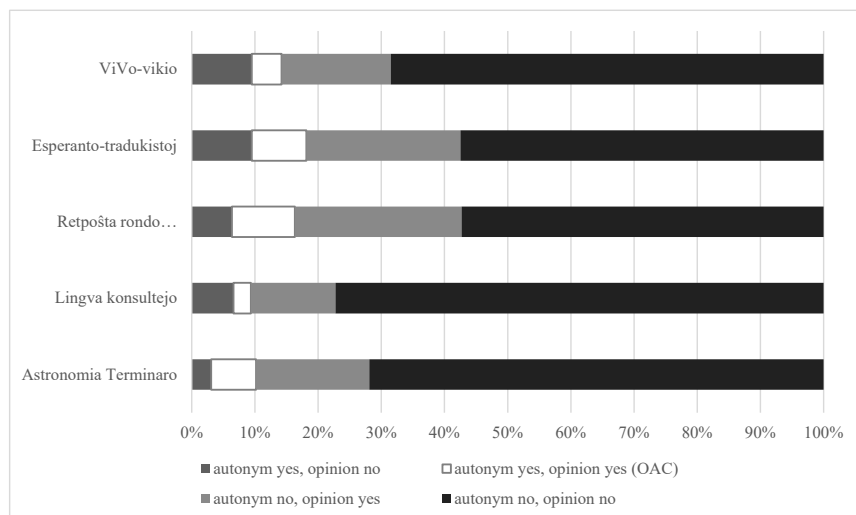


Figure 17. Contexts as classified in absolute numbers (table) and percentages (graph). The contexts selected for the analysis of this chapter are those containing both autonymy and opinion (opinionated autonym candidate contexts or OAC, white on the graph).

The 4,090 contributions served as units of analysis for the next chapter (qualitative analysis of lexical criteria in context). This analysis was also an opportunity to reflect on the indicators' performance. These indicators obtained far less good results on the entire corpus than they did on the test set. From the 4,090 contributions, I manually found that only 1,278 (about 31%) were true metalinguistic statements with an opinionated autonym (i.e., about the same proportion of metalinguistic statements with autonym expected in the corpus). A further set of 1,163 contexts (about 28%) were manually classified as a metalinguistic statement with autonym (thus nonopinionated): In some cases, these were clearly nonopinionated; in others it was hard to determine whether the speaker was expressing an opinion, so I classified them as nonopinionated. The corpus results are mixed and do not corroborate the results obtained on the test set. Further analyses would be needed to determine the underlying reasons.

7.8 Summary

In the present chapter, I offered a working proof of concept: I was able to use the lexical, syntactic, and paralinguistic peculiarities observed in my corpus pointing toward opinion and autonymy for identifying a small portion of contexts containing opinions and autonoms using natural language processing on a test set with relatively good precision. With all its imperfections, this proof of concept confirms my suggestion that it should be possible to systematically monitor speakers' opinions about lexical items using natural language processing. These first results pave the way toward further work on indicators of opinionated autonoms and toward the development of a largely automated application. The main advantage of such a system for language managers would be to have a tool for monitoring opinionated autonoms within short time constraints. However, the ideas put forward in the present chapter suffer several important limitations:

- The precision of indicators is not optimal. This precision could and should be improved, for example, using filtering techniques for eliminating noise as done in Rodríguez Penagos' thesis (2004b). As the type of noise may greatly vary from indicator to indicator, a specific filtering technique should be developed for each indicator, which would require substantial additional work. Also, and perhaps more importantly, whereas the indicators obtained relatively good results on the test set,

the results on the entire corpus are not satisfactory. Further analyses are needed to explain why and to adapt the indicators accordingly.

- Indicators pointing toward opinion in the present investigation are only of a lexical nature. The use of a lexicon is “necessary but not sufficient for sentiment analysis” (B. Liu, 2012, p. 12). The potential of other indicators could be explored as well; for instance, syntactic clues (e.g. sentence arrangement in Esperanto,³²⁷ see Privat, 2001, p. 23 [1930]) could be investigated.
- The subcorpora used in the present investigation are expected to contain opinionated autonyms. Once indicators are improved, they should also be tested on larger text collections that do not directly concern language. For example, one could imagine a web crawler that extracts opinionated autonym candidates for language managers.
- The proposed methodology finds contexts that contain both an opinion and an autonym. This is an approximation as the opinion may concern a target other than the autonym itself. As Liu stated (2012), “an opinion without its target being identified is of limited use” (p. 11). An improvement could be to integrate the work done by opinion mining and sentiment analysis on opinion targets to ensure that opinions and autonyms are related in the opinionated autonym candidates. This outcome would be particularly relevant for long contributions.
- Although the results are promising, most figures are not statistically significant because the number of identified cases is limited. Ideally, each indicator should be tested on larger individual samples with bigger manually annotated contexts.

327 For instance, putting the object at the beginning of a sentence to mark emphasis: “Tiun lampon vi aĉetis” (object-subject-verb, instead of “Vi aĉetis tiun lampon.”, subject-verb-object) may, in my opinion, point towards the expression of an opinion (among other things).

8 Results: Lexical criteria in context

8.1 Introduction

In Chapter 3, I argued for the monitoring of speakers' lexical opinions in natural settings (i.e., when speakers are not being explicitly observed by researchers or language managers. Such monitoring can serve to capture speakers' reaction toward lexical items (generally, or toward the target lexical item proposed by language managers)).

In the present investigation, I proposed to access speakers' lexical opinions in the corpus using natural language processing methods. To this end, in Chapter 7, I laid the foundations for the partly automated detection of what I have called metalinguistic statements with an opinionated autonym with an autonym or, shortly, opinionated autonoms.

The present chapter is the analysis of the detected opinionated autonoms proper. It starts by presenting the framework of analysis: previous related work in Section 8.2 and the analytical framework in Section 8.3. Specifically, the latter section explains why qualitative content analysis was chosen as an analytical method as well as the choice of the units of analysis.

Section 8.4 presents the results obtained from the analysis of the 4,090 opinionated autonym candidates—23 categories of lexical criteria speakers used in evaluating (accepting or rejecting) lexical items—and discusses them in light of existing research. In Chapter 9, I will then explain why these results matter to language managers.

8.2 Analysing opinionated autonoms

The analysis of opinionated autonym candidates extracted in Chapter 7 (see Section 7.7) is performed using a *qualitative content analysis approach*. I chose to use content analysis because it is suited for studying the contents of communication (Lichtman, 2013, p. 259). “In content analysis, researchers examine artifacts of social communication” (Berg, 2011, p. 240). Furthermore, content analysis should be used if the research question is “a descriptive ‘what’ question (such as: What are the interviewees saying here...)” (Schreier, 2012, p. 48) Also, content analysis is said to be well-suited to analyzing metalanguage:

Content-based approaches to discourse analysis have frequently been used to analyse directly-expressed language attitudes as they appear within discourse, often to lend weight to a quantitative analysis (e.g. Dailey-O’Cain 1997; Deminger 2000; Hoare 2000; Lammervo 2005). Like all discourse-analytic approaches to language attitudes, this approach requires a large corpus of data that the researcher must examine for the occurrence of stretches of conversation in which language attitudes are expressed. The researcher then analyses the content of the attitudes in each example, looks for overall patterns that emerge, and sorts these expressions of attitudes into categories according to the arguments he or she wants to make by providing examples from each category in the discussion. (Liebscher & Dailey-O’Cain, 2009, p. 197)

I opted for *qualitative* content analysis, rather than *quantitative* content analysis, because the principles I used for building the corpora are only suited for a qualitative analysis. As McEnery and Wilson (1997) mentioned, it is important to monitor corpora “because they are constantly changing in size and are less rigorously sampled than finite corpora they are not such a reliable source of quantitative (as opposed to qualitative) data about a language” (p. 23). The sampling techniques required by the two approaches differ: In quantitative content analysis, data must be selected using probabilistic approaches to ensure the validity of the statistical inference, whereas in qualitative content analysis texts are selected according to a purpose and a research question (Zhang & Wildemuth, 2009, p. 2).

Qualitative content analysis consists of the systematic analysis of text material. The material is analyzed step by step on the basis of a system of categories (Ramsenthaler, 2013).³²⁸ This system of categories is generally called a **coding scheme** or **coding frame**. It is “a way of structuring ... material, a way of differentiating between different meanings *vis-à-vis* [a] research question. It consists of main categories or dimensions and a number of subcategories for each dimension that specify the meanings in [the] material with respect to these main categories” (Schreier, 2012, p. 61) There exist various approaches to qualitative content

328 Zhang and Wildemuth (2009), for instance, suggest an 8-step process: (1) preparation of the data, (2) definition of the units of analysis, (3) development of categories and a coding scheme, (4) test of the coding scheme on a sample, (5) coding of all the text, (6) assessment of the coding consistency, (7) conclusions from the coded data, (8) report of the methods and findings.

analysis. As several works have already put lexical criteria under the microscope,³²⁹ the content analysis in the present investigation is directed.

Sometimes, existing theory or prior research exists about a phenomenon that is incomplete or would benefit from further description ... The goal of a directed approach to content analysis is to validate or extend conceptually a theoretical framework or theory ... Using existing theory or prior research, researchers begin by identifying key concepts or variables as initial coding categories ... code all highlighted passages using the predetermined codes. Any text that could not be categorized with the initial coding scheme would be given a new code. (Hsieh & Shannon, 2005, p. 1281)

Thus, the development of my coding scheme was both driven by theory (categories from the aforementioned works) and data (new categories from my corpus), as illustrated in Figure 18.

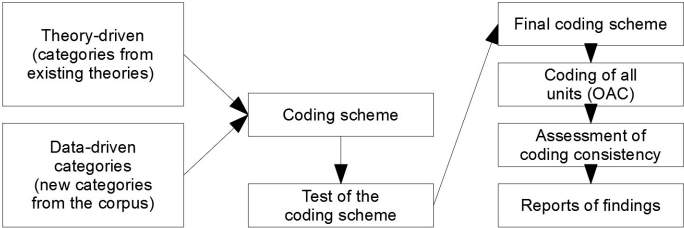


Figure 18. Process for directed content analysis (inspired from Hsieh & Shannon, 2005; Mayring, 2015, pp. 62–63; Zhang & Wildemuth, 2009).

The units were analyzed in terms of lexical criteria categories, but also in terms of polarization (positive or negative evaluation).

Approaching speakers’ subjective statements in the corpora poses a few difficulties because they reveal a complex argumentative structure: Sometimes it is difficult to determine whether a speaker is stating a fact or expressing an opinion. As

329 In her 2016 paper, Saint opted for an open-coding approach stating that to her knowledge “terminological opinions” had not been studied in a corpus like hers. I find it regrettable, because she proposes only six categories although comparable works could, in my opinion, have proved useful in the analysis of her corpus (see previous work above). Lexical criteria have been studied from other authors, and their findings should be integrated in any new study.

several scholars have noted, in “ordinary” speakers’ speech (e.g., descriptive statements), reported speech or exemplifications may also have an argumentative value (Doury, 2001; Rheault, 2004, p. 25; Vincent, 1994).

Therefore, difficulties arise in the interpretation of speakers’ statements about language; for instance, that of implicit subjectivity (see Kerbrat-Orecchioni, 2009, pp. 167–168). In a statement such as “it is not a good word,” a speaker tries to pass the evaluation off as an objective statement although it is utterly subjective.³³⁰ The interpretation is all the more difficult insofar as, in large-scale collaboration systems or in social networks, the length of some contributions may be extremely short.³³¹

In a sense, any lexical item of a language is of a subjective nature because words are nothing but symbols for replacing and interpreting referential objects, may they be real or imaginary (Kerbrat-Orecchioni, 2009, p. 79–80). Every language shapes the world in its own manner and every word is specific to a language community. Whereas overt objectivity is practically inexistent, subjectivity and objectivity in language find themselves on a gradual scale, every lexical item being marked by a smaller or larger amount of subjectivity. For instance, one shall agree that the sentence *This flower is red* appears to be more objective than *This flower is beautiful*, implying that some lexical items (e.g., red) may be intrinsically more objective than others (e.g., beautiful).

Following Kerbrat-Orecchioni (2009, p. 65), here I call **subjective statements** those that imply a personal vision and interpretation of the referring expression by a speaker. This roughly corresponds to the subjectivity category that Wiebe et al. named “evaluation,” which includes “emotions such as hope and hatred as well as evaluations, judgements, and opinions” (2001). For instance, the adjective *expensive* would not be interpreted in the same way for someone belonging to the upper class as it would for someone from the working class. *This car was expensive*, therefore, would fall into the range of what I call subjective statements. An objective statement, on the other hand, would be *This car cost \$20,000*.

Language mirrors a speaker’s system of subjective values (Zerkina, Lomakina, & Kostina, 2015). In a subjective statement, a speaker evaluates a referential object (real or imaginary) in a given context. The evaluation is based on the relative norm of the speaker. For instance, a mountain is *big* or *small* according to the general

330 Even if the speaker is completely absent from an utterance: an unpersonal description can be greatly subjective (Kerbrat-Orecchioni, 2009, p. 169).

331 Balahur Dobrescu et al. summarize some proposals regarding subjectivity and sentiment analysis in social media (2014).

idea of the size of a mountain for the speaker or based on a comparison with another object (here a mountain) that would constitute the norm for the speaker (Kerbrat-Orecchioni, 2009, p. 98).

This evaluation can be 1) axiological or nonaxiological and 2) implicit or explicit. Axiological statements are those that are accompanied by a value judgment, which can be either positive or negative (Kerbrat-Orecchioni, 2009, p. 102). Nonaxiological statements, on the other hand, only imply a qualitative or quantitative evaluation of a referential object. The subjective evaluation can be implicit or explicit (e.g., compare *This flower is beautiful* with *I think this flower is beautiful*). The first sentence, though subjective, could appear to be objective at the first glance. In the second one, the evaluation is clearly linked to a specific speaker and thus to their opinion. But even when the speaker is completely absent from an utterance: An impersonal description can be greatly subjective (Kerbrat-Orecchioni, 2009, p. 169). As Finegan (1995) explained, subjectivity is marked in subtle ways in language (and is not marked in the same way in all languages), such as through morphology, intonation, or word order, and is challenging to explicate (p. 3).

Scholars have resolved the methodological issue of ordinary speakers' argumentation in different ways. Rheault, for instance, directly tackled the problem by adapting existing models of argumentation to "ordinary" speakers' utterances (2010) with categories of statements (descriptive, prescriptive, evaluative, etc.). Remysen (2009) circumvented the challenge by excluding de facto implicit statements³³² from his analysis (p. 54), and this is the approach I adopt here. Each unit was assigned a polarity indication: positive, neutral, or negative. The units were tagged using RQDA, a free³³³ R package for qualitative data analysis.³³⁴

8.3 Speakers' lexical criteria

The results of the qualitative analysis fulfill the second objective of my investigation, which is to explore speakers' lexical opinions in context (see Section 1.2.2). The analysis yielded three types of information:

.....
332 For instance, statements that call upon rhetoric topoi.

333 BSD license.

334 I deliberately decided not to use Knowtator because it was not able to handle such a large set of data (program crashes).

- Which items are being discussed and upon which an opinion is given
- Overall, what criteria speakers used when expressing opinions on lexical items
- For each item, upon what aspects opinions are being given and the polarization of opinions

The 4,090 opinionated autonym candidates allowed for an observation of 2,233 opinionated autonoms in context and 3,886 further autonoms that were commented, but not opinionated, upon (see full lists in the appendices). Opinionated autonym candidates concerning extended units of meaning and topics that did not directly concern specific lexical items³³⁵ were intentionally excluded from the analysis. The types of opinionated autonym candidates that were excluded consisted of

- full sentences or sentence parts
- abbreviations
- function lexical items (prepositions, pronouns, conjunctions, grammatical articles ...)
- phraseological units: sayings, binomials, proverbs, stereotyped constructions with function verbs, winged words, communicative formulas³³⁶
- gender-related pronouns (e.g., *ŝi*, *ŝli*, *ri*, etc.) and affixes (e.g., *vir-*, *ge-*, *-iĉ*)³³⁷
- proper nouns (including country names)

There were also contexts that I intentionally excluded because I did not consider them to be relevant for my research:

- autonoms that were commented upon solely for their pronunciation
- autonoms that were commented upon in relation to grammar:
 - grammatical collocations (e.g., when speakers commented which preposition should be used with a specific verb)
 - transitivity (e.g., when speakers commented whether a verb was transitive or intransitive)

335 See 2.4.1 for a definition of lexical item (p. 55).

336 I am using Fiedler's terminology here (1999a).

337 On gender-related lexical items in Esperanto, see Fiedler (2014).

- use of the accusative (e.g., when speakers commented upon which case should be used in a specific situation)

The types of opinionated lexical items revealed by the analysis varied greatly, from items of everyday language (household items, name of holidays, etc.) to specialized terminology (evidently vocabulary related to astronomy in the AEKO network, but also, for example, legal or IT terms in other electronic networks). Discussions about the names of countries or nationalities, an endless debate in the Esperanto speech community, were also largely present (but excluded). I classified speakers' lexical criteria into five subcategories:

1. Properties of the lexical item
2. Reference to language use
3. Reference to other items of language
4. Reference to language instances or resources
5. Extralinguistic statements

8.3.1 Properties of the lexical item

Eleven distinct categories emerged from the results: internationality, clarity, preciseness, subjective qualities, Esperantic nature, pronunciation, length, grammatical acceptability, neological character, official status, and frequency.

8.3.1.1 *Internationality*

As Esperanto is a planned language created specifically for international communication, the criteria of internationality are ubiquitous in the results. The search for internationality finds its origins in Zamenhof's "dek-kvina regulo" (the fifteenth rule) in the *Fundamento* (1905):

The so-called "foreign" words, i.e. words which the greater number of languages have derived from the same source, undergo no change in the international language, beyond conforming to its system of orthography. — Such is the rule with regard to primary words, derivatives are better formed (from the primary word) according to the rules of the international grammar, e. g. teatr'o, "theatre", but teatr'a, "theatrical", (not teatr'al'a), etc.

It must be noted here, as explained in 4.3.1 (p. 134), that Esperanto uses loan creation (e.g., the language borrows on the basis of a foreign model): It borrows foreign morphemes (and not lexical items) that are adapted to Esperanto grammar to form new lexical items.

Interpreted literally, this criterion strictly concerns the form of lexical items and its application is evidenced by the corpus. For instance, here an item is evaluated positively because its form is thought to be international:

*Plue, la verbo “prezidi” estas bone subtenata de la kognatoj en pluraj lingvoj: angla – “to preside”
itala – “presiedere”
hispana – “presidir”
portugala – “presidir”
franca – “présider”
greka – “προεδρεύω” (proedrevu)
rusa – “председательствовать” (predsedatelstvovatj)*

*Plus, the verb “prezidi” is well supported by cognates of several languages:
English – “to preside”
Italian – “presiedere”
Spanish – “presidir”
Portuguese – “presidir”
French – “présider”
Greek – “προεδρεύω” (proedrevu)
Russian – “председательствовать” (predsedatelstvovatj)*

Although the internationality of the form of a lexical item is generally positively evaluated, this orthographical rule sometimes reaches its intended function if speakers note that the form is identical but the meaning varies from one language to another (e.g., the items “pedologio,” “hieroglifo,” and here “kronometro”):

“Kronometro” estas bela ekzemplo pri vorto (tiaj abundas), kiu estas forme internacia, sed sence ne. Ofte oni supozas, ke oni povas doni al tiaj kvazaŭ – internaciaĵoj ĉiajn sencojn, kiujn oni trovas en sia propra lingvo. La sola sufiĉe internacia signifo de “kronometro” estas “precizega horloĝo”. La vorto originas el la 18-a jarcento kaj disvastiĝis tutmonde kiel nomo de ŝipa kronometro (ŝipa horloĝo, longituda horloĝo), la horloĝo kiu finfine ebligis al navigistoj determini sian longitudon.

“Kronometro” is a good example of a word (there are many such cases) of which the form is international, but the meaning is not. People often assume they can give such international words any one of the meanings that they have in their own language.

The only relatively international meaning of “kronometro” is “very precise clock.” The word comes from the 18th century and spread all over the world as the name of a marine chronometer (a marine clock; a clock to determine longitude), the clock which eventually allowed shipmen to determine their longitude.

Sometimes speakers also extend the meaning of “international form” to lexical items whose formation methods are similar to formation methods found in other languages:

*Mi dirus “halthorloĝo”, ĉar eblas haltigi ĝin, kaj tiu vorto estas analogo al kunmetaĵo uzata en tre multaj lingvoj
I would say “halthorloĝo” [stop clock] because you can stop it, and this word is analogous to a compound used in many languages*

In the Esperanto speech community, the internationality of a lexical item is also often associated with its intelligibility for speakers of various languages:

*Konstelacio estas plibone komprenata internacie vorto.
Konstelacio is a word that is better understood internationally.*

The main argument for preferring international word forms (and meanings) is based on the desire to be understood by the greatest number of fellow speakers, which leads to the criteria of clarity.

It should be noted here that some speakers perceived that the criteria of internationality is problematic because its definition is unclear:

*La vorto “internacia” estas unu el la plej problemaj. Ne nur, ke kvin personoj havas kvin malsamajn opiniojn pri ĝia signifo, montriĝas eĉ, ke la sama persono komprenas ĝin malsame ĉe malsamaj problemoj.
The word “internacia” [international] is one of the most problematic ones. The problem is not only that five people have five different opinions about what it means, but it turns out that even a single person has a different understanding of it in different contexts.*

The criterion of internationality is mostly evaluated in an exclusively synchronic way by speakers, although lexical change in Esperanto and the languages from which it has borrowed lexical material has created **semantic false friends**; that is, words that are graphically and/or phonetically similar in various languages, but whose meanings diverge (Domínguez Chamizo & Nerlich, 2002, p. 1836).

8.3.1.2 Clarity (intelligibility)

Speakers repeatedly used the adjectives “klara” (clear) and “malklara” (unclear) to describe lexical items, for instance:

Ankaŭ mi pensas, ke ‘sojla maso’ aŭ ‘sokla amaso’ estas la plej klara esprimo pri tio.

I too think that “sojla maso” or “sokla amaso” is the clearest expression about this.

Plej klara estus “dom-invadanto”, aŭ simila, sed eble indus ekuzi novan verbon “skvati”, jam tre internacian.

The clearest [expression] would be “dom-invadanto,” or something similar, but maybe we could start using a new verb “skvati,” which is already relatively international.

However, the meaning they give to “clear” seems to vary. In the results, at times clarity (clear) appears to mean the use of a lexical item that can be understood by fellow speakers, and/or avoiding misunderstandings with other speakers. Clarity is then to be understood as the intelligibility of a lexical item for fellow speakers:

Mi uzus la terminon “instruistiĝa lernejo”, ĉar laŭ mi tiu termino estas la plej klara. Oni tuj scius pri kia lernejo temas.

I would use the term “instruistiĝa lernejo” because I think that this term is the clearest. You would immediately know what kind of school this is.

Tamen, mi preferas la vorton “kikerkaĉo”. Ĝi estas komprenebla por tiuj kiuj ne konas la pladon.

However, I prefer the word “kikerkaĉo.” It is intelligible for those who don’t know the dish.

La esprimon “senmarkaj medikamentoj” mi ne komprenus. “Senpatentaj” estas tute klara.

I would not understand the expression “senmarkaj medikamentoj”. “Senpatentaj” would be absolutely clear.

Other results suggest, however, that the notion of clarity is not systematically equivalent to that of intelligibility. For instance, the following speaker distinguishes between “klara” (clear) and “vaste komprenata” (largely understood):

Laŭ mi preferindas la esprimo “sojkazeo”. Ĉi tiu formo estas klara, vaste komprenata kaj uzata.

According to me, the expression “sojkazeo” should be preferred. This form is clear, largely understood, and used.

In some results, the notion of clarity seems to be used by speakers to refer to the transparency or motivation of a lexical item, such as in the following example:

Mi preferas la vorton “puŝlevo”; “kuŝapogo” ŝajnas al mi malklara.

I prefer the word “puŝlevo;” “kuŝapogo” seems unclear to me.

Laŭ mi, por iu ajn celo (muelado, elektrogenerado ktp) estas uzebla “ventopova maŝino”. Simpla “venta maŝino” aŭ “ventomaŝino” estas dubsenca, supozigante ankaŭ maŝinon, kiu produktas venton.

According to me, for any purpose (milling, generating power, etc.) you can use “ventopova maŝino”. The meaning of simple words such as “venta maŝino” or “ventomaŝino” is doubtful, leaving the possibility for any machine producing wind.

8.3.1.3 *Preciseness (transparency and motivation)*

The criterion of transparency or motivation seems to be more often expressed in terms of “preciseness” of lexical items:

Laŭ mi, “Usonfutbala pilko” pli precize priskribas la specon de piedpilko uzata por usonfutbalo.

I think that “Usonfutbala pilko” more precisely describes the type of ball used to play American football.

“Tensio-protektilo” estas nepreciza esprimo laŭ mi, ĉar protektas ne la tension, sed nian aparaton kontraŭ trotensio.

“Tensio-protektilo” [device to protect from tension] is not a precise expression for me because the protection is not against electric tension, but against surge.

Prefere ni uzu diversajn vortojn por diversaj sencoj.

La lingvo devas esti preciza.

We should preferably use different words for different meanings.

The language must be precise.

8.3.1.4 Subjective qualities

Speakers make use of a wide range of positive adjectives to qualify the lexical items of which they approve (“rimarkinda” remarkable, “bonega” very good, “eleganta” elegant, “normala” normal, and “sprita” witty) ...

mi tre ŝatas la vorton “nemalhavebla” kaj trovas ĝin sufiĉe eleganta :)

I really like the word “nemalhavebla” and find it to be relatively elegant :)

... and a wide range of negative adjectives for items of which they do not approve (“stranga” strange, “duba” doubtful, “ne tro eleganta” not really elegant, “malbona” not good, “malkera” not educated, “nenatura” unnatural, “artefarita” artificial, “perforta” violent”, “malbela” ugly, “malglata” harsh, and “pedanta” pedantic).

Pli stranga ŝajnas al mi la vorto “ĉopero” menciata sampaĝe, sed ŝajne cetere neuzata.

More strange to me is the word “ĉopero,” which is mentioned on the same page, but apparently is not used, either.

Interpretation is difficult in such cases, but the opinions that are expressed are clearly polarized. Another subjective aspect observed in the result is discrimination: A lexical item may be disqualified if a speaker evaluates it as discriminatory:

pro tio, la propono “hindismo” estas laŭ mi ne apoginda. ĝi _ne_ estas simpliga. ĝi estas diskriminacia.

For this reason, I believe that the suggestion “hindismo” should not be encouraged. It is NOT simplifying. It is discriminatory.

8.3.1.5 Esperantic nature

Remysen had found Frenchness to be an argument that columnists use to evaluate lexical items in the Canadian context (2009). In a similar way, the Esperantic nature of lexical items play a role for some speakers:

La termino “praeksplodo” estas bela kaj esperantega vorto por Big Bang. The term “praeksplodo” is beautiful and a very esperantic word for Big Bang.

In speakers’ terms, the adjective “esperanta” (esperantic) is usually used to mean a lexical item that has been created from indigenous roots rather than lexical items that would be created on the basis of foreign roots, as the following example illustrates:

“condominium” ŝajne jam fariĝis relative internacia termino; ne estas do surprize ke iu esperantigis ĝin al “kunderminiumo” (1), sed laŭ Google, tiu vorto ne populariĝis. Brazila retejo kreis pli esperantecan terminon: “kunproprietajo” (2) sed ne tuj evidentiĝas, ke temas pri loĝejo.

“condominium” has apparently become a relatively international term, so it is not surprising that someone used “kunderminiumo” in Esperanto (1). However, according to Google, this word has not become popular. A website in Brazil created a more esperantic term: “kunproprietajo” (2), but it is not self-evident at all that it refers to a place for living.

8.3.1.6 Pronunciation

Pronunciation also plays a role as a lexical criterion for speakers. Types of lexical items that received negative judgements in the corpus regarding pronunciation were foreign words (e.g., names of products, proper nouns):

Ni povas skribi iPhone ktp, sed kiel prononci ilin? Ĉu aj-fono aŭ i-fono [...] We can write iPhone, etc., but should we pronounce them [these words]? Would it be aj-fono [I-phono] or i-fono [ee-phono]?

However, lexical items with chains of consonants were also deemed (too) hard to pronounce:

Mi volas ankaŭ atentigi, ke la vorto “preciZPafisto” estas praktike neprononcebla, ĉar en ĝi estas unu apud la alia voĉa kaj sen[v]oĉa konsonantoj. I also want to draw attention to the fact that the word “preciZPafisto” [a candidate equivalent for sniper] is practically impossible to pronounce because it contains a voiced and an unvoiced consonant one after the other.

Se oni nepre sentas la bezonon uzi “matĉo” aŭ “maĉo”, mi rekomendus “maĉo”, ĉar la malfacilega prononco de “matĉo” laŭ mi estas pli grava problemo ol la homonimo de “maĉo” [...]

If there is an absolute need to use “matĉo” or “maĉo,” I would recommend “maĉo” because the really difficult pronunciation of “matĉo” is a more important problem than the fact that “maĉo” is a homonym.

8.3.1.7 Length

All things being equal, some speakers mention that they prefer more concise lexical items:

Inter “memplenumiĝanta” kaj “memplenuma”, mi preferas la pli mallongan vorton, simple ĉar ĝi estas pli mallonga, kaj egale klara.

Between “memplenumiĝanta” and “memplenuma,” I prefer the shortest word just because it is shorter yet equally clear.

“Concise” or “short” is a subjective notion as well, but when listing the lexical items that speakers positively evaluated as short, one sees that an item with around 10 letters is usually in the concise category.

Lexical item	Number of characters
lumeco	6
lavkuvo	7
logopedo	8
miso-sceno	10
presindaĵo	10
memplenuma	10
plumpilkado	11
pordotirilo	11
socia retejo	12

Table 27. Examples of lexical items that received positive comments for their conciseness.

In contrast, starting at around 14 characters, lexical items received negative comments in relation to their length. There are exceptions; for instance, the item “patrino” (mother), with only seven letters, was deemed to be too long for a word to be used in family situations where children would prefer the shorter version “panjo.”³³⁸ Table 28 offers a few examples from the results:

Lexical item	Number of caracters
patrino	7
introvertitulo	14
introvertiteco	14
konsiderindeco	14
stangodatumujo	14
transkontigado	14
naskiĝdatreveno	15
neŭrotransmisilo	16
perventa aparato	16
elparolkuracisto	16
neŭrotranssendilo	17
senpilota aviadilo	18
deviga malĉeesto ekstere	24
ekspansiigita polistireno	25
ringe bindita studmaterialo	27

Table 28. Examples of lexical items that were judged to be too long by speakers.

Above 20 characters, lexical items may receive extremely negatively polarized comments:

*Mi konsentas kun la duboj esprimataj pri “etendi”. “Ekspansiigita polistireno” ŝajnas ĝis nun plej trafa, sed kia vortomonstro!
I agree with the doubts brought up about “etendi.” “Ekspansiigita polistireno” now appears to be the most appropriate form, but what a monstrous word!*

It should also be noted that, in Esperanto, there are also relationships between the length of lexical items and polysemy: short lexical items are, on average, more likely to be polysemous (see Kück, 2009, pp. 77–78).

338 As in English, a mother is usually called “mom” at home by her children instead of “mother.”

8.3.1.8 Grammatical acceptability

Speakers further paid attention to Esperanto grammatical rules (as they understood them). They repeatedly rejected any lexical formations that seemed redundant, for instance:

Laŭ mi povus esti “kelkope” aŭ “plurope” aŭ “multope,” sed ne “grupope,” ĉar la vorto “grupo” jam diras, ke iuj estas almenaŭ kelkaj.

According to me, it could be “kelkope,” “plurope,” or “multope,” but not “grupope,” as the word “grupo” [“group”] already indicates that there are several people involved.

Tables 29 and 30 provide further examples of acceptability (and the lack thereof) in terms of grammar.

Lexical item	Criticism
grupope	Redundancy of -op-
eklumigi	Redundancy of ek-
distordi	Pleonasm
demokrateco	Unappropriate derivation of the root “demokrati-”
sinmortigi	Unappropriate back-formation
voĉdoni	Improper compound

Table 29. Examples of lexical items whose grammar was deemed to be inappropriate by speakers.

Lexical item	Criticism
gebopatroj	Compliant with Esperanto grammar ³³⁹
kioi	Verb is theoretically possible although inexistant
subtekstigi	Preferred to “subteksti” because “tekst-” is a noun root

Table 30. Examples of lexical items whose grammar was deemed to be appropriate by speakers.

.....

339 The discussion was mainly about “gebopatroj” versus “bogepatroj” to mean “parents-in-law”. Both “ge” (of both sexes, i.e. a man and a woman here) and “bo” (in-law) are prefixes in Esperanto, patr- is a root meaning “parent”, -o is the noun ending and -j the plural ending. Some speakers argued about the order of these prefixes within the lexical item.

8.3.1.9 Neologic character

In Esperanto, “neologismo” can have two meanings: both to refer to a new form or new meaning of a lexical item (the traditional meaning in linguistics) and to refer to a new root introduced into the language. Many speakers prefer to resort to existing roots rather than to import foreign lexical material and use so-called “neologismoj”:

Aldone, mi ĉiuokaze preferas eviti la neoficialan novradikon “semolo,” kaj uzas esprimojn kun “grio” aanstataŭe.

In addition, in any case, I prefer to avoid the unofficial new root “semolo,” and I use expressions with “grio” instead.

Instead, many speakers prefer creating a new word (which linguists would call a *neologism* but Esperanto speakers would not) by combining existing Esperanto roots and assigning the combination a new meaning:

Estas kompleksa fenomeno, tio, kion brazilanoj komprenas sub la vorto “favelo.” Nemirinde nederlandano tion ne komprenas. Tamen mi ne estas entuziasma pri la enkonduko de novaj radikoj. Mi plu uzas “ladkvartalon,” ĉar tiu “lad-” elvokas la provizorecon de la kabanjoj.

For Brazilians, the word “favelo” [favela] refers to a complex phenomenon. It’s not surprising that a Dutch person would not understand. However, I am not in favor of introducing new roots. I prefer using “ladkvartalo” [literally, “sheet-metal neighborhood”] because the root “lad-”[“sheet metal”] evokes the temporary nature of the shacks.

“Neologismoj” (new roots) are generally viewed negatively, and many Esperanto speakers use “neologismoj” only if they believe there is no other option:

Mi ĝenerale ne estas neologismema, sed mi ja sentas bezonon por nova vorto ĉi tie.

Generally speaking, I am not in favor of neologisms [new roots], but here I feel that I need a new word.

Mi vidis en rusa vortaro la varianton “paflertulo.” Verdire, tio ne tre plaĉas al mi. Laŭ mi nun estas kazo, kiam indas provi ekuzi neologismon. El la

antaŭaj proponoj ĉi tie mi emas uzi “snajpero.” Aŭ prefere eĉ “snajpisto,” ĉar ne temas pri “ero de snajpo.”

I saw “paflertulo” in a Russian dictionary. To be honest, I don’t really like it. This is a case in which it is worth trying to start using a neologismo [new root]. Of the previous suggestions, I would use “snajpero” or even “snajpisto” rather than “ero de snajpo.” [The root “snajpero” could also be analyzed as snajp-er-o, and the suffix “-er” is used to mean “single, individual, or unit,” which makes the suggestion ambiguous.]

Another issue is not directly linked to lexical criteria but is of the utmost interest for language managers: Around opinionated autonyms, speakers may indicate lexical items with new meanings (*neologisms* in linguists’ usage). For instance, they may apply a lexical item that they have found in use but that is missing from dictionaries:

La vorto “voki” en la senco de “telefoni” estas nek en PIV, nek en ReVo. sed mi trovis en:

<http://www.uea.org/kongresoj/2003/duabulteno.html>

kaj

<http://www.geocities.com/wfpilger/slangol3.htm#v>.

The word “voki” [“to call”] in the meaning of “telefoni” [“to make a phone call”] is in neither PIV nor ReVo, but I found it at

<http://www.uea.org/kongresoj/2003/duabulteno.html>

and

<http://www.geocities.com/wfpilger/slangol3.htm#v>.

Speakers further indicate when existing lexical items shift in meaning:

La vorto “sabeliko” estas Zamenhof-devena, sed PIV indikas “krispa brasiko” kiel difinon. Laŭ mi tute ne gravas, principe, ke la vorto alprenu novan signifon, tamen ...

The word “sabeliko” comes from Zamenhof, but PIV indicates “krispa brasiko” [crispy cabbage] as a definition. For me, it is okay, in principle, for the word to take on a new meaning, but ...

8.3.1.10 Official status

Some speakers choose to give preference to an item or a spelling from the Fundamento whenever possible:

Kial en la dua frazo oni uzas la formon “kaoso,” kvankam ReVo per referencoj ĉe “kaoso” rekomendas la F-tan vorton “ĥaoso”?

In the second sentence, why has the form “kaoso” been used when ReVo, in its entry on “kaoso,” recommends the word “ĥaoso” from the Fundamento?

These attitudes are not homogeneous, as others welcome lexical items that are not in the Fundamento as official items:

Jes, laŭ mi “fasado” aŭ ankaŭ la fundamenta “fasono” tute taŭgas.

Yes, according to me, both “fasado” and “fasono,” which is from the Fundamento, are completely appropriate.

8.3.1.11 Frequency

Many speakers use online bodies of text to determine the frequency of lexical items. This can be done via state-of-the-art corpora (e.g., Tekstaro.com), search engines (e.g., Google), or other sources (e.g., Wikipedia). Some speakers seem to be aware that such techniques are not representative of the whole language:

Mi havas la impreson ke “hiperligo” estas iom pedanta kaj efektive malofte uzata; ekzemple sur la paĝo http://eo.wikipedia.org/wiki/Helpo:Enhavo_aperas_pluraj_ligiloj,” sed neniu “hiper”-vortoj.

Mi serĉis “hiperligo” en la tuta vikipedio kaj trovis nur 5 aperojn! “Ligilo” aperas pli ol 20000 fojojn. “Ligo” aperas 10000 fojojn.

I have the impression that “hiperligo” is a bit pedantic and actually not used often; for example, the page http://eo.wikipedia.org/wiki/Helpo:Enhavo_aperas_pluraj_ligiloj contains several uses of “ligiloj” but no words with “hiper.”

I searched for “hiperligo” on all of Wikipedia and only found five occurrences! “Ligilo” appears more than 20,000 times. “Ligo” appears 10,000 times.

However, others do not hesitate to use the frequency argument to make a point in a discussion. For instance, some use a simple Google search as evidence to convince fellow speakers (or themselves) of the frequency of lexical material in real language use:

Krome, eta Gugla esploro pruvas, ke en reala lingvouzo, la neologismaj uzoj de la radiko "mobil/" jam estas pli oftaj ol la PIV-a.

In addition, a quick Google search proves that, in real language use, the neologic use of the root "mobil/" is already more frequent than the use given in PIV.

It should be noted that evaluating lexical items based on frequency measures goes against the principles of the so-called analytical school (see, for example, Philippe, 1991, p. 77–78) and stands in clear contradiction to Zamenhof's 22nd answer³⁴⁰ from 1907 (Zamenhof, 1990, Respondo 22):

Sed en Esperanto la "nekutimeco" ne prezentas gravan kaŭzon por neuzado ...

However, in Esperanto, the "unusual character" does not represent an important reason not to use something.

8.3.2 Reference to language use

Speakers also sometimes pay attention to which lexical items are being used and to who is using them. Here, I apply four categories: general language use, idiosyncratic use of lexical items, specific written language use, and language use by famous Esperantists (especially Zamenhof).

8.3.2.1 Language use in general

Generally, speakers view those lexical items that they believe the speech community uses often in a more positive manner than they view those that are not used or that are used only rarely. In the following excerpt, for instance, the speaker states that, although the item "artisma traduko" is appropriate, it is not used often, so the speaker decides to opt for another item ("arta traduko"):

.....

340 Zamenhof answered speakers' doubts and questions in a series of *Lingvaj Respondoj*.

“Artisma traduko” estas en si mem sufiĉe klara kaj utila koncepto, sed praktike ĝi ne estas uzata en tiu formo. Same pri la ĵus proponita “arteca traduko.” Mi decidis ŝanĝi la terminon al “arta traduko.” Tiu termino estas vaste uzata.

“Artisma traduko” [artistic translation] in itself is a relatively clear and useful concept, but it is not used in practice in this form. The same holds for “arteca traduko.” I thus decided to use the term “arta traduko” instead. This term is widely used.

If two items are competing, the criteria of general language use can help speakers to decide between them.

Sed nun ŝajnas temi pri “pipalo” kontraŭ “bodiarmo,” kaj tiam laŭ mi preferindas “bodiarmo,” ĉar tiu vorto jam estas uzata, kaj jam estas registrita en nia plej grava vortaro.

However, now it seems to be a choice between “pipalo” and “bodiarmo”; in that case, I think “bodiarmo” is preferable because this word is commonly used and is already registered in our most important dictionary.

To some speakers, this is one of the most important lexical criteria:

Se la ĝenerala publiko akceptos alian solvon, mi respektos ĝin (same kiel mi respektas venkon de komputilo kvankam mi mem antaŭe preferis komputoro, sed mi pretas akcepti malvenkon:), sed ĝis tiam mi restos ĉe mia esprimo.

If the general public accepts another solution in the future, I will respect that (just as I respected the victory of “komputilo” even though I preferred “komputoro,” as I am willing to accept defeat), but until then, I will stick to my expression.

However, language is not always sufficient to convince a speaker:

Mi ofte aŭdas Esperantistojn uzi la vorton “logotipo,” sed mi ne scias en kiu fonto ĝi aperas krom en homa uzado. Eble ĝi ne estas bona Esperanto.

I often hear Esperanto speakers use the word “logotipo,” but I am not sure what source it appears in, aside from this common use. Maybe it’s not good Esperanto.

8.3.2.2 Speakers' own language use

Declarations of the speakers' own usages abound in the results. Often, the speakers simply mention the lexical items that they use without arguing in its favor (for instance, when providing a suggestion to a fellow speaker). They only seldom provide explanations regarding the items that they intentionally use or avoid using.³⁴¹

Cetere mi neniam uzas "ttt-ejo", sed nur paĝo aŭ paĝaro, ĉar "ttt-ejo" estas terure nekomprenebla (por komencantoj) kaj malfacile elparolebla. Besides, I never use "ttt-ejo"—only "paĝo" or "paĝaro"—because "ttt-ejo" is awfully unintelligible (for beginners) and is hard to pronounce.

Kiu estas laŭ vi bona traduko por "pen drive"?

Mi kutime uzas la vorton "stangodatumujo", Mi scias, ke ĝi estas iom longa vorto, sed mi ŝatas ĝin, kaj ŝajnas al mi facile komprenebla.

What do you think is a good translation of "pen drive"?

I usually use the word "stangodatumujo" I know that it's a bit of a long word, but I like it, and it looks easily intelligible to me.

8.3.2.3 Use in written sources

Some of the excerpts in the results provide direct evidence that specific written sources directly influence speakers' lexical environments. The speakers cite a large variety of sources, from vague statements that a source came from "somewhere on the Internet" to much more precise statements. Table 31 provides a few examples.

341 Whether their declared language behavior matches their actual language behavior is an interesting question that is not addressed in the present investigation.

Quote	Source mentioned
<p>Ofte oni legas “EU-Komisiono,” kvankam la oficiala nomo estas “Komisiono de la Eŭropa Komunumo,” ĉar nejuristoj ofte ne konas la diferencon inter EU kaj EK.</p> <p><u>You often read “EU-Komisiono,”</u> but the official name is “Komisiono de la Eŭropa Komunumo”; people who are not jurists often cannot distinguish between the EU and this group.</p>	Somewhere
<p>En interreto mi trovis po unu ekzemplon de “neŭrotransmisiilo” kaj “neŭrotranssendilo” (iom pezaj).</p> <p><u>On the Internet, I found</u> examples of the use of “neŭrotransmisiilo” and “neŭrotranssendilo” (which are a bit cumbersome).</p>	The Internet
<p>Mi neniam antaŭe renkontis miskomprenon de la vorto “koro,” t. e., organo; tamen, en Universalaj Kongresoj mi renkontis afiŝojn pri la “Internacia Koruso.”</p> <p>Until now, I’ve never encountered a misunderstanding of the word “koro,” which means “organ,” but during the World Congresses, <u>I saw posters referring to the “Internacia Koruso.”</u></p>	Posters seen during congresses
<p>Cetere, Michiel Meeuws uzas “asistant-residento” <u>en sia traduko</u> “Saidjah kaj Adinda” (http://meeuw.org/saidjah/), sed mi dubas ĉu tio estas trafa elekto. Moreover, Michiel Meeuws uses “asistant-residento” <u>in his translation</u> “Saidjah kaj Adinda” (http://meeuw.org/saidjah/), but I doubt this is an appropriate choice.</p>	A translation found on the Internet
<p><i>En Monato</i> jam kelkfoje estis uzata la vorto “birilo.”</p> <p><u>Monato</u> [a monthly magazine] sometimes uses the word “birilo.”</p>	A monthly print magazine
<p><u>Raporto en revuo Esperanto de UEA</u> tradukis la terminon “European Ombudsman” per “Eŭropa Mediatoro.”</p> <p><u>A report in UEA’s magazine Esperanto</u> translated the term “European ombudsman” as “Eŭropa Mediatoro.”</p>	A report
<p>La vorto “nepereebbla” aperas en la romano Kastelo de Prelongo.</p> <p>The word <u>“nepereebbla”</u> appears in the novel <i>Kastelo de Prelongo</i>.</p>	A novel
<p><u>En la almanako</u>, oni legas la terminojn “planedeca nebulozo” kaj “planeda nebulozo.”</p> <p><u>The almanac</u> contains the terms “planedeca nebulozo” and “planeda nebulozo.”</p>	An almanac about astronomy

Table 31. Examples of written sources in which speakers find lexical items.

Speakers do not systematically provide opinions about the lexical items that they find in such sources, but this type of information is of great interest for language managers because it reveals the sources that the speakers consult regarding lexical material. These contexts provide additional information for explorations of speakers' lexical environments (which relates to the first objective of this investigation; see Section 1.2.1).

8.3.2.4 *Use by famous authors (Zamenhof)*

The opinions of Zamenhof, the initiator of the language, have a symbolic linguistic power in the Esperanto speech community, and speakers frequently refer to his writings (Lo Jacomo, 1981, p. 346), although Zamenhof did not claim authority over the language (see Privat, 1912, p. 21–22). His usages still play a major role in some speakers' lexical environments, and the lexical items that he used enjoy a certain amount of credibility:

Noto pri tiuj du vortoj: Z jam uzis “referaton,” do malgraŭ neoficialeco ĝi havas ian validecon

Regarding these two words, Zamenhof used “referato,” so, despite the fact that it is not official, it has some kind of validity.

Speakers repeatedly referred to the fact that Zamenhof used a lexical item as an argument in favor of that item:

Eble indas aldoni, ke la vorto “fajrilo” estas Zamenhofa.

It may be worth adding that Zamenhof used the word “fajrilo.”

However, for this criterion as well, some speakers remain critical and do not hesitate to question Zamenhof's lexical items:

Sinonimo de “knedujo,” eble iom arkaika, estas “pastujo,” kiun Zamenhof uzis en la traduko de la Malnova Testamento.

Zamenhof used a perhaps-outdated synonym for “knedujo,” “pastujo,” in his translation of the Old Testament.

Other speakers even clearly mention that they would prefer to give up some of Zamenhof's lexical items:

Por ĉi tio mi prefere uzus pli klaran “novvorto.” Tia uzo ekzistas ĉe Z, sed ĝi estas tre speciala kaj malmulte uzata.

For this, I would rather use a clearer novvort [“new word”]. Such a usage exists in Zamenhof, but it is very specific and rarely used.

8.3.3 Reference to other language items

8.3.3.1 Comparison to existing Esperanto lexical items

If a proposed lexical item is similar to an existing one, speakers usually evaluate it positively:

PIV2 eĉ jam enhavas la vorton “vintrostacio.” Do “plaĝstacio,” “somerstacio” kaj simile ŝajnas al mi bonaj solvoj.

PIV2 already has the word “vintrostacio” [vintr-o-staci-o]. Thus, “plaĝstacio” [plaĝ-staci-o], “somerstacio” [somer-staci-o], and similar words seem to be good solutions.

Mi dirus “neseksema,” ĉar ni jam diras “samsek[s]ema,” “alipseksema,” kaj “ambaŭseksema.”

I say “neseksema” [ne-seks-em-a] because we already say “samseksema” [sam-seks-em-a], “alipseksema” [ali-seks-em-a], and “ambaŭseksema” [ambaŭ-seks-em-a].

“Superfluido” estas io analogia al “supersolido,” “superlikvo,” “supergaso.” “Superfluido” [super-fluid-o] is analogous to “supersolido” [super-solid-o], “superlikvo” [super-likv-o], and “supergaso” [super-gas-o].

8.3.3.2 Comparison to existing foreign lexical items

Opinions are divided on the use of allogenisms³⁴² in Esperanto. Some speakers seem to be averse to any lexical items that resemble ones from other languages. This can be the case, for instance, if the lexical material comes from a dominant language such as English:

.....
³⁴² Allogenisms: “lexical constructions made in one language using material from another language” (Humbley, 2015, p. 35).

Mi neniam uzis la vorton “brodkasti,” ankaŭ por mi ĝi sonas kiel evitinda anglismo, k nia vivo nun estas troplenigita per kriplaj anglaj vortoj, uzataj senbezone.

I never use the word “brodkasti,” for it sounds like an Anglicism that should be avoided, and our lives are already hammered with deficient English words that are used in vain.

Oni kompreneble ne diru, aŭ skribu “haloveno” nek “halovino.” Tiuj du vortoj estas tro anglosonaj. La ĝusta traduko estas “festo de ĉiuj sanktuloj” aŭ “sanktulara antaŭvespero”!

Of course, you should neither say nor write “haloveno” or “halovino.” These two words sound too much like English. The correct translation is “festo de ĉiuj sanktuloj” or “sanktulara antaŭvespero”!

However, because Esperanto is a language targeted toward international communication, some speakers have the impression that certain lexical items have emerged due to the (undesired) influence of someone’s native language:

Oni emas misuzi la vorton “korekta” sub influo de sia denaska lingvo. People like to abuse of the word “korekta” due to the influence of their native languages.

Foreign languages nevertheless remain a source of inspiration for the creation of new lexical items, as in these examples:

Cetere, mi tre ŝatas la japanan manieron diri “vicfinalo” kaj “vicvicfinalo.” La kvazaŭkalkula maniero diri “duonfinalo,” “kvaronfinalo” ktp. ŝajnas al mi tre faka. “Tridekduonfinalon” jam neniu ordinara homo klare komprenas.

What’s more, I really like how Japanese speakers say “vicfinalo” and “vicvicfinalo.” Using “duonfinalo,” “kvaronfinalo,” and so on in a sort of mathematical way looks like jargon to me. When you reach the point of “tridekduonfinalo,” no ordinary person would understand you.

Some speakers are even in favor of Anglicisms:

Mi ja estas ruso, sed mi preferus, ke novaj vortoj estu prunteprenataj pli ofte ne el la rusa, sed el la angla, ĉar, bedaŭrinde (aŭ feliĉe, kiel oni preferas), la

lasta multe pli taŭgas por tio ... Konklude: oni ne provu blinde kondamni la anglan, sed oni utiligu tion pozitivan, kion ĝi enhavas.

Well, although I'm Russian, I prefer new words that are borrowed not from Russian but from English, as unfortunately (or fortunately, depending on your preference), English is better suited to this purpose ... In conclusion: rather than blindly condemn English, you should try to make the most of the positive things it has to offer.

8.3.3.3 Polysemic lexical items

For speakers, two types of polysemy constitute arguments against lexical items: internal and external. *Internal polysemy* refers to a root or lexical item that has more than one meaning in Esperanto. As a speaker explains in the following example (with a dash of humor), internal polysemy can be a real barrier within the Esperanto speech community:

Se filologio bakus esperanton, anstataŭ volapukismo, tiu arbara dio fariĝus "panoso," anstataŭ "pajno," el la genitiva formo (HE πάν, πανός).

La formo "pano" eĉ pli bonus, sed la maniuloj de la sistemo "unu vorto por ĉiu ideo" ne permesas tion. Ili timus manĝi la dion, fari el li sandviĉojn, kaj tio por la maniuloj estus hibriso (HE ύβρις) al esperanto. Oni povus esti kreinta la vorton "pandio," sed "ho, ne, ne eblus, ĉar tuj oni pensus, ke temas pri dio de la pano."

If a philologist had coined Esperanto expressions instead of Volapukisms, the god of the forest would be a "panoso" instead of a "pajno," from the genitive form (Greek: πάν, πανός).

The form "pano" [which means "bread" in Esperanto] would be even better, but the maniacs who apply the "one word for every idea" system don't allow it. They fear eating this god and making sandwiches out of him, as that would be a "hibris" (Greek: ύβρις) [an insult] to Esperanto. One could have created the word "pandio," but "oh no, no, this wouldn't be possible because one would immediately think this is the god of bread" [as "pandio" could be interpreted as "pan-di-o," literally "the god of bread"].

Not all Esperanto speakers are “maniacs” about polysemy, and some of them tolerate it, especially if they do not see another option:

Kvankam unu el la celoj de Zamenhof estis, ke ĉiu vorto en Esperanto havu nur unu signifon, mi trovas, ke tio ne ĉiam estas ebla.

Although one of Zamenhof’s goals was for every word in Esperanto to have only one meaning, I find that this is not always possible.

External polysemy, on the other hand, refers to cases in which the lexical item is assumed to be polysemous in Esperanto because speakers could interpret it differently, depending on their native languages:

Mi neniam uzas la vorton “brava,” precipe pro la plursenceco. Anglemuloj interpretas ĝin en “la brava soldato Ŝve’jk” kiel “kuraĝa,” dum nederlandanoj kaj germanoj interpretos ĝin kiel “(tro) ribel-malema”

*I never use the word “brava,” mostly because it is polysemous. Anglicized people interpret it as meaning “valiant,” as in *La Brava Soldato Ŝve’jk* [The Good Soldier Ŝvejk, a novel written by Jaroslav Hašek and translated into Esperanto], whereas Dutch and German people think it means “too subordinate.”*

8.3.4 Reference to language instances or resources

8.3.4.1 *The Akademio de Esperanto*

The speakers rarely mention the Akademio de Esperanto in the results, but some speakers remind others of this group’s recommendations:

Eble indus memori, ke “Ukrajn/o” kiel landnomo (kvankam mi konsentas, ke estas evitinda), tamen estas la de la Akademio rekomendita formo.

Maybe it is worth remembering “Ukrajn/o” as the name of a country (though I agree that it should be avoided), as that is still the form that the Akademio recommends.

8.3.4.2 The PIV dictionary

In the corpus results, the speakers' attitudes toward the PIV dictionary vary from very negative to extremely positive, including speakers who have absolutely no trust in PIV, as in these examples:

Ho, plue: mi neniam rekomendus ke oni uzu la PIV-on.

Well, regardless: I would never recommend that anyone use PIV.

Povas esti, mia PIV estas nur malnova, sed PIV ne estas fidinda. Fidu vian propran koron kaj inteligentecon, ĉu ne?

It could be [that the word "gepatro" is in PIV], as my PIV is old, but PIV is not reliable. You should trust your own heart and intelligence, don't you think?

Other speakers think that PIV is useful if used carefully:

Ĝi [PIV] do estas ekstreme danĝera laborilo, se oni senplie diras al si: "Tio enestas en la PIV, do tio estas uzinda." Sed ĝi estas tre utila laborilo, se alimaniere oni rilatas al ĝi.

It [the PIV dictionary] is an extremely dangerous tool if you simply say to yourself, "This is in PIV, so it should be used." However, it's a very useful tool if you adopt another approach toward it.

Finally, some speakers think that PIV remains the unconditional language reference:

Kaj mi denove asertas, ke mi multe amas vin ambaŭ, [persona nomo] kaj [persona nomo], sed mi pli amas Esperanton, kaj konsekvence por mi pli gravas PIV 2002 ol iu ajn pli demokratia, pli anonima kaj malpli fidinda retvortaro, ne gravas kiel fama ĝi estas por sia "virtuala" publiko. Dum la Akademio silentas, PIV (kun individua apogo de 20 akademianoj) restas la sola garantio, ke nia lingvo ne dispeciĝos.

I state once more that I love you both a lot, [personal name] and [personal name], but I love Esperanto more than I do you, so, for me, PIV 2002 is more important than any of the more democratic, more anonymous, and less trustworthy online dictionaries, regardless of how famous those are to the "virtual" public. Although the Akademio remains silent, PIV (with the

support of 20 members of the Akademio) remains the only guarantee that our language won't fall to pieces.

From these results, it seems that PIV is predominantly used as a reservoir of information rather than as a justification for the acceptance of lexical items. Some speakers do use PIV to argue in favor of lexical items, but they seem to be an exception in the corpus results:

*Ĝuste tian uzon de “certa” mi (kaj ankaŭ PIV2005) plene aprobas.
It is precisely this usage of “certa” that I (and PIV2005) completely approve of.*

8.3.4.3 Other language resources

As shown in Chapter 6 (Section 6.4.1), some focus group participants report using dictionaries (both general and specialized). This is confirmed in the results of the corpus, in which speakers mention the language resources where they found various autonymical lexical items. Table 32 provides a few illustrative examples of such language resources.

Dictionary referred to	Type of source
Vortaro de Esperanto (Kabe, 1911)	Monolingual general-language dictionary (print)
Gran Diccionario Español-Esperanto (Fernando de Diego) vortaro franca-Esperanto kaj Esperanto-franca (André Andrieu, 2000)	Multilingual general-language dictionary (print)
Dicionário Português-Esperanto (Túlio Flores) http://vortaro.brazilo.org/vtf/	Multilingual general-language dictionary (online)
Muzika Terminaro (Alfredo Aragon) 'La Sporta Lingvo en Esperanto' (Tibor Ujlaky-Nagy)	Monolingual specialized-language dictionary
Angla-Esperanta Medicina Terminaro de (Yamazoe Saburoo)	Multilingual specialized-language dictionary

Dictionary referred to	Type of source
ReVo Wikipedia ³⁴³ Lernu Vivo	Multilingual general-language dictionary (online)
vortaro-blogo http://vortaro-blogo.blogspot.com/2009/02/deja-vu.html	Dictionary blog
Nepivaj vortoj (André Cherpillod) ESPDIC (Paul Denisowski)	Lists and glossaries

Table 32. Examples of language resources in which speakers find lexical items.

Also in Chapter 6, some focus group members indicate that they use certain types of dictionaries as reservoirs of ideas (e.g., “One needs to approach these dictionaries with a particularly critical mind”). This attitude is also confirmed in the corpus results. Some speakers claim that language resources cannot be trusted:

Tamen, ni scias ke vortaraj difinoj estas malfidendaj, precipe en vortaroj Esperantaj. Aliflanke, mi ĉiam klopodas atenti tiujn difinojn, imagante ke ili povus esti la sola indikilo por mi, se mi ne konus aliajn fremdajn lingvojn, kaj estus ekz.. komencanto ĉina ... aŭ brita.

However, we know that definitions in dictionaries should not be trusted, especially those in Esperanto dictionaries. On the other hand, I try to always take these definitions into consideration, as I imagine that they could be the only indication—for instance, if I did not know any foreign languages, and I was a Chinese or British beginner.

However, many speakers use language resources because they see them as useful reservoirs of ideas:

La ofte kritikata Vikipedio tamen helpas, se taksu ĝiajn informojn prudente. Wikipedia, which is often criticized, still helps if you carefully consider the information that it contains.

343 Wikipedia is by definition not a dictionary, but it is often used as such by speakers on the look for language equivalents.

These results indicate that, just like PIV, these language resources are used as reservoirs of ideas rather than as prescriptive references.

8.3.5 Extralinguistic statements

8.3.5.1 *Need*

Quirion listed “response to a need” as a socioterminological factor. The corpus results confirm that speakers consider the filling of a lexical gap to be a strong argument in favor of a lexical item:

Lasu min diri la sekvan; min ĝenas, ke mi ĉiufoje devas ĉirkaŭskribi tiun “mobbing.” Eĉ, se tio laŭ gramatikaj pripensoj estas normala afero, mi sentas, ke ni ne havas vorton por fenomeno de la socio, kiu estas tre precize priskribebla. Do, estas la samo por mi, kvazaŭ ni ne havus la vorton “kafo” kiu estas same ĉiutaga kiel “mobbing.”

Let me tell you the following: I find it annoying that I always have to paraphrase the term “mobbing.” Even if paraphrasing is common from a grammatical point of view, I feel that I am lacking a word for a social phenomenon that can be described very precisely. For me, this is like not having a word for “coffee,” which is a daily concept just like “mobbing.”

8.3.5.2 *Context-specificity (variation)*

In the context-specificity category, I applied units for the observation of opinions on linguistic variation.³⁴⁴ Judgments linked to variations across time (*diachronic variation*), across social groups (*diastatic variation*), and across degrees of formality (*diaphasic variation*) all exist in the corpus.³⁴⁵ Table 33 provides examples.

344 Remysen also observed comments linked to linguistic variation in his corpus (2009).

345 Esperanto has no dialects (see Chapter 4), thus diatopic variation is not present in the results.

Context	Type of variation and judgement
<p>Cetere, mi dubas, ke la vorto “universale” estas konvena ĉi tie; por mi <u>ĝi aspektas kiel arkaika francaĵo</u>, kiu tradicie restas nur en la vortokombino “Universala Kongreso”—la moderna vorto estas “tutmonda,” “tutmonde.”</p> <p>Besides, I doubt that the word “universale” is appropriate here. To me, it looks like an outdated Gallicism based on the traditional word combination “Universala Kongreso”; the modern word is “tutmonda” or “tutmonde.”</p>	<p>Diachronic variation: Negative judgment because a lexical item is outdated</p>
<p>Tial fake kredeble “morfemsigno” estas pli trafa. Sed <u>por la komuna uzo “morfemsigno”</u> kredeble estas tro fakeca.</p> <p>Thus, in specialized contexts, “morfemsigno” is probably the most appropriate option. However, <u>in common use, “morfemsigno” is probably too specialized.</u></p>	<p>Diastratic variation: Negative judgment because a lexical item is too specialized</p>
<p>Mi pensas ke oni por esprimi la anglan “get high” diru “eŭforiĝi”; la adjektivo “high” do estas “eŭforiĝinta.”</p> <p><u>Eblas paroli iome pli neformale</u>, kaj diri “altumiĝi” / “estiĝi altuma.”</p> <p>ITo express the English phrase “get high,” you can say “eŭforiĝi”; thus, the adjective “high” is “eŭforiĝinta.”</p> <p><u>You can talk a bit less formally</u> by saying “altumiĝi” and “estiĝi altuma.”</p>	<p>Diaphasic variation: Proposal of distinct lexical items for formal and informal registers</p>

Table 33. Examples of statements about lexical items linked to different types of linguistic variation.

8.4 Summary

In this chapter, I revealed 23 categories of lexical criteria that Esperanto speakers can apply when accepting or rejecting a lexical item. I have done so using a corpus (a set of language naturally occurring data) to observe lexical criteria in context and to overcome the flaws of previous studies, which were based mostly on interviews and which thus suffered from the observer’s paradox (see Section 3.3.2).

Lexical criteria seem to have neither internal nor global consistency: Although speakers tend to possess their own ideologies, they sometimes use one criterion as

both an argument in favor of one lexical item and an argument against another item. The corpus results are also filled with conflicting ideologies and opinions—both interspeaker and intraspeaker. In other words, different speakers may use a criterion differently, and each speaker’s usage can vary according to the context (see the comment in Section 8.4.1.1).

In some ways, these lexical criteria raise more questions than they answer. How should interspeaker and intraspeaker inconsistencies be interpreted? Are the findings about the length of lexical items relevant for all of Esperanto? Is there an actual limit to the length of Esperanto lexical items (around 14 characters)? If so, can this be explained in terms of linguistic economy or cognitive overload?

In fact, each of the listed lexical criteria could be the subject of a dissertation in itself. However, I have argued throughout this thesis that language managers’ goal is applied: not to *know* but to *do*; thus, in the next chapter, I explain how language managers can leverage this data.

9 Using naturally occurring data to absorb uncertainty in deliberate lexical interventions

9.1 Introduction

In Chapter 3, I argued for the use of naturally occurring data, which refers to language data that is not produced from the explicit observations of researchers or language managers. In Section 3.2, I stated the nature of the problem: language managers must quickly make choices between mutually exclusive options (lexical items), even when they face structural uncertainty. A deliberate lexical intervention is an applied science problem. Language managers have the goal of reaching a certain degree of implantation for their target lexical item(s), so they seek to uncover the most efficient way to reach this goal. In other words, this goal is not to *know*, but to *do*. Therefore, the knowledge that language managers produce should help them efficiently achieve their goals. By analogy, successful marketers and entrepreneurs do not necessarily apply causal logic; rather, they adapt as they learn from the marketplace (see Section 3.2):

Strategies must change to leverage the specific requirements and behaviors of different groups along the diffusion curve. Product offerings may have to be adjusted over time and different adopter groups have to be told different stories about the benefits of the innovation. (Waarts, van Everdingen, & van Hillegersberg, 2002, p. 413)

Quo vadis language managers? The proposed solution is for language managers to continuously gather data from the speech community (following Auger, 1986, p. 52) by listening to the customers (i.e., speakers) so as to gather feedback and monitoring information directly from the marketplace (i.e., the speech community) in order to take prompt action if a deliberate lexical intervention seems to be ineffective. In Chapters 2 and 3, I mentioned that languages could be seen as complex adaptive systems. According to Eoyang and Berkas (1998), using such a system rather than a linear system involves changing the evaluator's role:

Complex adaptive dynamics do more than just require new tools and techniques for evaluation. They also transform the evaluator's role. Rather than being concerned with defining and measuring performance against specific outcomes, the evaluator takes on the task of designing and implementing

transforming feedback loops across the entire system. This role of transforming agent falls into two primary categories: absorbing uncertainty and making learning the primary outcome.

A metalinguistic statement with an opinionated autonym such as the ones obtained in the present investigation can act as a powerful tool for reducing uncertainty by gathering data from the speech community. By reducing uncertainty, here I mean increasing the confidence language managers may have in the probability of different outcomes for specific designational paradigms. In Chapter 8, I detailed how speakers' lexical criteria emerge from the corpus, allowing for a better comprehension of the acceptance or rejection of specific lexical items. However, other types of data also occurred around opinionated autonoms, which were related lexical source knowledge and lexical opinion (see definition of these concepts respectively in Sections 3.3.1 and 3.3.2).

The purpose of this chapter is twofold: explain the potential of the data analyzed in Chapter 8, but also of other types of data observed in the corpus of naturally occurring data. I introduce, one after another, types of data observed around opinionated autonoms and suggest how they could be used by language managers.

- Assess speakers' lexical-source knowledge (see Section 9.3.1)
- Understand speakers' lexical-source opinions (see Section 9.3.2)
- Assess speakers' lexical opinions (see Section 9.4)
- Understand speakers' lexical opinions (see Chapter 8)
- Assess speakers' lexical usage (via implantation studies and existing protocols)

In Section 9.6, I bring these considerations together into a model of action for language managers.

9.2 Assessing and understanding speakers' lexical knowledge

9.2.1 Data on lexical-source recall

The data on lexical-source recall indicate where each speaker found specific lexical items. This could be in a dictionary, as in the following example:

Kraüse (2007) havas “dentokrampo,” kaj tiu vorto troviĝas ankaŭ en Esperanta-hungara vortaro en la reto.

Kraüse (2007) has “dentokrampo,” and this word can be found in the online Esperanto–Hungarian dictionary, too.

However, it could also be from any of a host of other lexical sources, as shown in Section 8.4.2, including a book, as in this example:

La “Atlaso de la sunospektro” estas respektinda kaj saĝe prilaborita libro. Mi konsultis ĝin iomete antaŭ pluraj jaroj, sed nun mi ne havas ekzemplon. La fakto, ke la formo “suntorĉo” aperas en tiu verko, kaj la etimologia latina klarigo de M. Minnaert, estas sufiĉe por konvinki min.

The Atlaso de la Sunospektro [Atlas of the Solar Spectrum] is a respectable and wisely written book. I consulted it, to some extent, a few years ago, but now, I don’t have a copy. The fact that the form “suntorĉo” appears in this work, in addition to M. Minnaert’s etymological explanation, are sufficient to convince me.

The information found here tends to confirm the research-generated focus group results from Chapter 6. Here, however, the observations are covert. Such information is interesting for language managers because it corresponds to what I termed *lexical-source recall* in Section 3.3.1; scholars have not explored this topic, to my knowledge. This type of data indicates to language managers which lexical sources are present in speakers’ lexical environments and which (probably) are not. Language managers can then take action accordingly: For instance, if sources other than those that include the target lexical item keep appearing in the naturally occurring data, then the language managers know that they must better disseminate the lexical source that contains the target lexical item. This may sound like an obvious statement, but it is not: In Section 3.3.1, I showed that, at the level of lexical-source knowledge, scholars have repeatedly started from target lexical item sources instead of considering the extralexical situation in its entirety.

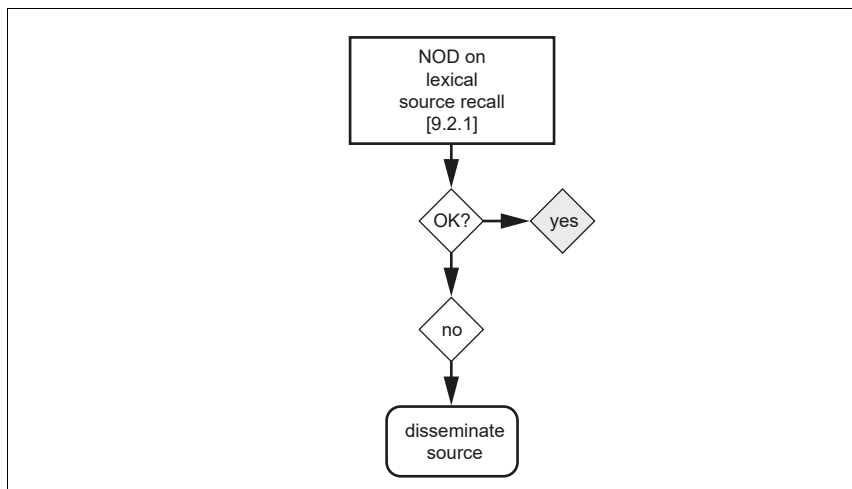


Figure 19. Using naturally occurring data (NOD) to assess lexical-source recall and better disseminate the lexical source containing target lexical item if the assessment is not satisfactory.

9.2.2 Data on lexical-source opinion

Speakers may mention not only the sources where they found specific lexical items but also the ways in which they evaluated those sources (i.e., lexical-source opinions). These opinions can be positive, as in the following instances:

La sporta lingvo en Esperanto (1972) de Tibor Újlaky-Nagy, verko laux mi tre rekomendinda, kaj desxutebla cxe www.fw.hu/eventoj/steb/sportalingvo.rtf uzas la esprimojn “egalstato” kaj “sendecida,” laux la oportuno de la koncerna kunteksto.

*I would highly recommend Tibor Újlaky-Nagy’s *La Sporta Lingvo en Esperanto (1972)* [a sports dictionary], which can be downloaded from www.fw.hu/eventoj/steb/sportalingvo.rtf. It uses the expressions “egalstato” and “sendecida,” depending on what’s appropriate in the given context.*

Ha, tiun vorton “kirko” konas ankaŭ la ReVo (fonto multege pli fidinda ol la PIV). Dankon pro la atentigo.

Ah, this word “kirko” [referring to a Christian church] is also present in ReVo (a source that is much more reliable than the PIV). Thanks for your remark.

However, these opinions can also be negative. Here, for instance, a speaker regrets that a lexical source has registered a lexical item that is almost never used, such that its meaning cannot be determined with certainty:

*“Vorko” estas unu el la multaj monstroj de PIV. La vorto estas tiom mal-
multe uzata, ke oni ne povas diri ion fidindan pri ĝia signifo en la reala
lingvo. Ĝi mankis ankoraŭ en PIV-1970 kaj enŝteliĝis en PIVon tra la PIV-
Suplemento (1987), kie ĝi ankoraŭ havis duan signifon “parto de fortikaĵo.”*
*“Vorko” is one of numerous monsters in PIV. This word is used so rarely
that you cannot say anything reliable about its meaning in real language
use. It was absent in the 1970 edition of PIV and intrusively entered the PIV
through the supplement (1987), where it still had a second meaning: “part
of a fortification.”*

Why is this type of information interesting for language managers? Well, this may again sound like a truism, but if speakers do not positively evaluate the lexical sources through which the language managers try to disseminate the target lexical item, then the speakers are unlikely to use them, and the target lexical item is unlikely to be adopted among the target speech community. If the perceived quality of a lexical source is high, however, then speakers likely will be loyal to this source (i.e., they are likely to use it again). Language managers can leverage this type of hint to start evaluating the perceived quality of the lexical sources that they will use in the dissemination of the target lexical item. If speakers have negative opinions of a lexical source, language managers can improve one of two aspects: the lexical source itself or the speakers’ perceptions of that source. In the above example, improving the lexical source could involve removing the lexical item “vorko” from the dictionary or adding a note to explain that it is rarely used within the speech community.

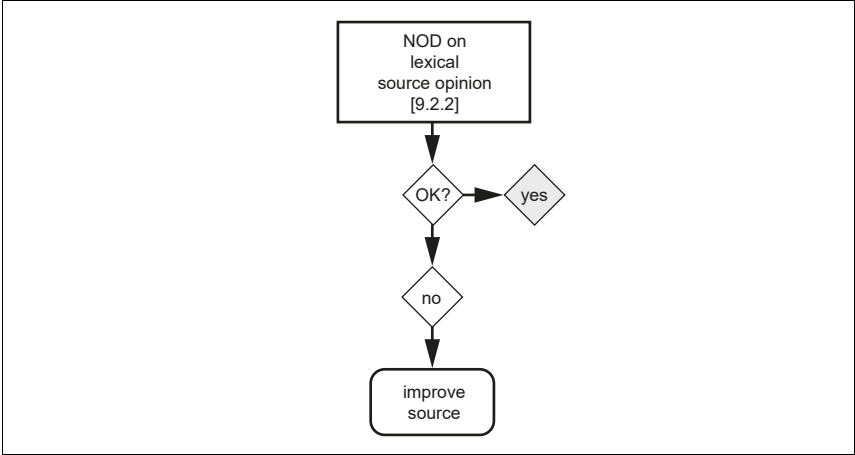


Figure 20. Using naturally occurring data (NOD) to understand lexical-source opinion and improve the perception of lexical sources containing target lexical item in the case of negative perception.

9.3 Assessing speakers’ lexical opinions

For language managers, it is important to know whether lexical items and their products are having positive or negative effects on the market. There are two possible approaches here: evaluating the opinions for each autonym and summarizing the opinions for an entire designational paradigm (e.g., comparing the target lexical item to competing lexical items).

9.3.1 Data on lexical item polarization at the lexical item level

An analysis of the contributions of opinionated autonoms is straightforward because the results of opinionated autonym extraction (see Chapter 7) can be easily filtered according to a specific autonym form.

Based on the data, I identified three categories for autonymical lexical items that induce

- only negative opinions,
- only positive opinions, or
- mixed opinions.

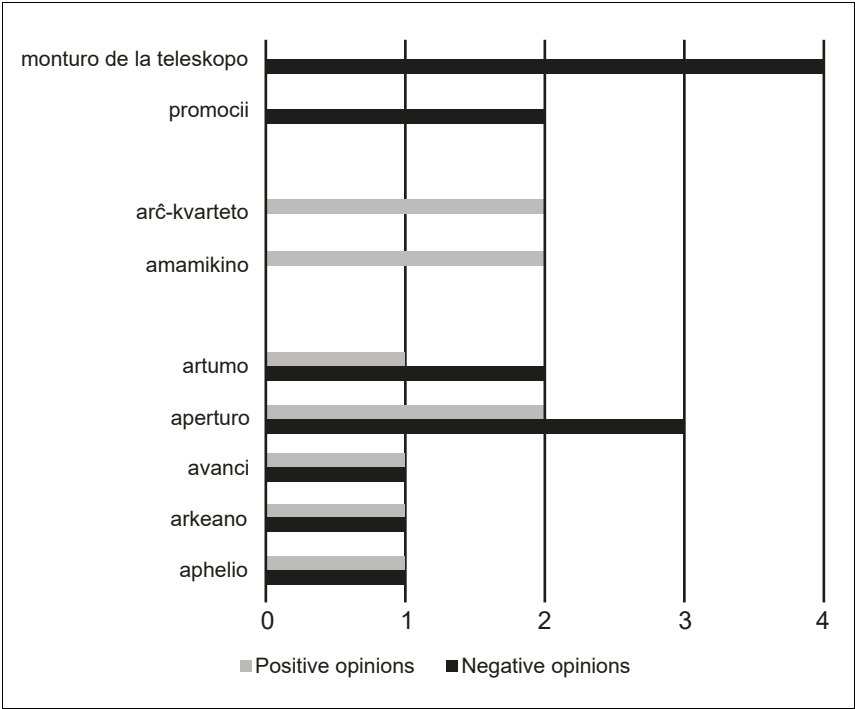


Figure 21. Sample opinionated autonymical lexical items grouped in three clusters (positive, negative, and mixed opinions).

Such clusters help language managers identify the trends for each lexical item. For instance, in the results, “monturo de la teleskopo” has four negative opinions and no positive opinions, so its implantation in the speech community will likely encounter barriers from speakers. By contrast, items such as “arĉ-kvarteto” and “amamikino” have predominantly favorable preliminary opinions. The method I used in my investigation (see Chapter 7) led to qualitative data, but language managers who are interested in quantitative data could adapt their methodology (e.g., corpus parameters) accordingly.³⁴⁶

.....
346 For a quantitative analysis with statistically significant results, further aspects should be taken into account, among which the frequency with which specific individuals post. Participation in electronic networks of practice is uneven, generally and also in the Esperanto speech community. Derks (2017), for instance, has distinguished five types of mailing list participants in the Esperanto speech community according to posting frequency: lurkers, single contributors, repeat messengers, prolifics and

I suggest the following: If target lexical items generally receive positive opinions, then language managers should take no actions because the goal is to ensure that lexical implantation works—not to understand why it works. On the contrary, language managers should act if (and only if) a target lexical item receives predominantly negative feedback. To identify the activities that they should undertake, language managers can seek to understand why speakers’ opinions are negative about each specific target lexical item (see Section 9.4) before taking action.

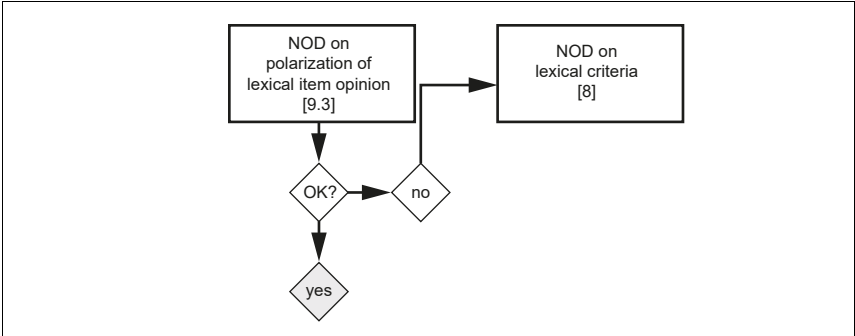


Figure 22. Using naturally occurring data (NOD) to assess lexical item polarization at the level of the lexical item and, if polarization is negative, move toward understanding speakers’ lexical opinion through lexical criteria (9.4).

9.3.2 Data on lexical item polarization at the designational paradigm level

Language managers can also choose to analyze an entire designational paradigm. This is much less straightforward (and much more time-consuming) than analyzing a single lexical item because the autonyms of a given designational paradigm have to be (manually) linked to a specific concept.

This approach is similar to Quirion’s implantation analysis for designational paradigms (2003a), but it occurs at an earlier stage, which I labeled as *lexical opinion* in Section 3.3.2. For example, in several opinionated autonym contexts, speakers look for an equivalent of “luminosity” (the amount of energy that a star emits). Four propositions occur in the results: “lumeco,” “lumpotenco,” “lumintenseco,”

.....
 specialists. Repeat messengers, for instance, could post similar contents several times in order to try and prove their points to other network members. For quantitative analyses, thus, each opinion or lexical criteria should be clearly associated with a specific individual (e.g. via an ID).

and “lumintenso.” The results indicate that the candidate with the greatest potential is “lumeco,” as a speaker already uses it and as it is short, clear, and relevant in the context of astronomy. The three other candidates all have negative comments. This does not guarantee that the candidate “lumeco” will win out in the long run, but it does show a clear indication that it is currently the preferred form.

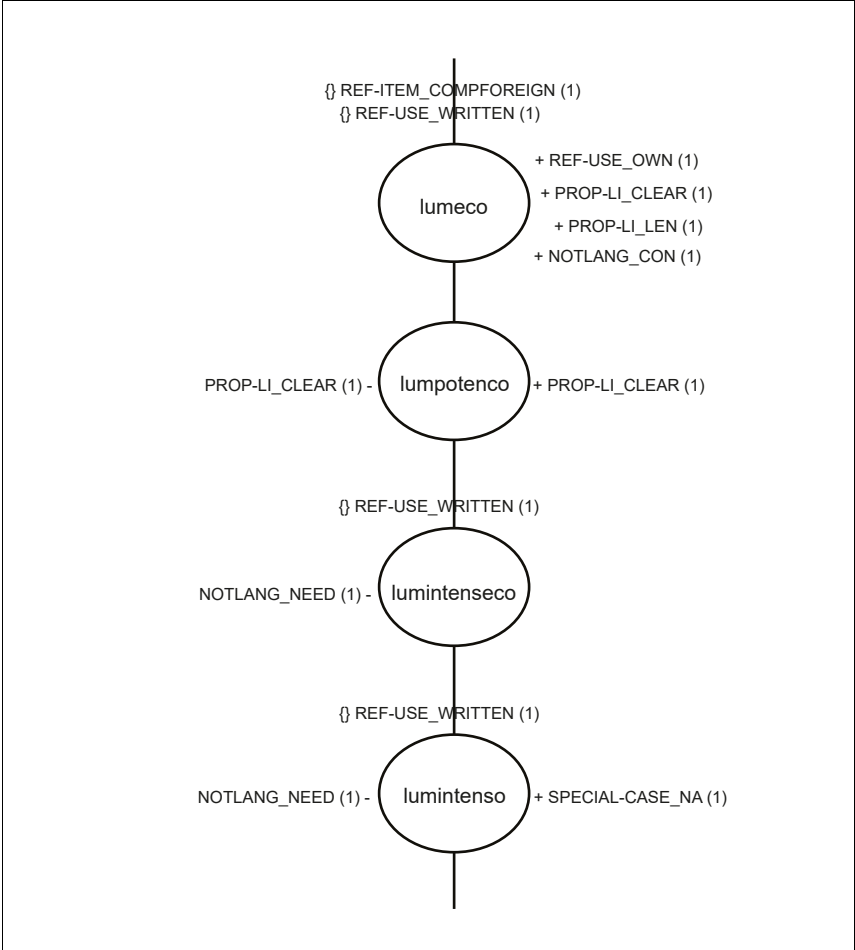


Figure 23. Comparison of statements made about four lexical item candidates competing as equivalents of “luminosity” (the amount of energy emitted by a star). “Lumeco” appears to be the best candidate in this set of results.

9.4 Understanding speakers' lexical opinions

When language managers realize (based on opinionated information) that a target lexical item is not achieving good results among the target speech community (see Section 9.3), this information can prove useful, as it can help them to conduct a more profound analysis before taking action. In Chapter 8, I presented the range of lexical criteria that speakers may apply to a lexical item. Consider a specific example: In this study's results, the lexical item "pedologio" receives four negatively polarized statements and one neutral statement:

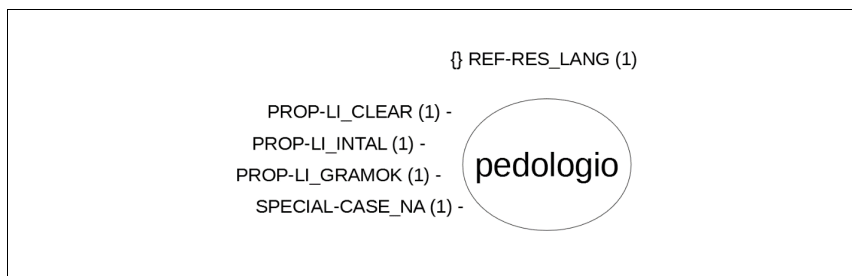


Figure 24. Summary of statements concerning the autonym “pedologio,” which is four negatively polarized opinionated statements and a neutral statement.

Here, language managers can zoom in on this lexical item and analyze the criteria that the speakers use when expressing negative opinions. In this context, speakers discuss the potential use of “pedologio” as an equivalent for “soil science.” Language managers who want to understand why the opinions about “pedologio” are rather negative in this context can retrieve the relevant excerpts from a database, as illustrated in Table 34:

pedologio		
{}	REF-REF, LANG	<p>Koncerne “pedologio”: la senco “studoj pri kreskanta homo ...” jam estas en ReVo—pri kio mi ne kulpas. la senco “grundoscienco” ne estas en ReVo, sed mi vidas en interreto ke ĝi estas uzata—pri kio mi ne kulpas.</p> <p>As to “pedologio,” the meaning “studies of growing humans” is already in ReVo—this is not my fault. The meaning “grundoscienco” [“soil science”] is not in ReVo, but I see that it is used on the Internet—this also is not my fault.</p>
–	PROP-LI, INTAL	<p>Tamen la okazo de “pedologio” estas ege malsimila. La vorto estas miskvalita strukture (ĉar ped- normale temas pri infanoj), ĝi ne estas inter-nacia, por ĝi ekzistas pli bona kaj jam bone establita termino.</p>
–	PROP-LI, GRAMOK	<p>However, the case of “pedologio” is quite different. The word is of poor quality from a structural point of view (because “ped-” is usually about children); it also is not international, and there is already a better, currently used term for it.</p>
–	PROP-LI, CLEAR	<p>“Pedologio” estas bazita sur alia pedo- (sed ankaŭ “pedofilio” implicas alispecan “-filio”n, alispecan amon). La rezulto tamen estas egala: la vorto misgvidas al tute alia sencokampo, kaj terminologie tio estas granda fuŝo.</p> <p>“Pedologio” is based on another meaning of “pedo-” (but “pedofilio” implies another type of “-filio,” or “love”). The result is the same, however: The word leads to a completely different semantic field; terminologically speaking, this is a big mess.</p>
–	SPECIAL- CASE_NA	<p>Mi tuj registris la vortojn “grundologio” kaj “pedologio” kiel evitindaj en: http://www.bonallingvo.it/index.php/Simplaj_samsignifaj_vortoj.</p> <p>I immediately registered the words “grundologio” and “pedologio” as not recommended, per http://www.bonallingvo.it/index.php/Simplaj_samsignifaj_vortoj.</p>

Table 34. Summary of statements about the opinionated lexical item “pedologio.”

In the results, the speakers’ criticisms of “pedologio” concern its lack of internationality and the grammatical problem posed by the prefix “ped-,” which is predominantly used to mean “relating to children.” This polysemy of the prefix is misleading, and one speaker mentions a resource that identifies the word as one

that should be avoided. Finally, another speaker mentions that the lexical item with the meaning “soil science” is not in the ReVo dictionary but can be found on the Internet.

With this analytical tool, language managers can understand why a target speech community might reject a specific lexical item. They can then review the target lexical item accordingly. For a revision, several actions can be envisaged, including replacing the target lexical item with another one that better corresponds to speakers’ expectations and working to change speakers’ opinions about the target lexical item. Lexical opinions can be learned, just as attitudes can be:

Although constitutional and physiological factors have to be kept in mind (for instance extroversion seems to be hereditary to some extent), attitudes are learnt, which is why parents and education become influential factors in this respect, their influence being such that the attitudes originated in these social milieus happen to be particularly resistant. Other socializing factors to consider are friends, peers and the mass media, especially television nowadays. (Lasagabaster, 2008, p. 400)

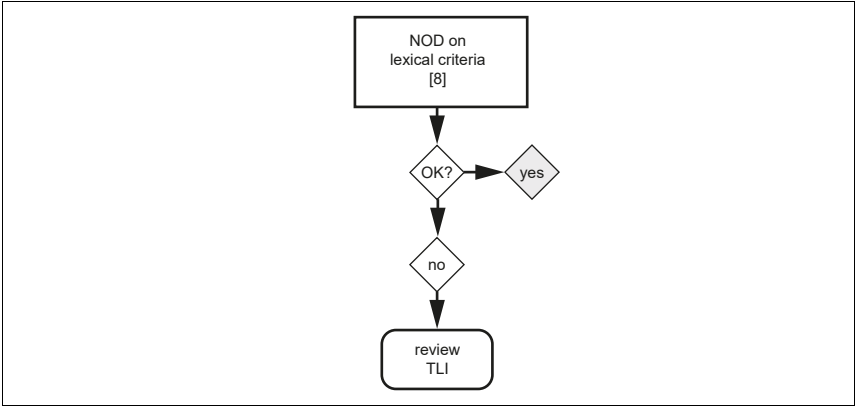


Figure 25. Using naturally occurring data (NOD) to decide whether a given target lexical item needs revision.

9.5 A model for action based on naturally occurring data

Regarding these five types of naturally occurring data that language managers can use during the lexical-implantation process: Where should they start?

As I explained in Chapter 3 (Section 3.2), several scholars have put forth the idea that language should be considered a complex, adaptive system, which implies that it is probably impossible to predict the details of lexical changes. I further explained that most deliberate lexical interventions are equivalent to a supply-side market, wherein the key to effectiveness is not planning but rather having the capability to both monitor the process of gathering information from the marketplace and provide quick reactions to unforeseen changes. Here, I repeat what I suggested in Chapter 3: language managers should change their tactics as a target lexical item progresses from aim to achievement. To this end, they should, following the analogy to marketers, process market information and recursively make data-based decisions.

One of the points that I wish to stress is that assessment should be prioritized over understanding, as understanding is not essential when speakers have positive assessments. The goal of language managers is to do, not to know, and language managers must act quickly. Thus, assessment and understanding should be clearly separated. The process of understanding can and should be skipped if an assessment provides satisfactory results. Assessments should, however, take place on a regular basis, in the form of a monitoring:

Monitoring provides the feedback loop for learning about the system; learning is sought not for its own sake but primarily to better achieve management objectives. In this case, monitoring should be designed to reduce the critical uncertainties in models of the managed system. (Lyons, Runge, Laskowski, & Kendall, 2008, p. 1683)

language managers must acknowledge uncertainty to address it, so they must process data from the speech community during the lexical-implantation process. I cannot stress enough that scholars and language managers who are pursuing the goal of implanting target lexical items should concentrate less on clarifying the details of every implantation factor or on trying to predict the results of lexical interventions, and more on reacting to changes in the target speech community and on monitoring the linguistic situation (especially at the beginning of the lexical intervention).

Broad features can become knowable, but as speakers explained in the corpus, a person can use the same lexical criterion differently depending on the situation. Here, I quote this prime example again:

La vorto “internacia” estas unu el la plej problemaj. Ne nur, ke kvin personoj havas kvin malsamajn opiniojn pri ĝia signifo, montriĝas eĉ, ke la sama persono komprenas ĝin malsame ĉe malsamaj problemoj.

The word “internacia” [international] is one of the most problematic ones. The problem is not only that five people have five opinions about what it means; it also turns out that even one person can have different understandings of the word in different contexts.

This is a strong argument for monitoring each specific target lexical item case instead of engaging in a general search for knowledge

Language managers should gather information from the speech community not so that they can know but so that they can do. They should do so throughout the lexical-implantation process to adapt to the language’s complex, adaptive system whenever needed. Figure 26 brings together the data types presented above and illustrates how language managers can leverage naturally occurring data.

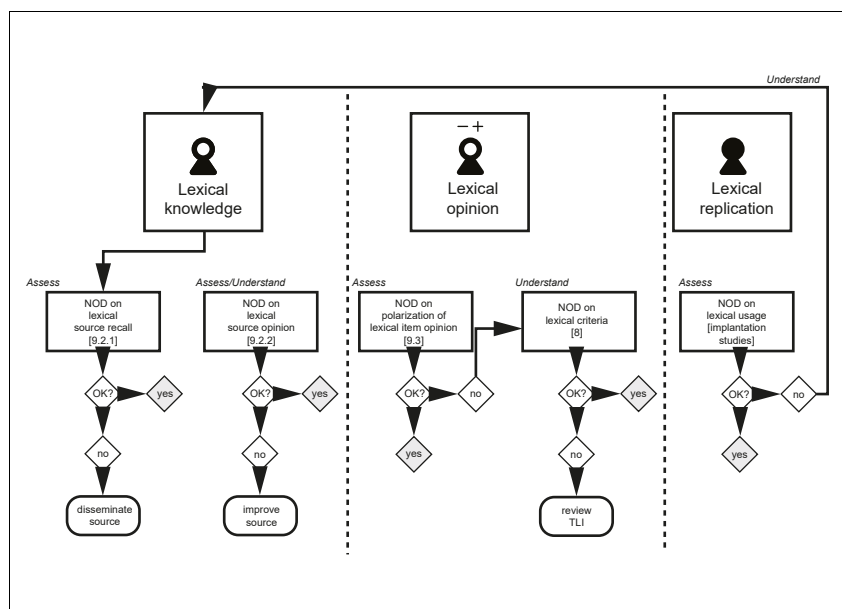


Figure 26. Model for action for language managers based on naturally occurring data (NOD). A “yes” means that no further activities should be undertaken.

Conclusion

10 Conclusion and future research

10.1 Objectives and methods

In this investigation, I proposed language managers to explore, in context, speakers' lexical environments and lexical opinions using naturally occurring data. This proposal was necessary, for although language managers have long tried to conduct deliberate lexical interventions with the aim of changing speakers' lexical usages, some have had relatively little success.³⁴⁷ The main contribution of this work is the conclusion that the success of a deliberate lexical intervention cannot necessarily be known in advance; therefore, I propose a shift in research focus: language managers should aim to *do*, not to *know*. Accordingly, language managers should find a way to make decisions quickly when addressing structural uncertainty.

In this investigation, I also proposed to prevent uncertainty by monitoring data from the speech community on both lexical environments (mainly so as to assess the sources of lexical information that speakers use) and lexical opinions (so as to assess, on a case-by-case basis, the criteria that speakers use when evaluating and choosing lexical items).

In this investigation, I used both a focus group study (so as to better understand speakers' lexical environments, including extralexical situations) and naturally occurring data (so as to gain a broader understanding of speakers' lexical criteria in context—thus avoiding the observer's paradox).

10.2 Summary of main findings and achievements

Deliberate lexical interventions

Based on existing research, I conceptualized the notion of deliberate lexical interventions and listed the problems that language managers are facing when conducting deliberate lexical interventions; for instance, members of the target speech community often do not know the target lexical items, or know them but do not use them. I also argued that language should be considered a CAS. One conse-

³⁴⁷ Needless to say this depends on how success is defined (see discussion under 2.5.1, p. 70). However, cases in which the dissemination of target lexical items is zero (see example e.g. in Gresa Barbero, 2016) can, in my view, be called failures.

quence of this viewpoint is the impossibility of foreseeing lexical change in great detail (which leads to the probable impossibility of developing a predictive model). I proposed a solution to this problem: minimizing structural uncertainty (i.e., finding strategies that are suitable given partial ignorance of the future and that are feasible for use with strict time constraints). Uncertainty can also be reduced by exploring speakers' lexical environments and opinions in context.

Speakers' lexical environments

Language managers struggle with low levels of lexical knowledge. I thus examined how speakers learn new lexical material and made the following observations:

- External searches for information are not systematic and may be replaced with various other strategies, including lexical creation.
- Each speaker has unique modes of accessing lexical sources, and these modes can vary from speaker to speaker.
- Lexical sources can be traditional (e.g., dictionaries) or alternative (e.g., collaborative dictionaries, Wikipedia, or search engines) and can also include fellow speakers.
- Some speakers are not selective when choosing lexical sources, as long as each source meets their lexical needs.
- Some speakers use lexical sources as reservoirs of ideas and consciously choose whether to use or ignore their contents on a case-by-case basis.

The detection of speakers' lexical opinions in context

This investigation is intended to allow language managers to reduce uncertainty within strict time constraints and to eliminate the observer's paradox that weakened many previous studies. To this end, I developed a proof of concept to monitor speakers' lexical opinions, as found in corpora; I combined lexical, syntactic, and paralinguistic peculiarities that pointed toward opinion and autonomy, using natural language processing to identify contexts containing opinions and autonyms with relatively good precision (above 0.8 for detecting opinionated autonym candidates in the test set). This proof of concept suggests that it is possible to use natural language processing to both systematically monitor speakers' opinions of lexical items and further automate the application of natural language processing. This method is innovative because it overcomes the observer's paradox.

Speakers' lexical opinions in context

Lexical opinions are present throughout the corpus, with 23 categories of lexical criteria revealed according to the naturally occurring data. Speakers use these lexical criteria to accept or reject lexical items. Interspeaker consistency is not the rule, however, as speakers have conflicting ideologies and criteria. In addition, intraspeaker inconsistencies exist. Thus, I conducted a case-by-case observation and monitored the level of the lexical items during the implantation process.

10.3 Future research

The focus group study and the observation of the contexts containing opinionated autonoms reveal the presence of data types other than lexical criteria, including

- data on lexical-source recall, regarding the sources that speakers consult to find lexical items;
- lexical-source opinions, regarding what speakers think of the lexical sources; and
- lexical-item polarization, regarding the speakers' opinions about a specific lexical item (or set of items) in a designational paradigm.

Language managers can and should use such data to address uncertainty. Those who wish to undertake deliberate lexical interventions should explicitly acknowledge the uncertainty and should design an approach to address it. This approach can, as shown in this investigation, involve collecting data from the target speech community throughout the lexical-implantation process to enable quick reactions to changes, such as adaptations in strategies and deliberate lexical intervention activities. This is needed because language is a complex, adaptive system.

I propose using two main perspectives when continuing this research: first, improving this study's methods to more quickly collect information from the speech community throughout the lexical-implantation process, and second, collaborating directly with members (speakers) of the target speech community.

Improving the proof of concept

The proof of concept could be improved by:

- conducting statistically relevant tests to extract opinionated autonym candidates and testing the application on texts with main topics other than language;
- improving the precision of indicators (e.g., by eliminating noise through filtering techniques);
- adding indicators that are not lexical (e.g., syntactic clues) when detecting opinions;
- automatically detecting Esperanto opinions' orientations and targets; and
- adapting this study's tool for use with other languages.

Collaboration with members (speakers) of the target speech community

One future area of research involves investigating lexical cocreation (collaborative or semicollaborative lexicography)—in other words, determining how best to put speakers to work. This is the most challenging goal for any language manager, which is why I did not start with it. Because this is a very ambitious goal and because new proposals are always being made in the Esperanto speech community (the most recent one, to my knowledge, is Cramer, 2018), researchers must remember that neither collaboration nor consultation is guaranteed. The Esperanto speech community knows this well. This challenge may be even greater in the Web era than before because, when “online, people can switch behaviors as soon as they see something better” (Li & Bernoff, 2011, p. 12).

References

- Abel, A., & Meyer, C. M.** (2013). The dynamics outside the paper: user contributions to online dictionaries. *Proceedings of ELex 2013*, 179–194. Tallinn, Estonia.
- Abel, A., & Meyer, C. M.** (2016). Nutzerbeteiligung. In A. Klosa & C. Müller-Spitzer (Ed.), *Internetlexikografie. Ein Kompendium* (pp. 249–290).
- Achard-Bayle, G., & Lecolle, M.** (Ed.). (2009). *Sentiment linguistique. Discours spontanés sur le lexique*. Metz: Université Paul Verlaine.
- Achard-Bayle, G., & Paveau, M.-A.** (2008). Présentation. La linguistique «hors du temple». *Pratiques*, 139–140.
- Adegbija, E.** (2014). Marketing new lexical terminology in Nigeria: Some practical considerations. In M. Pütz (Ed.), *Discrimination through Language in Africa?, Perspectives on the Namibian Experience* (Reprint 2013, pp. 101–122). <https://doi.org/10.1515/9783110906677>
- Akrich, M.** (2012). Les listes de discussion comme communautés en ligne : outils de description et méthodes d'analyse. *Papiers de recherche du CSI*, numéro 25. Retrieved from http://www.csi.ensmp.fr/working-papers/WP/WP_CSI_025.pdf
- Allony-Fainberg, Y.** (1974). Official Hebrew terms for parts of the car: a study of knowledge, usage and attitudes. In J. A. Fishman (Ed.), *Advances in Language Planning* (pp. 493–517). The Hague: De Gruyter Mouton.
- Amey, P.** (1996). Polyphonie et scénographie des discours sur l'énergie nucléaire. La mise en scène de l'énonciation dans les discours experts et profanes (Mémoire de diplôme en sociologie). Université de Genève, Genève.
- Ammon, U., Dittmar, N., Mattheier, K. J., & Trudgill, P.** (2008). Sprache und Massenkommunikation. In *Sociolinguistics – An International Handbook of the Science of Language and Society* (2nd compl. rev. and extend. ed., Vol. 1, pp. 1615–1628). Retrieved from <https://www.degruyter.com/viewbooktoc/product/178201>
- Antia, B. E.** (2000). Terminology and language planning: an alternative framework of practice and discourse. Amsterdam [etc.]: J. Benjamins.

- Antos, G.** (1996). Laien-Linguistik. Studien zu Sprach- und Kommunikationsproblemen im Alltag. Am Beispiel von Sprachratgebern und Kommunikationstraining. Tübingen: Max Niemeyer Verlag.
- Arendt, B.** (2010). Niederdeutschdiskurse : Spracheinstellungen im Kontext von Laien, Printmedien und Politik. Berlin: Schmidt, Erich.
- Arendt, B., & Kiesendahl, J.** (2015). Sprachkritische Kommentare in der Forenkommunikation – Form, Funktion und Wirkung. In J. Bücker, E. Diedrichsen, & C. Spieß (Ed.), *Perspektiven linguistischer Sprachkritik* (pp. 159–198). Stuttgart: Ibidem-Verl.
- Arndt, T., & Wöllbrink, B.** (2013). *Der Terminologiekreis. Unternehmensweite Abstimmungsprozesse planen und leiten*. Präsentation an der tekom-Herbsttagung, Wiesbaden. Retrieved from http://tagungen.tekom.de/fileadmin/tx_doccon/slides/413_Der_Terminologiekreis_Unternehmensweite_Abstimmungsprozesse_planen_und_leiten.pdf
- Astronomia Esperanto-Klubo (AEKO).** (2000, 2012). Astronomia Terminaro. Accessed on 2016/12/18: <https://groups.yahoo.com/neo/groups/aeko-at/info>
- Astronomia Esperanto-Klubo (AEKO), & Rocha-Pinto, H. J.** (n. d.). Astronomia Terminaro 1.5. Accessed on 2016/12/8: <http://www.eventoj.hu/steb/vortaroj/astronomia-terminaro.htm>
- Auger, P.** (1986). Francisation et terminologie : l'aménagement terminologique. In G. Rondeau & J. C. Sager (Ed.), *Termia 84: terminologie et coopération internationale: la terminologie, outil indispensable au transfert des technologies: colloque international de terminologie, Luxembourg, 27–29 août 1984 = Terminology and international cooperation: the terminology dimension of the transfer of technology: [...]* (pp. 47–55). Québec: Girsterm.
- Authier-Revuz, J.** (2003). Le Fait autonymique : Langage, langue, discours. Quelques repères. In J. Authier-Revuz, M. Doury, & S. Reboul-Touré (Ed.), *Parler des mots. Le fait autonymique en discours* (pp. 67–96). Paris: Presses Sorbonne Nouvelle.
- Back, O.** (1996). Plansprachen. In Hans Goebel, Peter H. Nelde, Zdeněk Starý, & Wolfgang Wölck (Ed.), *Kontaktlinguistik: ein internationales Handbuch zeitgenössischer Forschung = Contact linguistics: an international handbook of contemporary research = Linguistique de contact: manuel international des recherches contemporaines* (pp. 881–887). Berlin: Walter de Gruyter.
- Badir, S.** (2014). Système à tous les étages. *Signata*, 4 [*Que peut le métalangage?*], 155–175.

- Baker, C.** (1992). *Attitudes and Language*. Clevedon, Philadelphia [sic], Adelaide: Multilingual Matters.
- Baker, M. J.** (2014). *Marketing Strategy & Management* (Fifth Edition). London: Palgrave.
- Balahur Dobrescu, A.** (2011). Methods and Resources for Sentiment Analysis in Multilingual Documents of Different Text Types. Universitat d'Alacant.
- Balahur Dobrescu, A., Mihalcea, R., & Montoyo, A.** (2014). Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications. *Computer Speech and Language*, 28, 1–6.
- Ballarin, S.** (2009). *Approche sociolinguistique de la production terminologique au Québec. Application du modèle glottopolitique* (Thèse de doctorat en traduction, Université de Montréal, Faculté des arts et des sciences, Département de linguistique et de traduction). Retrieved from <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/6692>
- Bartsch, R.** (1987). *Sprachnormen : Theorie und Praxis* (Studienausg.). Tübingen: Niemeyer.
- Beaulieu-Marianni, M.-P.** (2012). Conceptualisation, représentations et processus de production dans l'apprentissage de l'anglais : quels enjeux pour le lexique ? (Thèse de doctorat, École doctorale Lettre, langues et spectacles). Université Paris-Ouest Nanterre La Défense.
- Beckert, J.** (1996). What is sociological about economic sociology? Uncertainty and the embeddedness of economic action. *Theory and Society*, (25), 803–840.
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Ellis, N. C., Holland, J., ... Schoenemann, T.** (2009). Language Is a Complex Adaptive System. Position Paper. *Language Learning*, 59(Suppl. 1), 1–26.
- Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C.** (2003). A Neural Probabilistic Language Model. *Journal of Machine Learning Research*, (3), 1137–1155.
- Berg, B. L.** (2011). *An Introduction to Content Analysis* (8th ed.). Retrieved from <http://depts.washington.edu/uwmcnair/chapter11.content.analysis.pdf>
- Bergenholtz, H., Bothma, T. J. D., & Gouws, R. H.** (2015). Phases and Steps in the Access to Data in Information Tools. *Lexikos*, (25), 1–30.
- Bergenholtz, H., & Gouws, R.** (2010). A New Perspective on the Access Process. *Hermes – Journal of Language and Communication Studies*, (44).
- Bergenholtz, H., Nielsen, S., & Tarp, S.** (Ed.). (2009). *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang.

- Berry, R.** (2009). Making the Most of Metalanguage. *Language Awareness*, 14(1), 3–20.
- Betz, W.** (1960). Sprachlenkung und Sprachentwicklung. In Sprache und Wissenschaft: Vorträge gehalten auf der Tagung der Joachim Jungius-Gesellschaft der Wissenschaften, Hamburg, am 29. und 30. Oktober 1959 (pp. 85–100). Göttingen: Vandenhoeck & Ruprecht.
- Beuge, P.** (2014). Was ist gutes Deutsch aus Sicht Linguistischer Laien? *Zeitschrift für Dialektologie und Linguistik*, 81(2), 129–150.
- Bhatia, T. K., & Ritchie, W. C.** (Ed.). (2013). *The Handbook of Bilingualism and Multilingualism*. Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.
- Bhreathnach, Ú.** (2011). *A Best-Practice Model for Term Planning* (PhD in the Faculty of Humanities and Social Sciences, Dublin City University). Retrieved from http://doras.dcu.ie/16548/1/UBhreathnach_Thesis_Vol_1.pdf
- Bibeau, G.** (1983). La normalisation terminologique et le néo-purisme au Québec. *Québec français*, 49, 14–17.
- Biber, D.** (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8(4), 243–257.
- Biber, D., Conrad, S., & Reppen, R.** (1998a). *Corpus Linguistics* (2004 (fourth printing)).
- Biber, D., Conrad, S., & Reppen, R.** (1998b). *Corpus linguistics. Investigating language structure and use*. Cambridge: Cambridge University Press.
- Bick, E.** (2016). A Morphological Lexicon of Esperanto with Morpheme Frequencies. In *LREC2016, Tenth International Conference on Language Resources and Evaluation*, 1075–1078. Portorož.
- Bickart, B., & Schindler, R. M.** (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, 15(3), 31–40.
- Blanke, D.** (1981). Plansprache und Nationalsprache – Einige Probleme der Wortbildung des Esperanto und des Deutschen in konfrontativer Darstellung (Dissertation A). Humboldt-Universität zu Berlin.
- Blanke, D.** (1985). *Internationale Plansprachen. Eine Einführung*. Berlin: Akademie-Verlag.
- Blanke, D.** (1998a). Interlinguistik und Plansprachen. *Sitzungsberichte der Leibniz-Sozietät*, 21(2), 45–76.
- Blanke, D.** (1998b). *Pri la aktualata stato de interlingvistiko* (T. Carlevaro, Ed.). Bellinzona: Hans Dubois.

- Blanke, D.** (2006). Interlinguistische Beiträge. Zum Wesen und zur Funktion internationaler Plansprachen. (S. Fiedler, Ed.). Frankfurt am Main: Peter Lang.
- Blanke, D.** (2009). Causes of the relative success of Esperanto. *Language Problems & Language Planning*, 33(3), 251–266.
- Blanke, D.** (2010). Sprachwandel im Esperanto – gezeigt an Beispielen aus der Lexik. In K. Reinke & C. Sinner (Ed.), *Sprache als Spiegel der Gesellschaft* (pp. 51–77). München: Peniope, 2010.
- Blanke, D., & Blanke, W.** (2012). Fachsprachliche Kommunikation in Esperanto. Fachkommunikation – interlinguistische Aspekte. Beiträge der 21. Jahrestagung der Gesellschaft für Interlinguistik e.V., November 2011 in Berlin (Interlinguistische Informationen, Beiheft 19). Vortrag an der Tagung Fachkommunikation – interlinguistische Aspekte, Berlin.
- Blanke, D., & Blanke, W.** (2015). Is scholarly communication possible in a so-called « artificial » language? *Interdisciplinary Description of Complex System*, 13(2), 216–235.
- Blanke, W.** (1997). Terminologiaj principoj kaj ilia normigo – kelkaj aspektoj. In I. Koutny & M. Kovács (Ed.), *Struktura kaj socilingvistika esploro de Esperanto* (pp. 117–125). Budapest: Steleto & ILEI.
- Blanke, W.** (2008a). Esperanto – Terminologie und Terminologearbeit. New York: Mondial.
- Blanke, W.** (2008b). Terminologia Esperanto-Centro. Efforts for Terminological Standardization in the Planned Language. In *Esperanto – Terminologie und Terminologearbeit* (pp. 48–61). New York: Mondial.
- Blanke, W.** (2013). Pri terminologia laboro en Esperanto. Elektitaj publikajoj (D. Blanke, Ed.). New York: Mondial.
- Blythe, R. A., & Croft, W.** (2009). The Speech Community in Evolutionary Language Dynamics. In N. C. Ellis & D. Larsen-Freeman (Ed.), *Language as a Complex Adaptive System* (pp. 47–63). Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.
- Bock, K., & Levelt, W.** (1994). Language production: Grammatical encoding. In *Handbook of Psycholinguistics* (pp. 945–984). San Diego: Academic Press.
- Boeije, H.** (2010). *Analysis in qualitative research* (1st ed.). London, Thousand Oaks, New Delhi, Singapore: SAGE Publications.
- Bonnin, J. E.** (2012). *Decentralization of the linguistic norm online. The Royal Spanish Academy challenged on the Internet* (N° 36). Retrieved from https://www.tilburguniversity.edu/upload/15770a63-e1fc-41be-ab65-e027acc7bc00_tpcs%20paper36.pdf

- Bonnin, J. E.** (2014). Pensar el castellano en Internet: discursos sobre la norma en los foros de WordReference.com. In E. Narvaja de Arnoux & S. Nothstein (Ed.), *Temas de glotopolítica. Integración regional sudamericana y panhispanismo* (pp. 351–372). Retrieved from https://www.academia.edu/17054397/Pensar_el_castellano_en_Internet_discursos_sobre_la_norma_en_los_foros_de_WordReference.com
- Bormann, W.** (1999). Novaj defioj antaŭ la Akademio de Esperanto. In D. Blanke, R. McCoy, & O. Buller (Ed.), *Por aktiva lingvopolitiko. Aktoj de la Lingvopolitika Seminario en la 81-a Universala Kongreso de Esperanto, Prago 1996* (pp. 35–42). Rotterdam: Universala Esperanto-Asocio.
- Bosredon, B., & Tamba, I.** (1998). L'autonymie linguistique. Un exemple de transfert théorique. *Sémiotiques*, (14), 171–186.
- Bosson, G.** (2000). Écrire dans une langue régionale. L'expérience piémontaise vue de l'extérieur. In S. Gilardino & B. Villata (Ed.), *II° Convegno internazionale sulla lingua e la letteratura del Piemonte – Vercelli 2000* (pp. 159–193). Vercelli: VercelliViva.
- Boulangier, J.-C.** (1984). Quelques observations sur l'innovation lexicale spontanée et l'innovation lexicale planifiée. *La Banque des mots*, 27, 3–29.
- Boulogne Declaration.** (2014, June 5). Retrieved from aktuale.info website: <https://web.archive.org/web/20140506075349/http://aktuale.info/en/biblioteko/dokumentoj/1905>
- Bourdieu, P.** (1977a). L'économie des échanges linguistiques. *Langue française*, 34, 17–34.
- Bourdieu, P.** (1977b). The economics of linguistic exchanges (R. Nice, Transl.). *Social Science Information*, 16(6), 645–668.
- Bourdieu, P.** (1982). Ce que parler veut dire. L'économie des échanges linguistiques. Paris: Fayard.
- Bovaird, T.** (2008). Emergent Strategic Management and Planning Mechanisms in Complex Adaptive Systems. *Public Management Review*, 10(3), 319–340.
- Brinton, L. J., & Traugott, E. C.** (2005). *Lexicalization and Language Change*. Cambridge University Press.
- Brosch, C.** (2008). Vortfarado el la vidpunktoj de la hindeŭropista lingvoscienco kaj de esperantologio (Słowotwórstwo z punktu widzenia indeoeuropeistyki i esperantologii) (Universitata diplomlaboraĵo). Uniwersytet im. Adama Mickiewicza w Poznaniu, Poznań.
- Brosch, C.** (2009). Komposition und Derivation in altindogermanischen Sprachen und im Esperanto. In S. Fiedler (Ed.), *Esperanto und andere Sprachen im Vergleich* (pp. 83–97). Berlin: Gesellschaft für Interlinguistik.

- Cabré, M. Teresa.** (2010). La implantación de la terminología: conceptos, estrategias, recursos y estudios realizados sobre el francés de quebec y el catalán. In X. Dins Alberdi & P. Salaburu (Ed.), *Ugarteburu terminologia jardunaldiak* (pp. 2–18). Bilbao: Euskal Herriko Unibertsitatea (Universidad del País Vasco).
- Cabré, María Teresa.** (2002). La neologia, avui: el naixement d'una disciplina. In María Teresa Cabré, J. Freixa, & E. Solé (Ed.), *Lèxic i neologia* (pp. 29–40). Barcelona: Observatori de Neologia, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, María Teresa.** (2003). Theories of terminology. Their description, prescription and explanation. *Terminology*, 9(2), 163–200.
- Camacho, J.** (1999). La Mava Lingvo: Neologismoj kaj Malneologismoj en Esperanto. In V. Benczik (Ed.), *Lingva arto. Jubilea libro omaĝe al William Auld kaj Marjorie Boulton*. Rotterdam: UEA.
- Cameron, D.** (1995). *Verbal hygiene*. London ; New York: Routledge.
- Candel, D.** (2005). La néologie d'aménagement : l'expérience française. In L. Depecker & V. Dubois (Ed.), *Les néologies contemporaines* (pp. 55–77). Paris: Société française de terminologie.
- Canut, C.** (1998). Pour une analyse des productions épilinguistiques. *Cahiers de praxématique*, (31), 69–90.
- Canut, C.** (2000). Subjectivité, imaginaires et fantasmes des langues : la mise en discours « épilinguistique ». *Langage et société*, 2000/3(93), 71–97.
- Carnap, R.** (1934). *Logische Syntax der Sprache*. Wien: Springer.
- Cartoni, B.** (2008). *De l'incomplétude lexicale en traduction automatique : vers une approche morphosémantique multilingue* (Université de Genève). Retrieved from <http://archive-ouverte.unige.ch/downloader/vital/pdf/tmp/59qgb8g8elchghne03s3jfo5h5/out.pdf>
- Casper, K.** (2002). *Spracheinstellungen : Theorie und Messung*. Heidelberg: K. Casper.
- Castillo, C., Donato, D., Gionis, A., Murdock, V., & Silvestri, F.** (2007). Know your neighbors: web spam detection using the web topology. *SIGIR '07 Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 423–430. Retrieved from <https://dl.acm.org/citation.cfm?id=1277814>
- Chambliss, D. F., & Schutt, R. K.** (2013). *Making Sense of the Social World. Methods of Investigation* (4th ed.). Los Angeles, London, New Delhi, Singapore, Washington DC: SAGE Publications.

- Chansou, M.** (1993). Étude d'implantation des arrêtés de terminologie. Domaines : audiovisuel et publicité. In L. Depecker (Ed.), *La mesure des mots. Cinq études d'implantation terminologique*. (pp. 133–233). Rouen: Publications de l'Université de Rouen.
- Chansou, M.** (2003). L'aménagement lexical en France pendant la période contemporaine : (1950–1994), étude de sociolexicologie. Paris: Champion.
- Charron, C., Favier, J., & Li, C.** (2006). *Social Computing. How Networks Erode Institutional Power, And What to Do About It*. Retrieved from <http://forr.com/gsw1-18>
- Cherpillod, A.** (1988). *Nepivaj vortoj* (2a eldono). Courgenard.
- Choi, T. Y., Dooley, K. J., & Rungtusanatham, M.** (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management*, 19(3), 351–366.
- Chrdle, P.** (2013). *Profesia uzo de Esperanto kaj ĝiaj specifaj trajtoj* (3rd ed.). Dobřichovice: KAVA-PECH.
- Christensen, K. S.** (1985). Coping with Uncertainty in Planning. *Journal of the American Planning Association*, 51(1), 63–73.
- Clark, L. A., & Watson, D.** (1995). Constructing Validity: Basic Issues in Objective Scale Development. *Psychological Assessment*, 7(3), 309–319.
- Cobarrubias, J.** (1983). Ethical Issues in Status Planning. In J. Cobarrubias & J. A. Fishman (Eds.), *Progress in Language Planning: International Perspectives* (pp. 41–86). Berlin, New York, Amsterdam: Mouton Publishers.
- Cole, C. A., Gaeth, G., & Singh, S. N.** (1986). Measuring prior knowledge. *NA – Advances in Consumer Research*, 13, 64–66.
- Coluzzi, P.** (2007). Minority Language Planning and Micronationalism in Italy. An analysis of the situation fo Friulian, Cimbrian and Western Lombard with reference to Spanish Minority Languages. Oxford, Bern, Berlin, Brussels, Frankfurt am Main, New York, Vienna: Peter Lang.
- Commissione sulla lingua internazionale (detta esperanto).** (1995). Diffusione del documento conclusivo della Commissione per la promozione della Lingua Internazionale detta « Esperanto ». *Bolletino Ufficiale, Ministero della pubblica istruzione*, 122(21–22).
- Conroy, M. J., Runge, M. C., Nichols, J. D., Stodola, K. W., & Cooper, R. J.** (2011). Conservation in the face of climate change: The roles of alternative models, monitoring, and adaptation in confronting and reducing uncertainty. *Biological Conservation*, (144), 1204–1213.
- Convery, I.** (2012). A review of research ethics in internet-based research. *Practitioner Research in Higher Education*, 6(1), 50–57.

- Corsetti, R.** (1996). A mother tongue spoken mainly by fathers. *Language Problems & Language Planning*, 20, 263–273.
- Corsetti, R.** (1999). Kreoliĝo de Esperanto inter personaj gustoj kaj oficiala normo. In D. Blanke, R. McCoy, & O. Buller (Ed.), *Por aktiva lingvopolitiko. Aktoj de la Lingvopolitika Seminario en la 81-a Universala Kongreso de Esperanto, Prago 1996* (pp. 43–60). Rotterdam: Universala Esperanto-Asocio.
- Coseriu, E.** (1977). Sistema, norma y habla. In F. Abad Nebot (Ed.), *Lecturas de sociolingüística* (pp. 43–51). Madrid.
- Costa, J.** (2016). Sobre el «marc de les condicions sociolingüístiques favorables» per a la implantació de la normativa lingüística catalana [On the «favourable sociolinguistic conditions framework» for implementing the linguistic norms]. *Caplletra*, (61), 125–154.
- Couillard, D.** (2006). *Managing in a Sea of Uncertainty: Leadership, Learning, and Resources for the High Tech Firm*. Québec: Presses internationales Polytechnique.
- Coupland, N., & Jaworski, A.** (1993). Sociolinguistic perspectives on metalanguage. In J. A. Lucy (Ed.), *Reflexive language: reported speech and metapragmatics* (pp. 15–51). Cambridge: University press.
- Couturat, L., & Leau, L.** (1903). *Histoire de la langue universelle*. Paris: Librairie Hachette et Cie.
- Cova, B., Dalli, D., & Zwick, D.** (2011). Critical perspectives on consumers' role as «producers»: Broadening the debate on value co-creation in marketing processes. *Marketing Theory*, 11(3), 1–11.
- Cramer, M.** (2018). Kiel krei fidindan Esperanto-vortaron? *Język. Komunikacja. Informacja*, (13), 81–100.
- Crawford, P., & Bryce, P.** (2003). Project monitoring and evaluation: a method for enhancing the efficiency and effectiveness of aid project implementation. *International Journal of Project Management*, 21, 363–373.
- Creese, S.** (2013). Exploring the Relationship between Language Change and Dictionary Compilation in the Age of the Collaborative Dictionary. *Proceedings of ELex 2013*, 392–406. Tallinn, Estonia.
- Cristea, D., Forăscu, C., Răschip, M., & Zock, M.** (2008). How to Evaluate and Raise the Quality in a Collaborative Lexicographic Approach. *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008*. Presentation at the International Conference on Language Resources and Evaluation, LREC 2008, Marrakech, Morocco.
- Croft, W.** (2000). *Explaining language change. An evolutionary approach*. Essex: Pearson Education Limited.

- Cruse, D. A.** (1986). *Lexical Semantics*. Cambridge, New York: Cambridge University Press.
- Crystal, D.** (1985). How many millions? The statistics of English today. *English Today*, 1, 7–9.
- Crystal, D.** (2008). Two thousand million? Updates on the statistics of English. *English Today*, 24(1), 3–6.
- Cukor-Avila, P.** (2000). Revisiting the Observer's Paradox. *American Speech*, 75(3), 253–254.
- Daoust, D.** (1986). Le changement terminologique planifié : un cas particulier de changement linguistique. *Revue québécoise de linguistique*, 15(2), 247–269.
- Daoust, D.** (1995). Quelques facteurs sociodémographiques qui sous-tendent le choix des langues pour la terminologie en milieu de travail et leur interrelation avec des attitudes. *Meta : journal des traducteurs / Meta: Translators' Journal*, 40(2), 260–283.
- Dasgupta, P.** (1993). Idiomaticity and Esperanto texts: an empirical study. *Linguistics*, 31, 367–386.
- De Ketele, J.-M., & Roegiers, X.** (2009). *Méthodologie du recueil d'informations. Fondements des méthodes d'observation, de questionnaire, d'interview et d'étude de documents (4^e édition)*. Bruxelles: de Boeck.
- De Pelsmacker, P., Geuens, M., & Van den Bergh, J.** (2005). *Foundations of Marketing Communications. A European Perspective*. Essex: Pearson Education Limited.
- de Saussure, F.** (1966). *Course in general linguistics* (3rd ed.; C. Bally, A. Sechehaye, & A. Riedlinger, Ed.; W. Baskin, Transl.). New York: McGraw-Hill Book Co.
- de Saussure, R.** (1908). *Die «Mängel» des Esperanto* (J. Schmid, Transl.). Bern: Böhler & Co., 1917.
- de Saussure, R.** (1910). *La construction logique des mots en espéranto: réponse à des critiques*. Genève: Librairie Kündig.
- de Saussure, R.** (1985). *La vort-strukturo en Esperanto* (Reeldono). Brazilia.
- de Singly, F.** (2012). *L'enquête et ses méthodes. Le questionnaire* (3rd ed.). Armand Colin.
- de Vaus, D.** (2014). *Surveys in Social Research* (6th Edition). London & New York: Routledge.
- de Vecchi, D.** (2004). La terminologie de la communication d'entreprise : bases d'une pragmatérminologie. *Cahier du C.I.E.L.*, (2004), 71–84.

- de Vecchi, D.** (2016). Approche pragmateterminologique des termes des entreprises et des organisations. *Pragmateterminologie*, (12), Pragmateterminologie.
- Deklaracio pri la Esenco de la Esperantismo.** (1905). Accessed on 2016/12/11: http://www.akademio-de-esperanto.org/fundamento/bulonja_deklaracio.html
- Delavigne, V.** (2000). Le double jeu de l'autonymie. *Actes du colloque international « Le fait autonymique dans les langues et les discours »*. Retrieved from <http://www.cavi.univ-paris3.fr/ilpga/autonymie/actes.htm>
- Delavigne, V.** (2001). *Les mots du nucléaire. Contribution sociotermiologique à une analyse des discours de vulgarisation* (Thèse pour obtenir le grade de docteur de l'Université de Rouen. Discipline : linguistique). Université de Rouen.
- Delavigne, V.** (2003). Quand le terme entre en vulgarisation. *Terminologie et Intelligence artificielle (TIA)*, 80–91. Strasbourg.
- Delavigne, V.** (2017). Term usage and sociotermiological variation. The impact of social and local issues on the movement of terms. In P. Drouin, A. Francœur, J. Humbley, & A. Picton (Ed.), *Multiple Perspectives on Terminological Variation* (pp. 31–55). Amsterdam, Philadelphia: John Benjamins.
- Demeyere, B.** (2010). ViVo. In K. Noskova & P. Baláž (Ed.), *Modernaj teknologioj por Esperanto* (pp. 105–114). Modra (SK): ESPERO.
- Demeyere, B.** (n. d.). ViVo (Vikia Vortaro). Accessed on 2016/12/18: <http://www.kono.be/vivo/>
- Depecker, L.** (Ed.). (1997). *La mesure des mots. Cinq études d'implantation terminologique*. Rouen: Publications de l'Université de Rouen.
- Depecker, L.** (2005). Quels principes dégager pour traiter la néologie ? Les néologies contemporaines: actes du colloque du 15 octobre 2004 organisé à l'École normale supérieure de la rue d'Ulm, 13–26. Paris: Le savoir des mots.
- Depecker, L.** (2013). Pour une ethnoterminologie. In J. Quirion, L. Depecker, & L.-J. Rousseau (Ed.), *Dans tous les sens du terme* (pp. 14–29). Ottawa: Presses de l'Université d'Ottawa.
- Derks, J.** (2017). *Daŭro, demokrateco kaj komunumeco en esperantaj dissendolistoj*. Uniwersytet im. Adama Mickiewicza w Poznaniu, Poznan.
- Desmet, I.** (2005). Variabilité et variation en terminologie et langues spécialisées : discours, textes et contextes. *Préactes des 7es journées scientifiques AUF-LTT. Mots, termes et contextes*. Présenté à Bruxelles. Retrieved from <http://perso.univ-lyon2.fr/~thoiron/JS%20LTT%202005/pdf/Desmet.pdf>

- Desmet, I.** (2007). Terminologie, culture et société. Éléments pour une théorie variationniste de la terminologie et des langues de spécialité. *Cahiers du Rifal*, (26), 3–13.
- Díaz Hormingo, M. T.** (2012). Lexical Creation and Euphemism: Regarding the Distinction Denominative or Referential Neology vs. Stylistic or Expressive Neology. *Lexis. Journal in English Lexicology*, (7).
<https://doi.org/10.4000/lexis.371>
- Dierks, A.** (2017). *Re-Modeling the Brand Purchase Funnel*.
https://doi.org/10.1007/978-3-658-17822-2_2
- Diki-Kidiri, M., Joly, H., & Murcia, C.** (1981). *Guide de la néologie*. Paris: Conseil international de la langue française.
- Ding, X., Liu, B., & Yu, P. S.** (2008). A Holistic Lexicon-Based Approach to Opinion Mining. *WSDM '08 Proceedings of the 2008 International Conference on Web Search and Data Mining*, 231–239. Palo Alto: ACM.
- Doan, A., Ramakrishnan, R., & Halevy, A. Y.** (2011). Mass Collaboration Systems on the World-Wide Web. *ACM*, 54(4), 86–96.
- Domanskiy, D.** (2016). Коммуникативно-функциональные характеристики глаголов-неологизмов во французском медийном дискурсе [Communicative and functional characteristics of verbal neologisms in the discourse of the French media] (На соискание ученой степени кандидата филологических наук [PhD thesis in philological sciences]). Moscow City University, Moscow.
- Domínguez Chamizo, P. J., & Nerlich, B.** (2002). False friends: their origin and semantics in some selected languages. *Journal of Pragmatics*, 34, 1833–1849.
- Doury, M.** (2001). *La fonction argumentative des échanges rapportés*. Présenté à La circulation des discours : les frontières du discours rapporté, Bruxelles.
- Drame, A.** (2009). Terminology Policies and Communication for Social Change: Promoting Linguistic Diversity and Terminology in South Africa (Universität Wien). Retrieved from http://othes.univie.ac.at/8922/1/2009-10-04_0647867.pdf
- Dreyfus, H. L., & Dreyfus, S. E.** (1986). *Mind over Machine: The Power of Human Intuition and Expertise in the Era of the Computer*. Oxford: Basil Blackwell.
- Drezen, E. K.** (1930). *Analiza historio de Esperanto-Movado*. L'Omnibuso.
- Duin, O.** (2006). *Esperanto in theorie en praktijk* (Bachelorscriptie). Universiteit van Amsterdam, Amsterdam.

- Duliĉenko, A.** (1990). *Международные вспомогательные языки* [International auxiliary languages]. Tallinn: Valgus.
- Duliĉenko, A. D.** (1989). Ethnic language and planned language. On the particulars of the structural-genetic and the functional aspect. In *Trends in Linguistics. Studies and Monographs: Vol. 42. Interlinguistics. Aspects of the Science of Planned Languages* (pp. 47–61). Berlin, New York: Mouton de Gruyter.
- Duliĉenko, A. D.** (1997). Esperanto: A Unique Model for General Linguistics. In H. Tonkin (Ed.), *Esperanto, interlinguistics, and planned language* (pp. 66–69). Lanham, Maryland: University Press of America.
- Dury, P.** (2013). Que montre l'étude de la variation d'une terminologie dans le temps. Quelques pistes de réflexion appliquées au domaine médical. *Debate Terminológico*, (9), 2–10.
- Dylan Project.** (2006). Working papers. Accessed on 2019/04/03: http://www.dylan-project.org/Dylan_en/dissemination/page102/page102.php
- Eastin, M. S.** (2002). Diffusion of e-commerce: an analysis of the adoption of four e-commerce activities. *Telematics and Informatics*, (19), 251–267.
- Edwards, J.** (2010). *Minority Languages and Group Identity. Cases and Categories*. Amsterdam/Philadelphia: John Benjamins.
- Edwards, J.** (2012). *Multilingualism. Understand Linguistic Diversity*. London, New York: Continuum International Publishing Group.
- Eichholz, R.** (1995). Principoj por selekto de fakaj terminoj. *La Stato kaj Estonteco de la Internacia Lingvo Esperanto*, 72–88. Prago: KAVA-PECH.
- Eichhorn, P., & Towers, I.** (2018). *Principles of Management. Efficiency and Effectiveness in the Private and Public Sector*. Switzerland: Springer.
- Ellis, N. C.** (2011). The emergence of language as a complex adaptative system. In J. Simpson (Ed.), *The Routledge Handbook of Applied Linguistics* (pp. 654–667). London, New York: Routledge.
- Engelberg, S., & Storrer, A.** (2016). Typologie von Internetwörterbüchern und -portalen. In A. Klosa & C. Müller-Spitzer (Ed.), *Internetlexikografie. Ein Compendium* (pp. 31–64).
- Eoyang, G. H., & Berkas, T.** (1998). Evaluation in a complex adaptive system. In *Managing complexity in organizations* (pp. 313–335). Westport, Conn: Quorum Books.
- Esperantic Studies Foundation.** (n. d.). Tekstaro de Esperanto. Accessed on 2018/05/05: <http://tekstaro.com/>
- Esperanto-tradukistoj.** (2003, 2016). Accessed on 2016/12/18: <https://groups.yahoo.com/neo/groups/tradukado/info>

- Estopà, R., Coromina, E., & Mestres, J. M.** (2010). La neologia especialitzada, repte constant per al mediador lingüístic. In E. Coromina i Pou & J. M. Mestres i Serra (Ed.), *Aspectes de terminologia, neologia i traducció* (pp. 15–40). Barcelona, Vic: Institut d'Estudis Catalans, Universitat de Vic.
- Fiebleman, J. K.** (1961). Pure Science, Applied Science, Technology, Engineering: An Attempt at Definitions. *Technology and Culture*, 2(4), 305–317.
- Felber, H., & Lang, F.** (1979). Würdigung der Person und des Wissenschaftlers. In *Terminologie als angewandte Sprachwissenschaft. Gedenkschrift für Univ.-Prof Dr. Eugen Wüster* (pp. 15–28). München: K. G. Verlag Saur KG.
- Fernández Silva, S.** (2013). La influencia del área disciplinar en la variación terminológica: Un estudio en un corpus interdisciplinario sobre pesca. *Revista Signos*, 43(83), 361–388.
- Fettes, M.** (1997). Esperanto and Language Policy: Exploring the Issues. *Language Problems and Language Planning*, 21(1), 66–77.
- Fettes, M.** (2013). *The ethics of Esperanto. Norms of conduct in a cosmopolitanized speech community*. Retrieved from <http://edukado.net/biblioteko/filmejo?iid=233>
- Fiedler, S.** (1998). Zu Merkmalen der Esperanto-Sprachgemeinschaft (Ergebnisse einer Probandenbefragung). In U. Becker (Ed.), *Beiheft 4 « Soziokulturelle Aspekte von Plansprachen »* (pp. 22–29). Berlin.
- Fiedler, S.** (1999a). Phraseology in Planned Languages. *Language Problems and Language Planning*, 23(2), 175–187.
- Fiedler, S.** (1999b). Plansprache und Phraseologie. Empirische Untersuchungen zu reproduziertem Sprachmaterial im Esperanto. Frankfurt am Main: Peter Lang.
- Fiedler, S.** (1999c). Pri Esperanto-frazeologio. *Esperantologio / Esperanto Studies*, 1, 48–51.
- Fiedler, S.** (2002). On the main characteristics of Esperanto communication. In K. Knapp & C. Meierkord (Ed.), *Lingua Franca Communication* (pp. 53–86). Frankfurt/Main: Peter Lang.
- Fiedler, S.** (2006). Standardization and self-regulation in an international speech community: the case of Esperanto. *International Journal of the Sociology of Language*, (117), 67–90.
- Fiedler, S.** (2012). The Esperanto denaskulo. The status of the native speaker of Esperanto within and beyond the planned language community. *Language Problems and Language Planning*, 36(1), 69–84.

- Fiedler, S.** (2014). Geschlecht im Esperanto. Eine sprachwissenschaftliche Betrachtung zu genderspezifischen Bezeichnungen in einer Plansprache. In C. Brosch & S. Fiedler (Ed.), *Interlinguistik im 21. Jahrhundert. Beiträge der 23. Jahrestagung der Gesellschaft für Interlinguistik e.V.* (pp. 45–60). Berlin.
- Fiedler, S.** (2015). The topic of planned languages (Esperanto) in the current specialist literature. *Language Problems & Language Planning*, 39(1), 84–104. <https://doi.org/10.1075/lplp.39.1.05fie>
- Finegan, E.** (1995). Subjectivity and subjectivisation: an introduction. In D. Stein & S. Wright (Ed.), *Subjectivity and subjectivisation: linguistic perspectives*. Cambridge, New York: Cambridge University Press.
- Finney Rutten, L. J., Arora, N. K., Bakos, A. D., Aziz, N., & Rowland, J.** (2005). Information needs and sources of information among cancer patients: a systematic review of research (1980–2003). *Patient Education and Counseling*, 57(3), 250–261.
- Firth, J. R.** (1957). A synopsis of linguistic theory 1930–1955. In *Studies in Linguistic Analysis* (pp. 1–32). Oxford: Basil Blackwell.
- Fischer Hubert, D.** (2001). Le lexique de l'informatique et de l'internet en français. In IV Congrès internacional sobre llengües per a finalitats específiques. The Language of International Communication. Español de los negocios (pp. 115–120). Barcelona: Publicacions de la Universitat de Barcelona.
- Flick, U.** (1998). *An Introduction to Qualitative Research* (1st ed.). London, Thousand Oaks, New Delhi: SAGE Publications.
- Flinz, C.** (2014). Wörterbuchbenutzung: Ergebnisse einer Umfrage bei italienischen DaF-Lernern. *Proceedings of the XVI EURALEX International Congress: The User in Focus*, 213–224. Retrieved from www.euralex.org/elx_proceedings/Euralex2014/euralex_2014_014_p_213.pdf
- Forster, P. G.** (1982). *The Esperanto Movement* (Reprint). The Hague, Paris, New York: Mouton Publishers.
- FranceTerme.** (n.d.). Suggérez vos termes : wikiLF. Accessed on 2013/11/28: <http://wikilf.culture.fr/suggerez-vos-termes>
- FranceTerme.** (n.d.). wikiLF. Accessed on 2013/10/17: <http://wikilf.culture.fr/>
- Freixa i Aymerich, J. F.** (2015). La Implantació del lèxic normatiu: estudi complementari del lèxic implantat no normatiu [The implantation of standardized lexicon: a complementary study of non-standardized implanted lexical units]. *Treballs de sociolingüística catalana*, (25), 65–80.
- Freixa, J.** (2006). Causes of denominative variation in terminology. A typology proposal. *Terminology*, 12(1), 51–77.

- Frøkjær, E., Hertzum, M., & Hornbæk, K.** (2000). Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 345–352. <https://doi.org/10.1145/332040.332455>
- Fuertes, I.** (2008). Towards Harmonization or Standardization in Governmental Accounting? The International Public Sector Accounting Standards Board Experience. *Journal of Comparative Policy Analysis: Research and Practice*, 10, 327–345.
- Fuertes-Olivera, P. A., & Tarp, S.** (2008). La teoría funcional de la Lexicografía y sus consecuencias para los diccionarios de economía del español. *Revista de Lexicografía*, (XIV), 75–95.
- Fugger, B.** (1980). Die Einstellung der Franzosen zur französischen Sprachpolitik. In *Zur Geschichte des gesprochenen Französisch und zur Sprachlenkung im Gegenwartsfranzösischen* (pp. 58–78). Wiesbaden: Franz Steiner Verlag.
- Fugger, B.** (1983). Sprachentwicklung – Sprachbeeinflussung – Sprachbewusstsein. Eine soziolinguistische Untersuchung zur französischen Sprachpolitik. *Fachsprache*, (3), 128–137.
- García Llamas, B. I.** (2015). La Teoría Funcional de la Lexicografía y Su Aplicación a la Construcción de Diccionarios Especializados. Una Propuesta de Diccionario de Comercio Internacional (Tesis doctoral). Universidad nacional de educación a distancia (UNED).
- Gardin, B.** (1974). La néologie. Aspects sociolinguistiques. *Langages*, (36 (La néologie lexicale)), 67–73.
- Gaudin, F.** (1993a). Pour une socioterminologie: des problèmes sémantiques aux pratiques institutionnelles. Rouen: Université de Rouen.
- Gaudin, F.** (1993b). Socioterminologie : propos et propositions épistémologiques. *La Langage et l'Homme*, XXVIII(4 (spécial SOCIO-TERMINOLOGIE)), 247–258.
- Gaudin, F.** (2003). Socioterminologie : une approche sociolinguistique de la terminologie. Louvain-la-Neuve: Duculot.
- Gaudin, F.** (2005). La socioterminologie. *Langages*, 39e année(157), 80–92. <https://doi.org/10.3406/lgge.2005.976>
- Gaudin, F.** (2007). Quelques mots sur la socioterminologie. *Cahiers du Rifal*, (26), 26–35.
- Gaudin, F., & Guespin, L.** (1993). Étude d'implantation des arrêtés de terminologie. Domaine : génie génétique. In L. Depecker (Ed.), *La mesure*

- des mots. Cinq études d'implantation terminologique.* (pp. 1–45). Rouen: Publications de l'Université de Rouen.
- Gazzola, M.** (2011). The economic evaluation of efficiency and fairness in the management of multilingual communication: the case of Patent Offices. Université de Genève, Genève.
- Geeraerts, D.** (1994). Varieties of lexical variation. *Euralex*, 78–83. Retrieved from <http://euralex.org/publications/varieties-of-lexical-variation/>
- Geeraerts, D., Grondelaers, S., & Bakema, P.** (2010). *The Structure of Lexical Variation. Meaning, Naming, and Context* (Reprint).
- Gendron, A.-M., & Messina, C.** (2015). Neologia e norme. Criteria terminologici per la creazione e la valutazione dei neologismi. *VALS-ASLA*, (Numéro spécial, tome 1), 111–127.
- Gervais, M., & Fessler, D.** (2017). On the deep structure of social affect: Attitudes, emotions, sentiments, and the case of “contempt”. *Behavioral and Brain Sciences*, (40), 1–18.
- Gill, S. K.** (2013). *Language Policy Challenges in Multi-Ethnic Malaysia*.
- Gilreath, C. T.** (1992). Harmonization of Terminology: An Overview of Principles. *International Classification*, 19(3), 135–139.
- Gledhill, C.** (2000). *The Grammar of Esperanto. A corpus-based description*. Retrieved from http://stl.recherche.univ-lille3.fr/sitespersonnels/gledhill/Esperanto_a_corpus-based_description_GLEDHILL.pdf
- Gledhill, C.** (2014). Phraseology as a measure of emergent norm: The case of Esperanto. In J. C. Herreras (Ed.), *Politiques linguistiques et langues autochtones d'enseignement dans l'Europe des vingt-sept* (pp. 317–348). Presses universitaires de Valenciennes.
- Glunk, R.** (1966). Erfolg und Mißerfolg der nationalsozialistischen Sprachlenkung. *Zeitschrift für deutsche Sprache*, (22), 57–73.
- Glunk, R.** (1967a). Erfolg und Mißerfolg der nationalsozialistischen Sprachlenkung (Fortsetzung). *Zeitschrift für deutsche Sprache*, (23), 83–113.
- Glunk, R.** (1967b). Erfolg und Mißerfolg der nationalsozialistischen Sprachlenkung (Fortsetzung). *Zeitschrift für deutsche Sprache*, (23), 178–188.
- Gobbo, F.** (1997). Il dilemma dell'esperanto. Tra vocazione ausiliaria e naturalizzazione. Università degli Studi di Torino, Torino.
- Gobbo, F.** (2009a). Fondamenti di interlinguistica ed esperantologia: Pianificazione linguistica e lingue pianificate. Milano: Raffaello Cortina.
- Gobbo, F.** (2009b). Pianificare una lingua pianificata. L'Accademia di Esperanto tra tradizione e innovazione. *Multilinguismo e Società, Anno 2010*.

- Gobbo, F.** (2017). The lexicography of Esperanto. In P. Hanks & G.-M. De Schryver (Ed.), *International Handbook of Modern Lexis and Lexicography* (pp. 1–10). Germany: Springer-Verlag GmbH.
- Goebel, H.** (2012). Quelques remarques relatives aux concepts Abstand et Ausbau de Heinz Kloss. In U. Ammon (Ed.), *Status and Function of Languages and Language Varieties* (Reprint 2012, p. 278–290). <https://doi.org/10.1515/9783110860252>
- Goes, H.** (2007). Afero de Espero. Konciza historio de la Esperanto-movado en Afriko. Rotterdam: Universala Esperanto-Asocio.
- Golden, B.** (1990). La kreado, aprobo kaj oficialigo de vortoj en Esperanto. In R. Eichholz (Ed.), *Akademioj studoj 1988–1990* (pp. 197–199). Bailieboro (CA): Esperanto Press.
- Gonseth, F.** (1975). Le référentiel, univers obligé de médiatisation. Lausanne: L'âge d'homme.
- Gorter, D.** (2006). Introduction: The Study of the Linguistic Landscape as a New Approach to Multilingualism. *Linguistic Landscape: A New Approach to Multilingualism*, 3(1), 1–6.
- Gouadec, D., Crespel, A., & Colombel, H.** (1993). Étude d'implantation des arrêtés de terminologie. Domaine: informatique. In L. Depecker (Ed.), *La mesure des mots. Cinq études d'implantation terminologique*. (pp. 235–493). Rouen: Publications de l'Université de Rouen.
- Gregori, L., & Panunzi, A.** (2017). Measuring the Italian-English lexical gap for action verbs and its impact on translation. *Proceedings of the Workshop on Sense, Concept and Entity Representations and Their Applications*, 102–109. Valencia, Spain: The Association for Computational Linguistics.
- Gresa Barbero, D.** (2016). Estudi d'implantació terminològica. La gastronomia del Pròxim Orient i el nord d'Àfrica als restaurants de Gràcia i l'Eixample de Barcelona. Barcelona: scaterm.
- Griesbach, T.** (2006). Unwort und laienlinguistische Wortkritik. Zur Erforschung des sprachkritischen Denkens in Deutschland. Aachen: Shaker Verlag.
- Grin, F.** (2005). L'enseignement des langues étrangères comme politique publique. Haut conseil de l'évaluation de l'école.
- Grin, F., & Gazzola, M.** (2007). Efficiency and fairness (Working Papers of the Dylan Project, Working Paper No 1, pp. 9–31).
- Grosjean-Maupin, E., Esselin, A., Grenkamp-Kornfeld, S., & Waringhien, G.** (1930). *Plena Vortaro de Esperanto* (Unua Eldono). Paris: Sennacieca Asocio Tutmonda.

- Guest, G., Bunce, A., & Johnson, L.** (2006). How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods*, 18(1), 59–82.
- Guilbert, L.** (1965). *La formation du vocabulaire de l'aviation*. Université de Paris, Paris.
- Guilbert, L.** (1973). Théorie du néologisme. *Cahiers de l'Association internationale des études francaises*, (25), 9–29. <https://doi.org/10.3406/caief.1973.1020>
- Guilbert, L.** (1975). *La créativité lexicale*. Paris: Larousse.
- Guilford, J.** (1997). Les attitudes des jeunes Français à propos des emprunts à l'anglais. *La Linguistique*, 33(2), 117–135.
- Guinard, T.** (2016). An Algorithm for Morphological Segmentation of Esperanto Words. *The Prague Bulletin of Mathematical Linguistics*, (105), 63–76.
- Gumperz, J. J.** (2009). The Speech Community. In A. Duranti (Ed.), *Linguistic Anthropology: a Reader* (pp. 66–73). Malden: John Wiley & Sons.
- Haag, M., & Brodersen, B.** (2013). *Terminologiarbeit im Unternehmen – wie fängt man da an?* Workshop an der tekomp Jahrestagung 2013. Retrieved from http://tagungen.tekom.de/fileadmin/tx_doccon/slides/331_Terminologiemangement_im_Unternehmen_wie_f_ngt_man_da_an_.pdf
- Haarmann, H.** (2012). Functional Aspects of Language Varieties – A Theoretical-Methodological Approach. In U. Ammon (Ed.), *Status and Function of Languages and Language Varieties* (Reprint 2012, p. 153–193). <https://doi.org/10.1515/9783110860252>
- Habert, B.** (2000). Des corpus représentatifs : de quoi, pour quoi, comment ? In M. Billiger (Ed.), *Linguistique sur corpus* (pp. 11–58). Perpignan: Presses universitaires de Perpignan.
- Habert, B., Nazarenko, A., & Salem, A.** (1997). *Les linguistiques de corpus*. Paris: Armand Colin.
- Harris, Z.** (1991). *A theory of language and information : a mathematical approach*. Oxford: Clarendon Press.
- Haszpra, O.** (2001). Kritika trarigardo de la provoj modernigi la Fundamentan alfabeton de la komencoj ĝis 2001. In Akademio de Esperanto (Ed.), *Lingva Planado kaj Leksikologio* (pp. 133–162). Zagreb: Fonto.
- Heckhausen, J., & Heckhausen, H.** (2018). *Motivation and Action*. New York: Springer.
- Helfrich, U.** (1993). Neologismen auf dem Prüfstand: ein Modell zur Ermittlung der Akzeptanz französischer Neologismen. Wilhelmsfeld: Egert.

- Heller, M.** (1978). *Projet pilote : l'usage et la connaissance de la terminologie de l'automobile à Montréal : variation et distribution sociale*. Office de la langue française.
- Henriksen, B.** (1999). Three dimensions of vocabulary development. *SSLA*, (21), 303–317.
- Heringer, H. J., & Wimmer, R.** (2015). *Sprachkritik: eine Einführung*. Paderborn: Fink.
- Hermans, A.** (1994). Diffusion et propagation. *Terminologies nouvelles*, 12, 40–44.
- Hermans, A.** (1995). Sociologie des discours scientifiques. Quelques réflexions. *Meta : journal des traducteurs / Meta : Translators' Journal*, 40(2), 224–228.
- Hermans, A., & Vansteelandt, A.** (1999). Néologie traductive. *Terminologies nouvelles*, (20), 37–43.
- Herrera, V.** (2005). La néologie d'aménagement : l'expérience des communautés de langues néolatines. In L. Depecker & V. Dubois (Ed.), *Les néologies contemporaines* (pp. 101–109). Paris: Société française de terminologie.
- Herwartz, R.** (2007). *Erfolgreiches Terminologie-Management im Unternehmen*. Vortrag an der tekom Regionalgruppe Nord. Retrieved from <http://docplayer.org/16595114-Erfolgreiches-terminologie-management-im-unternehmen.html>
- Hoffmann, H.** (2001). Komenco de nociosistemigo ĉe fervojo. *Scienca Revuo*, 2, 19–26.
- Holz, G.** (1951). Joachim Heinrich Campe als Sprachreiner und Wortschöpfer (Dissertation). Tübingen.
- Houdebine, A.-M., & Fodor, F.** (2013). Normes et imaginaires linguistiques dans l'élaboration des langues française et hongroise. In W. M. Ayres & M. Sejjido (Ed.), *Bon usage et variation sociolinguistique : perspectives diachroniques et traditions nationales* (pp. 289–298). Lyon: ENS éd.
- Houdebine-Gravaud, A.-M.** (2004). L'imaginaire linguistique : un niveau d'analyse et un point de vue théorique. In A.-M. Houdebine-Gravaud (Ed.), *L'imaginaire linguistique* (pp. 9–21). Paris: L'Harmattan.
- Hsieh, H.-F., & Shannon, S. E.** (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Hu, M., & Liu, B.** (2004). Mining and Summarizing Customer Reviews. *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177. <https://doi.org/10.1145/1014052.1014073>

- Hughes, J., & Lang, K. R.** (2004). Issues in Online Focus Groups: Lessons Learned from an Empirical Study of Peer-to-Peer Filesharing Systems Users. *European Journal of Business Research Methods*, 2(2), 95–110.
- Humbley, J.** (2004). La réception de l'œuvre d'Eugen Wüster dans les pays de langue française. *Cahiers du CIEL*, 33–51.
- Humbley, J.** (2006). Metaphor and secondary term formation. *Cahiers Du CIEL*, (2000–2003), 199–212.
- Humbley, J.** (2015). Allogenisms: The major category of « true » false loans. In C. Furiassi & H. Gottlieb (Ed.), *Pseudo-English: Studies on False Anglicisms in Europe* (pp. 36–58). De Gruyter.
- Humphreys, A., & Grayson, K.** (2008). The Intersecting Roles of Consumer and Producer: A Critical Perspective on Co-production, Co-creation and Prosumption. *Sociology Compass*, (2).
- Hundt, M.** (2010). New norms – How new grammatical constructions emerge. In A. N. Lenz & A. Plewnia (Ed.), *Grammar between Norm and Variation* (pp. 27–58). Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien: Peter Lang.
- Hungarian Central Statistical Office.** (2011). *Population by language knowledge and sex*. Retrieved from http://www.ksh.hu/nepszamlalas/docs/tables/regional/00/00_1_1_4_2_en.xls
- Hutchby, I., & Wooffitt, R.** (2008). *Conversation Analysis*. Cambridge: Polity.
- Hymes, D.** (1964). Introduction: Toward Ethnographies of Communication. *American Anthropologist*, 66(6), 1–34.
- Informoj por redaktantoj kaj redaktontoj.** (n. d.). Accessed on 2016/08/12: <http://purl.org/net/voko/revo/dok/redinfo.html>
- Internacia Fervojista Esperanto-Federacio (IFEF).** (2015, novembre). *Terminara Kuriero (TeKu) 152. Komunikilo de la Terminara Sekcio en la Faka Komisiono de IFEF*. Retrieved from http://ifef.free.fr/spip/IMG/pdf/TeKu_152.pdf
- Internacia Fervojista Esperanto-Federacio (IFEF).** (2016, March). *Terminara Kuriero (TeKu) 153. Komunikilo de la Terminara Sekcio en la Faka Komisiono de IFEF*. Retrieved from http://ifef.free.fr/spip/IMG/pdf/TeKu_153.pdf
- International Organization for Standardization (ISO).** (2007). *ISO/TR 22134:2007. Practical guidelines for socioterminology*.
- Ischreyt, H.** (1965). Studien zum Verhältnis von Sprache und Technik : institutionelle Sprachlenkung in der Terminologie der Technik. Düsseldorf: Schwann.

- ISO/TC 37.** (2009). *ISO 704:2009 (E). Terminology work — Principles and methods*. International Organization for Standardization (ISO).
- Jackson, H.** (2013). *The Bloomsbury Companion to Lexicography*. London, New Delhi, New York, Sidney: Bloomsbury.
- Jacquet, A., & Rosier, L.** (2014). Les commentaires métalinguistiques des internautes sur les sites d'information belges. *Studii de lingvistică*, 4, 111–130.
- Jakobson, R.** (1960). Linguistics and Poetics. In Sebeok, T.A. (Ed.), *Style in Language* (pp. 350–377). New York: Wiley.
- Jakobson, R.** (1980). Metalanguage as a linguistic problem. In *Michigan Studies in the Humanities. the framework of language* (pp. 81–92).
- Jansen, B. J.** (2011). Bidding on the buying funnel for sponsored search and keyword advertising. *Journal of Electronic Commerce Research*, 12(1).
- Jansen, W.** (2007). *Woordvolgorde in het Esperanto. Normen, taalgebruik en universalialia*. Universiteit van Amsterdam: LOT.
- Janton, P.** (1977). *L'espéranto* (2^e édition). Paris: Presses universitaires de France.
- Jaworski, B., & Kohli, A. K.** (2006). Co-creating the voice of the customer. In R. F. Lusch & S. Vargo (Ed.), *The Service-dominant Logic of Marketing: Dialog, Debate, and Directions* (pp. 109–116). London & New York: Routledge.
- Johnston, L., & Hewstone, M.** (1992). Cognitive Models of Stereotype Change. 3. Subtyping and the Perceived Typicality of Disconfirming Group Members. *Journal of Experimental Social Psychology*, (28), 360–386.
- Jones, R. P., & Runyan, R. C.** (2016). Conceptualizing a path-to-purchase framework and exploring its role in shopper segmentation. *International Journal of Retail & Distribution Management*, 44(8), 776–798.
<https://doi.org/10.1108/IJRDM-09-2015-0148>
- Juhász, J.** (1970). *Probleme der Interferenz*. München: Max Hueber Verlag.
- Kabatek, J.** (1996). *Die Sprecher als Linguisten, Interferenz- und Sprachwandelphänomene dargestellt am Galicischen der Gegenwart* (Reprint 2012). Retrieved from <https://www.degruyter.com/viewbooktoc/product/51930?rskey=Zhkjaw&result=6>
- Kadoja, H.** (2013). *Dekonstruado de esperanta frazeologio* (C. O. Kiselman & M. Maradan, Ed.). Rotterdam: Universala Esperanto-Asocio.
- Kalocsay, K., & Waringhien, G.** (1980). *Plena analiza gramatiko de Esperanto* (Kvara, tralaborita, eldono). Rotterdam: Universala Esperanto-Asocio.

- Kaplan, R. B.** (1997). *Language planning: from practice to theory*. Clevedon: Multilingual Matters.
- Karabacak, E.** (2009). Acceptance of terminology sanctioned by the Turkish Language Society. A study of the use of economic terms in Turkish newspapers. *Terminology*, 15(2), 145–178.
- Keller, R.** (2003). *Sprachwandel* (3. Auflage). Tübingen, Basel: A. Francke Verlag.
- Kelley, K., Clark, B., Brown, V., & Sitzia, J.** (2003). Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care*, 15(3), 261–266. <https://doi.org/DD1193>
- Kerbrat-Orecchioni, C.** (2009). *L'énonciation* (4^e édition). Paris: Armand Colin.
- Kilian, J., Niehr, T., & Schiewe, J.** (2010). Laienlinguistische Sprachkritik. In *Germanistische Arbeitshefte: Vol. 43. Sprachkritik, Ansätze und Methoden der kritischen Sprachbetrachtung* (pp. 56–94). Retrieved from <https://www.degruyter.com/viewbooktoc/product/43623?rskey=ZjFohC&result=14>
- Kim, S.-M., & Hovy, E.** (2004). Determining the Sentiment of Opinions. *Proceedings of the COLING Conference*. Presentation in Geneva.
- Kimura, G. C.** (2012). Esperanto and minority languages. A sociolinguistic comparison. *Language Problems and Language Planning*, 36(2), 167–181.
- Kiselman, C.** (1991). Vortkreaĵ procezoj de Esperanto. *Scienca Revuo*, 42, 95–109.
- Klein, L. R., & Ford, G. T.** (2003). Consumer Search for Information in the Digital Age: an Empirical Study of Prepurchase Search for Automobiles. *Journal of Interactive Marketing*, 17(3), 29–49.
- Klosa, A., & Tiberius, C.** (2016). Der lexikographische Prozess. In A. Klosa & C. Müller-Spitzer (Ed.), *Internetlexikografie. Ein Kompendium* (pp. 65–110).
- Konsilio de Eŭropo.** (2007). *Laŭstupa memtaksilo* (edukato.net, Transl.). Retrieved from <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168045bb54>
- Kristensson, P., Matthing, J., & Johansson, N.** (2008). Key Strategies for the Successful Involvement of Customers in the Co-Creation of New Technology-Based Services. *International Journal of Service Industry Management*, 19(4), 474–491.
- Krueger, R. A.** (1998a). *Analyzing & Reporting Focus Group Results*. Thousand Oaks: SAGE Publications.
- Krueger, R. A.** (1998b). *Developing Questions for Focus Groups*. Thousand Oaks: SAGE Publications.

- Krueger, R. A.** (1998c). *Moderating Focus Groups*. Thousand Oaks: SAGE Publications.
- Kück, A.** (2009). *Quantitative Studien zum Esperanto unter besonderer Berücksichtigung der Wortbekanntheit* (Dissertation zur Erlangung des akademischen Grades Doktor der Philosophie). Universität Trier.
- Kunkel, M.** (2015). *The medium is the corpus – Zur Nutzung des Internets als Textkorpus in laienlinguistischen Diskussionsforen*. In S. Hardy, S. Herling, & C. Patzelt (Ed.), *Laienlinguistik im frankophonen Internet* (pp. 201–224). Berlin: Frank & Timme.
- Künzli, A.** (2001). Experts versus novices : l'utilisation de sources d'information pendant le processus de traduction. *Meta : journal des traducteurs / Meta: Translators' Journal*, 46(3), 507–523.
- Kurfess, I., & Schmacht, C.** (2012). *Terminologearbeit als Basis für erfolgreiches Wissensmanagement*. Vortrag an den 8. Stuttgarter Wissensmanagement-Tagen, Stuttgart. Retrieved from https://www.cognitas.de/fileadmin/upload/pdf/Vortraege/WiMa_cognitas_IK_CS.pdf
- Labov, W.** (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.
- Landqvist, F., & Teigland, R.** (2005). Collective Action in Electronic Networks of Practice: An Empirical Study of Three Online Social Structures. In P. Van Den Besselaar, G. De Michelis, J. Preece, & C. Simone (Ed.), *Communities and Technologies 2005* (pp. 359–375). Retrieved from http://dx.doi.org/10.1007/1-4020-3591-8_19
- Landrø, J.** (2008). Purismus als Strategie europäischer Sprachpolitik. Eine Untersuchung am Beispiel der « Kleinsprachen » Sardisch, Isländisch, Lettisch. Hamburg: Verlag Dr. Kovač.
- Language Policy Unit, Council of Europe.** (n. d.). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Retrieved from www.coe.int/lang-CEFR
- Lapenna, I., Lins, U., Carlevaro, T., & Sadler, V.** (Ed.). (1974). *Esperanto en perspektivo. Faktoj kaj analizoj pri la internacia lingvo*. London, Rotterdam: Universala Esperanto Asocio.
- Larsen-Freeman, D.** (2011). A complexity theory approach to second language development/acquisition. In D. Atkinson (Ed.), *Alternative Approaches to Second Language Acquisition* (pp. 48–72). New York: Routledge.
- Lasagabaster, D.** (2008). Attitude. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Ed.), *Sociolinguistics / Soziolinguistik. An International Handbook of the Science of Language and Society / Ein internationales*

- Handbuch zur Wissenschaft von Sprache und Gesellschaft (2nd compl. rev. and extend. ed., pp. 399–405). Retrieved from <https://www.degruyter.com/viewbooktoc/product/178201>
- Lauffer, B., & Paribakht, T. S.** (1998). The Relationship Between Passive and Active Vocabularies: Effects of Language Learning Context. *Language Learning*, 48(3), 365–391. <https://doi.org/10.1111/0023-8333.00046>
- Leblanc, B., & Bilodeau, C.** (2009). Discours épiterminologique : indice du degré de réussite d’implantation des terminologies françaises. *Neologica*, (3), 167–182.
- Lee, B.** (1993). Metalanguages and subjectivities. In J. A. Lucy (Ed.), *Reflexive language : reported speech and metapragmatics* (pp. 365–391). Cambridge: University press.
- Lemnitzer, L.** (2010). Neologismenlexikographie und das Internet. *Lexicographica*, 26, 65–78.
- Lenz, A. N.** (2003). Struktur und Dynamik des Substandards. Eine Studie zum Westmitteldeutschen (Wittlich/Eifel). Stuttgart: Franz Steiner Verlag.
- Leroyer, P., & Möller, B.** (2006). Intégration du genre dans le dictionnaire d’entreprise: Le cas des *case stories*. In M. Gotti & D. S. Giannoni (Ed.), *New Trends in Specialized Discourse Analysis* (pp. 91–108). Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien: Peter Lang.
- Levelt, W. J. M.** (1995). *Speaking: from intention to articulation* (4th print.). Cambridge ; London: The MIT Press.
- Levin, S.** (2002). Complex adaptive systems: Exploring the known, the unknown and the unknowable. *Bulletin (New Series) of the American Mathematical Society*, 40(1), 3–19.
- Lew, R., & De Schryver, G.-M.** (2014). Dictionary Users in the Digital Revolution. *International Journal of Lexicography*, 27(4), 341–359.
- Leyk, J.** (1980). Problemoj de la evoluo de Esperanto. In I. Szerdahelyi (Ed.), *Miscellanea interlinguistica* (pp. 458–464). Budapest: Tankönyvkiadó.
- Li, C., & Bernoff, J.** (2011). *Groundswell : winning in a world transformed by social technologies*. Boston (MA): Harvard Business Review Press.
- Liamputtong, P.** (2011). *Focus Group Methodology. Principles and Practice*. Thousand Oaks: SAGE Publications.
- Lichtman, M.** (2013). *Qualitative Research in Education. A User’s Guide* (3rd ed.). Los Angeles: Sage.
- Liebscher, G., & Dailey-O’Cain, J.** (2009). Language attitudes in interaction. *Journal of Sociolinguistics*, 13(2), 195–222. <https://doi.org/10.1111/j.1467-9841.2009.00404.x>

- Limaye, M., & Pompian, R.** (1991). Brevity Versus Clarity: The Comprehensibility of Nominal Compounds in Business and Technical Prose. 28(1), 7–21.
- Lindstedt, J.** (1983). Recenza eseo. Kalman Kalocsay - Gaston Waringhien: Plena Analiza Gramatiko de Esperanto. Rotterdam 1980: Universala Esperanto-Asocio. Tria parto. *Planlingvistiko*, (6), 7–14.
- Lindstedt, J.** (2006). *Native Esperanto as a Test Case for Natural Language*. Retrieved from http://www.ling.helsinki.fi/sky/julkaisut/SKY2006_1/1FK60.1.5.LINDSTEDT.pdf
- Lindstedt, J.** (2010). Esperanto as a Family Language. In F. Dervin (Ed.), *Lingua francae: La véhicularité linguistique pour vivre, travailler et étudier* (pp. 69–80). Retrieved from https://helda.helsinki.fi/bitstream/handle/10138/25665/Lindstedt_2010_Family.pdf?sequence=2
- Lingva Konsultejo** [Social Network]. (n. d.). Accessed on 2017/09/17: <https://www.facebook.com/groups/konsultejo>
- Lingva manlibro pri verkado de Revo-artikoloj.** (n. d.). Accessed on 2016/08/12: http://purl.org/net/voko/revo/dok/lingva_manlibro.html
- Lipczuk, R.** (2007). Geschichte und Gegenwart des Fremdwortpurismus in Deutschland und Polen. Frankfurt a.M: P. Lang.
- Liu, B.** (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- Liu, B., & Zhang, L.** (2012). A Survey of Opinion Mining and Sentiment Analysis. In C. Aggarwal & C. Zhai (Ed.), *Mining Text Data* (pp. 415–463). Boston (MA): Springer.
- Liu, D., Li, Y., & Thomas, M. A.** (2017). A Roadmap for Natural Language Processing Research in Information Systems. *Proceedings of the 50th Hawaii International Conference on System Sciences*, 1112–1121. <https://doi.org/10.24251/HICSS.2017.132>
- Liu, H.** (2001). Creoles, Pidgins, and Planned Languages. Language Evolution under Special Conditions. In K. Schubert (Ed.), *Planned Languages: from Concept to Reality* (pp. 121–177). Brussel: Hogeschool voor Wetenschap en Kunst.
- Lo Jacomo, F.** (1981). *Liberté ou autorité dans l'évolution de l'espéranto* (Thèse de doctorat). Université Paris 5, René Descartes, Sciences humaines, Sorbonne, Paris.
- Lorente, M.** (2013). Terminología in vivo y variación funcional. Ugarteburu Terminologia Jardunaldiak (V). Terminologia naturala eta terminologia planifikatua euskararen normalizazioari begira. Présenté à Bizkaia Aretoan.

- Retrieved from <http://www.ehu.es/documents/2430735/2730483/LIBURUAehuei13-02.pdf>
- Loubier, C.** (1994). L'implantation du français : du microcosme au macrocosme. *Terminologies nouvelles*, 12, 20–39.
- Lucy, J. A.** (1993). Reflexive language and the human disciplines. In J. A. Lucy (Ed.), *Reflexive language : reported speech and metapragmatics* (pp. 9–32). Cambridge: University press.
- Lusch, R. F., & Vargo, S.** (Ed.). (2006). *The Service-dominant Logic of Marketing: Dialog, Debate, and Directions*. London & New York: Routledge.
- Lyons, J. E., Runge, M. C., Laskowski, H. P., & Kendall, W. L.** (2008). Monitoring in the Context of Structured Decision-Making and Adaptive Management. *Journal of Wildlife Management*, 72(8), 1683–1692.
- Macenery, T., & Hardie, A.** (2012). *Corpus linguistics: method, theory and practice*. Cambridge ; New York: Cambridge University Press.
- Magris, M.** (2009). Übersetzungsfehler und deren Auswirkungen in der Praxis. In M. Forstner, H. Lee-Jahnke, & P. A. Schmitt (Ed.), *CIUTI-Forum 2008: Enhancing Translation Quality : Ways, Means, Methods* (pp. 301–312). Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien: Peter Lang.
- Malovec, M.** (1999). *Gramatiko de Esperanto*. Retrieved from http://www.literatura.bucek.name/malovec/malovec_gramatiko-esp.pdf
- Manders, W.** (1976). Interlinguistik und Esperanto. In R. Haupenthal (Ed. and transl.), *Plansprachen. Beiträge zur Interlinguistik* (pp. 234–242). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Manning, C. D., & Schütze, H.** (1999). *Foundations of Statistical Natural Language Processing*. Cambridge (Massachusetts): MIT Press.
- Maradan, M., & Mueller, J.** (2012). Esperanto – auch für die Fachkommunikation? *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*, (2012–5), 46–50.
- Maras, T.** (2015). Die sprachliche Norm und ihre (In)Varianten als feste Grundlagen für die Evaluation von sprachlichen Ausdrücken. *Acta Universitatis Lodziensis. Folia Germanica*, (Acta Universitatis Lodziensis. Folia Germanica 11), 71–79.
- Marcellesi, C.** (1971). *Approche synchronique du vocabulaire de l'informatique (3^e génération)* (Thèse de III^e cycle). Université Paris-Nanterre (Paris X), Paris.
- Marpsat, M., & Razafindratsima, N.** (2010). Survey methods for hard-to-reach populations: introduction to the special issue. *Methodological Innovations Online*, 5(2), 3–16.

- Martin, A.** (1992). Théorie de la diffusion des innovations et implantation terminologique. *Terminologies nouvelles*, (7), 34–41.
- Martin, A.** (1994). Théorie de la diffusion des innovations et implantation terminologique (II). *Terminologies nouvelles*, (11), 33–39.
- Martin, A.** (1998). Les mots et leurs doubles : étude d'implantation de la terminologie officialisée dans le domaine de l'éducation au Québec. Gouvernement du Québec, Office de la langue française.
- Massion, F.** (2009). Terminologiemanagement: Luxus oder Muss? Von der Theorie zur Praxis. In F. Mayer & U. Seewald-Heeg (Ed.), *Terminologiemanagement. Von der Theorie zur Praxis* (pp. 27–34). Berlin: Bundesverband der Dolmetscher und Übersetzer e.V. (BDÜ).
- Mattos, G.** (1999). Esenco kaj estonteco de la Fundamento de Esperanto. *Esperantologio / Esperanto Studies*, 1, 21–37.
- Maurais, J.** (1993). Terminology and Language Planning. In H. Sonneveld & K. L. Loening (Ed.), *Terminology. Applications in interdisciplinary communication* (pp. 111–125). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Mayar, M.-P.** (2005). La néologie d'aménagement : le cas de la Communauté française de Belgique. In L. Depecker & V. Dubois (Ed.), *Les néologies contemporaines* (pp. 79–99). Paris: Société française de terminologie.
- Mayer, H.** (1987). La eterna neologisma diskuto: Kalocsay kaj la sekvo. Vieno: Krammer.
- Maying, P.** (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., aktual. und überarb. Aufl). Weinheim: Beltz.
- McEneary, T., & Wilson, A.** (1997). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- McIvor, O.** (2009). Strategies for Indigenous language revitalization and maintenance. In *Encyclopedia of Language and Literacy Development* (pp. 1–12). Retrieved from www.literacyencyclopedia.ca/pdfs/topic.php?topId=265
- McKelvie, A., Haynie, J. M., & Gustavsson, V.** (2011). Unpacking the uncertainty construct: Implications for entrepreneurial action. *Journal of Business Venturing*, (26), 273–292.
- Meier, S.** (2015). Gesprächskritik? Normative Gesprächsreflexion als Gegenstand der Kommunikationsgeschichte – am Beispiel der Gesprächsform *Diskussion*. In J. Bückler, E. Diedrichsen, & C. Spieß (Ed.), *Perspektiven linguistischer Sprachkritik* (pp. 101–126). Stuttgart: Ibidem-Verl.

- Mejri, S.** (2011). Néologie et unité lexicale : renouvellement théorique, polylexicalité et emploi. *Langages*, 3(183), 25–37.
- Mejri, S., & Sablayrolles, J.-F.** (2011). Présentation : Néologie, nouveaux modèles théoriques et NTIC. *Langages*, 3(183), 3–9.
- Melchior, L.** (2012). Halbkollaborativität und Online-Lexikographie. Ansätze und Überlegungen zu Wörterbuchredaktion und Wörterbuchforschung am Beispiel LEO Deutsch-Italienisch. *Lexicographica*, 28, 337–372.
- Mertz, E., & Yovel, J.** (2003). Metalinguistic awareness. In J.-O. Östman, J. Verschueren, J. Blommaert, & C. Bulcaen (Ed.), *Handbook of Pragmatics* (pp. 713). Amsterdam: Benjamins.
- Meunier, D., & Rosier, L.** (2014). Quand le savoir s’emmêle... La construction discursive de la norme chez les locuteurs non experts. *Les Carnets du Cediscor. Publication du Centre de recherches sur la didacticité des discours ordinaires*, (12), 99–113.
- Meyer, C. M., & Gurevych, I.** (2012). Wiktionary: A new rival for expert-built lexicons? Exploring the possibilities of collaborative lexicography. In S. Granger & M. Paquot (Ed.), *Electronic Lexicography* (1st edition, pp. 259–291). Oxford University Press.
- MGA.** (2011). *Pri lingvo kaj movado (debato)*. Retrieved from <http://www.youtube.com/watch?v=kniln-sGf-Y>
- Michel, J.-B., Shen, Y. K., Presser Aiden, A., Veres, A., Gray, M. K., The Google Books Team, ... Lieberman Aiden, E.** (2011). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science*, 331(6014), 176–182.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J.** (2013). *Efficient Estimation of Word Representations in Vector Space*. Retrieved from <https://arxiv.org/pdf/1301.3781.pdf>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J.** (2013). *Distributed Representations of Words and Phrases and their Compositionality*. Retrieved from <https://arxiv.org/abs/1310.4546>
- Milroy, J.** (1992). *Linguistic Variation and Change: On the Historical Sociolinguistics of English*. Oxford ; Cambridge Mass.: B. Blackwell.
- Milroy, J.** (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, 5(4), 530–555.
- Miner, K.** (2006). *La morfemefiko de Kalocsay-Schubert*. Retrieved from http://www.airbits.com/~miner/ME_package/me.html
- Miner, K.** (2011). The impossibility of an Esperanto linguistics / La neebleco de priesperanta lingvoscienco. *Interlingvistikaj Kajeroj*, 2(1), 26–51.

- Moch, G.** (1897). La question de la langue internationale et sa solution par l'Esperanto. Paris: V. Giard & E. Brière.
- Monnerot-Dumaine, M.** (1960). Précis d'interlinguistique générale et spéciale. Paris: Maloine.
- Montané March, M. A.** (2012). Terminologia i implantació : anàlisi d'alguns factors que influencien l'ús dels termes normalitzats de la informàtica i les TIC en llengua catalana (Tesi doctoral UPF, Universitat Pompeu Fabra). Retrieved from <http://www.tdx.cat/bitstream/handle/10803/97295/tamm.pdf?sequence=1>
- Moore, R.** (2015). *Citizen Sociolinguistics: The Return of the Repressed [Tillburg Papers in Culture Studies 148]*. Retrieved from https://www.tilburguniversity.edu/upload/9d58dc0e-f768-4292-9f6e-82023e6dd16b_TPCS_148_Moore.pdf
- Morales, H. L.** (2014). Lexical Availability Studies. In *Educational Linguistics: Vol. 14. Lexical Availability in English and Spanish as a Second Language* (pp. 1–11). https://doi.org/10.1007/978-94-007-7158-1_1
- Morgan, D. L.** (1998a). *Planning Focus Groups*. Thousand Oaks: SAGE Publications.
- Morgan, D. L.** (1998b). *The Focus Group Guidebook*. Thousand Oaks: SAGE Publications.
- Morsly, D.** (1990). Attitudes et représentations linguistiques. *La Linguistique*, 26(2), 77–86.
- Mortureux, M.-F.** (1987). Les résistances à la néologie terminologique : système lexical et facteurs socioculturels. *Meta : journal des traducteurs / Meta: Translators' Journal*, 32(3), 250–254.
- Mortureux, M.-F.** (1993). Paradigmes désignationnels. *Semen. Revue de sémiolinguistique des textes et discours*, (8).
- Moser, H.** (1967). Sprache – Freiheit oder Lenkung? Zum Verhältnis von Sprachnorm, Sprachwandel, Sprachpflege. Mannheim: Bibliographische Institut.
- Muñoz Martín, F. J., & Valdivieso Blanco, M.** (2006). La importación lingüística en una relación asimétrica Español e Inglés, dos socios desiguales. *Panace@: Revista de Medicina, Lenguaje y Traducción*, 7(24), 285–296.
- Munske, H. H.** (2015). Lexikologie. Wortbildung. Orthographie. Ausgewählte sprachwissenschaftliche Schriften (1970–2015) (J. Lee, Ed.). Erlangen: FAU University Press.

- Murano, M.** (2014). La lexicographie 2.0 : nous sommes tous lexicographes ?
In *Docteurs et recherche...une aventure qui continue* (pp. 147–162). Trieste: Edizioni Università di Trieste.
- Nahir, M.** (1977). The Five Aspects of Language Planning – A classification.
Language Problems & Language Planning, 1(2), 107–123.
- Nahir, M.** (1998). Micro language planning and the revival of Hebrew: A schematic framework. *Language in Society*, (27), 335–357.
- Nahir, M.** (2002). Corpus planning and codification in the Hebrew Revival.
Language Problems and Language Planning, 26(3), 271–298.
- Nakache, D., Metais, E., & Timsit, J. F.** (2005). Evaluation and NLP.
Database and Expert Systems Applications, 417–422.
https://doi.org/10.1007/11546924_61
- Neergaard, P.** (1955). Polemiko: Rimarkoj pri kelkaj terminologiaj principoj.
Studo kun speciala konsidero al biontologio kaj Noto de la redaktoro.
Scienca Revuo, 7(26), 43–58.
- Nekvapil, J.** (2011). The history and theory of language planning. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (Vol. 2, p. 871–887). Retrieved from https://www.researchgate.net/publication/295912768_The_history_and_theory_of_language_planning
- Nesi, H.** (2012a). Alternative e-dictionaries: uncovering dark practices. In S. Granger & M. Paquot (Ed.), *Electronic Lexicography* (pp. 363–378). Oxford: Oxford University Press.
- Nesi, H.** (2012b). Alternative e-dictionaries: Uncovering dark practices. In S. Granger & M. Paquot (Ed.), *Electronic Lexicography* (pp. 363–378). Oxford: Oxford University Press.
- Nettle, D.** (1999). Using Social Impact Theory to simulate language change.
Lingua, (108), 95–117.
- Neuland, E.** (1993). Sprachgefühl, Spracheinstellungen, Sprachbewußtsein. Zur Relevanz « subjektiver Faktoren » für Sprachvariation und Sprachwandel. In K. J. Mattheier, K.-P. Wegera, W. Hoffmann, J. Macha, & H.-J. Solms (Ed.), *Vielfalt des Deutschen. Festschrift für Werner Besch* (pp. 723–748). Frankfurt am Main, Berlin, Bern, New York, Paris, Wien: Peter Lang.
- Ní Ghearáin, H.** (2011). The problematic relationship between institutionalised Irish terminology development and the Gaeltacht speech community: dynamics of acceptance and estrangement. *Language Policy*, 10, 305–323.
- Nic Phóidín, C., Ó Cleircín, G., & Bhreathnach, Ú.** (2010). Building on a terminology resource – the Irish experience. In A. Dykstra & T. Schoonheim

- (Ed.), *Proceedings of the XIV Euralex International Congress* (pp. 954-965). Ljouwert: Fryske Akademy.
- Nida, E. A.** (2001). Dynamic Equivalence in Translating. In C. Sin-Wai & D. E. Pollard (Ed.), *An Encyclopedia of Translation. Chinese-English. English-Chinese* (pp. 223–230). Hong Kong: The Chinese University Press.
- Niedzielski, N., & Preston, D. R.** (Ed.). (1999). *Folk Linguistics*. Berlin: Mouton de Gruyter.
- Niiniluoto, I.** (1993). The aim and structure of applied research. *Erkenntnis*, 38(1), 1–21.
- Nissilä, N., & Pilke, N.** (2017). Rättsvans är ej så lyckligt” Terminologisk förankring genom TNCs rundfrågor till sakkunniga. In *Svenskans beskrivning* (pp. 243–254). Göteborg: Institutionen för svenska språket.
- Nogué Pich, M., & Vila i Moreno, F. X.** (2007a). Buscant, encara, la via idònia? La implantació dels neologismes terminològics en el camp de l'escalada esportiva. In F. X. Vila i Moreno, M. Nogué Pich, & I. Vila i Moreno (Ed.), *Estudis d'implantació terminològica: una aproximació en l'àmbit dels esports* (pp. 91–159). Vic : Barcelona: Eumo Editorial ; Termcat.
- Nogué Pich, M., & Vila i Moreno, F. X.** (2007b). Entre *el hockey i l'hoquei*. La implantació dels neologismes terminològics en el camp de l'hoquei sobre herba. In F. X. Vila i Moreno, M. Nogué Pich, & I. Vila i Moreno (Ed.), *Estudis d'implantació terminològica: una aproximació en l'àmbit dels esports* (pp. 161–240). Vic : Barcelona: Eumo Editorial ; Termcat.
- Nord, B.** (2002). Hilfsmittel beim Übersetzen. Eine empirische Studie zum Rechercheverhalten professioneller Übersetzer. Frankfurt: Peter Lang.
- Normantas, V.** (2013). Sources of text for Esperanto corpora. *Esperantologio / Esperanto Studies*, 6.
- Obreja, C.** (2012). Imaginaire linguistique et dynamique de la langue dans le discours de la presse écrite. *ANADISS*, 13.
- Oljanov, V.** (Ed.). (1985). Lingvaj disputoj kaj lingvo-evoluo. In *Interlingvistiko kaj esperantologio* (pp. 91–96). Sofia: Bulgara Esperantista Asocio.
- Olshavsky, R., & Wymer, W.** (1995). The Desire for New Information From External Sources. In S. MacKenzie & R. Stayman (Ed.), *Proceedings of the Society for Consumer Psychology* (pp. 17–27). Bloomington (IN): Printmaster.
- Osthus, D.** (2003). Le bon usage d'Internet – Discours et conscience normatifs dans des débats virtuels. In D. Osthus, C. Polzin-Haumann, & C. Schmitt (Ed.), *La norme linguistique. Théorie – pratique – médias – enseignement*.

- Actes du colloque tenu à Bonn le 6 et le 7 décembre 2002* (pp. 139–152). Bonn: Romanistischer Verlag.
- Otman, G.** (1991). Quelques aspects du métier de terminologue. *Cahiers de linguistique sociale*, (18), 159–180.
- Ottum, B. D., & Moore, W. L.** (1997). The role of market information in new product success/failure. *Elsevier Journal of Product Innovation Management*, 14(4), 258–273.
- Pabst, B.** (2014). Das Fundamento als Maßstab sprachlicher Richtigkeit im Esperanto. In C. Brosch & S. Fiedler (Ed.), *Interlinguistik im 21. Jahrhundert. Beiträge der 23. Jahrestagung der Gesellschaft für Interlinguistik e. V.* (pp. 45–60). Berlin.
- Parkvall, M.** (2010). How European is Esperanto? A typological study. *Language Problems & Language Planning*, 34(1), 63–79.
- Patent-De.** (2008). Hubgesteuerte Scheibenwischanlage für Kraftfahrzeuge – Dokument DE3742635C2. Accessed on 2017/07/19: <http://www.patent-de.com/19891019/DE3742635C2.html>
- Paveau, M.-A.** (2011). Do non-linguists practice linguistics? *AILA Review*, 24 (*Applied Folk Linguistics*), 40–54.
- Paveau, M.-A., & Rosier, L.** (2008). *La langue française. Passions et polémiques*. Paris: Vuibert.
- Peek, L., & Fothergill, A.** (2009). Using focus groups: lessons from studying daycare centers, 9–11, and Hurricane Katrina. *Qualitative Research*, 9(1), 31–59.
- Peter, P. J., & Olson, J. C.** (2010). *Consumer Behavior & Marketing Strategy* (9th ed.). Irwin: McGraw-Hill.
- Petersen, A. M., Tenenbaum, J., Havlin, S., & Stanley, H. E.** (2012). Statistical Laws Governing Fluctuations in Word Use from Word Birth to Word Death. *Scientific Reports*, 2.
- Peterson, R. A., & Merino, M. C.** (2003). Consumer Information Search Behavior and the Internet. *Psychology & Marketing*, 20(2), 99–121.
- Philippe, B.** (1991). Sprachwandel bei einer Plansprache am Beispiel des Esperanto. Universität Konstanz, Konstanz.
- Picone, M. D.** (1996). *Anglicisms, neologisms and dynamic French* (Vol. 18). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Picton, A.** (2009). Diachronie en langue de spécialité. Définition d'une méthode linguistique outillée pour repérer l'évolution des connaissances en corpus. Un exemple appliqué au domaine spatial (Thèse en Sciences du langage). Université de Toulouse, Toulouse.

- Picton, A.** (2014). Métalangage et sentiment linguistique des experts : regard en langue de spécialité. *Le discours et la langue*, 6(1), 77–90.
- Piron, C.** (1989). Who are the speakers of Esperanto? In *Trends in Linguistics. Studies and Monographs: Vol. 42. Interlinguistics. Aspects of the Science of Planned Languages* (pp. 157–172). Berlin, New York: Mouton de Gruyter.
- Polle, F.** (1889). Wie denkt das Volk über die Sprache?: Gemeinverständliche Beiträge zur Beantwortung dieser Frage. Leipzig: Teubner.
- Pool, J., & Grofman, B.** (1989). Linguistic artificiality and cognitive competence. In K. Schubert & D. Maxwell (Ed.), *Interlinguistics. Aspects of the Science of Planned Languages* (pp. 145–156). Berlin, New York: Mouton de Gruyter.
- Preston, D. R.** (2004). Folk metalanguage. In A. Jaworski, N. Coupland, & D. Galasiński (Ed.), *Metalanguage: social and ideological perspectives* (pp. 75–101). Berlin: Mouton de Gruyter.
- Preston, D. R.** (2011). Methods in (applied) folk linguistics. Getting into the mind of the folk. *AILA Review*, 24 (*Applied Folk Linguistics*), 15–39.
- Privat, E.** (1912). *Historio de la lingvo Esperanto. Unua parto. 1887–1900* (Vol. 2). Leipzig: Universala Esperantia Librejo.
- Privat, E.** (1927). *Historio de la lingvo Esperanto. Dua parto. La Movado 1900–1927* (Vol. 2). Leipzig: Ferdinand Hirt & Sohn.
- Privat, E.** (2001). *Esprimo de sentoj en Esperanto. Resumo de kurso farita ĉe la Universitato de Ĝenevo, en la lingvosciencia fakultato dum 1929 kaj 1930* (F. Luin, Ed.). Tyresö: Inko.
- Pruvost, J., & Sablayrolles, J.-F.** (2003). *Les néologismes* (1^{re} édition). Paris: Presses universitaires de France.
- Prys, D., & Jones, D. B.** (2007). *Guidelines for the Standardization of Terminology for the Welsh Assembly*. Retrieved from <http://techiath.bangor.ac.uk/wp-content/uploads/2016/01/Guidelines-for-Standardization.pdf>
- Public Works and Government Services Canada.** (2016). « Crash Cushion ». In *Terminium Plus*. Retrieved from http://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-fra.html?lang=fra&i=1&srchtxt=crash+cushion&index=alt&context2nd_wet=1#resultrecs.
- Puškar, K.** (2015). Esperanto (s)en perspektivo? Croatian Esperantists on the International Language Esperanto. *Interdisciplinary Description of Complex System*, 13(2), 322–341.
- Quemada, B.** (1971). À propos de la néologie. Essai de délimitation des objectifs et des moyens d'action. *La Banque des mots*, 1, 137–150.

- Quirion, J.** (2003a). La mesure de l'implantation terminologique : proposition d'un protocole. Étude terminométrique du domaine des transports au Québec. Office québécois de la langue française.
- Quirion, J.** (2003b). Methodology for the design of a standard research protocol for measuring terminology usage. *Terminology*, 9(1), 29–49.
- Quirion, J.** (2004). État de la question sur la nature des facteurs d'implantation terminologique. *Actes du Colloque international sur la traduction : Traduction et Francophonie*, 193–200. Retrieved from http://www.colloque.net/archives/2003/volume_1/quirion.pdf
- Quirion, J.** (2005). L'automatisation de la terminométrie : premiers résultats. *Actes du 5^e Congrès international : Langue et terminologie helléniques*, 61–70. Nicosie: Techniko Epimelitirio Elladas.
- Quirion, J.** (2006). La localisation, palimpseste de l'aménagement terminologique ? Stratégies d'implantation terminologique et marketing. *Meta : journal des traducteurs / Meta: Translators' Journal*, 51(4), 824–837.
- Quirion, J.** (2010). Les résultats des études d'implantation, d'hier à aujourd'hui. In L. Depecker & V. Dubois (Ed.), *L'évaluation des politiques linguistiques* (pp. 69–80). Paris: Société française de terminologie.
- Quirion, J.** (2012). Néologie traductive, néologie aménagiste et néologie collaborative massive : l'unité dans la disparité. *Neologica*, (6), 129–141.
- Quirion, J.** (2014a). La mesure de la variation terminologique comme indice de l'évolution des connaissances dans un environnement bilingue. In R. Temmerman & M. Van Campenhoudt (Ed.), *Dynamics and Terminology: An interdisciplinary perspective on monolingual and multilingual culture-bound communication* (pp. 281–302). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Quirion, J.** (2014b). Pour une néologie sociale. In P. Dury, J. C. de Hoyos, J. Makri-Morel, F. Maniez, V. Renner, & M. Belén Villar Díaz (Ed.), *La néologie en langue de spécialité : détection, implantation et circulation des nouveaux termes* (pp. 101–120). Lyon: Publications du CRTT.
- Rampton, B.** (2010). Speech Community. In J. Jaspers, J. Verschueren, & J.-O. Östman (Ed.), *Society and Language Use* (pp. 274–303). Amsterdam: John Benjamins.
- Ramsenthaler, C.** (2013). Was ist « Qualitative Inhaltsanalyse? » In H. Kolbe & C. Dunger (Ed.), *Der Patient am Lebensende. Eine qualitative Inhaltsanalyse* (pp. 23–42).
- Rašić, N.** (1994). La rondo familia. Sociologiaj esploroj en Esperantio. Pisa: Edistudio.

- Rat für Deutschsprachige Terminologie.** (2004). *Professional profile for terminologists*. Retrieved from www.iim.fh-koeln.de/radt/Dokumente/RaDT_Berufsprofil_englisch.pdf
- Ratchford, B. T., Talukdar, D., & Myung-So, L.** (2001). A Model of Consumer Choice of the Internet as an Information Source. *International Journal of Electronic Commerce*, 5(3), 7–21.
- Read, S., Dew, N., Sarasvathy, S. D., Song, M., & Wiltbank, R.** (2009). Marketing Under Uncertainty: The Logic of an Effectual Approach. *Journal of Marketing*, 73(3), 1–18.
- Reboul-Touré, S.** (2004). Écrire la vulgarisation scientifique aujourd’hui. *Sciences Médias et Société*, 195–212. Lyon.
- Reid, D. J., & Reid, F. J. M.** (2005). Online focus groups. An in-depth comparison of computer-mediated and conventional focus group discussions. *International Journal of Market Research*, 47(2), 131–162.
- Remysen, W.** (2009). Description et évaluation de l’usage canadien dans les chroniques de langage : contribution à l’étude de l’imaginaire linguistique des chroniqueurs canadiens-français (Thèse de doctorat en linguistique). Université Laval, Québec.
- Remysen, W.** (2011). L’application du modèle de l’Imaginaire linguistique à des corpus écrits : le cas des chroniques de langage dans la presse québécoise. *langage & société*, 135(1), 47–65.
- Rey-Debove, J.** (1971). Étude linguistique et sémiotique des dictionnaires français contemporains. The Hague, Paris: Mouton.
- Rey-Debove, J.** (1978). Le métalangage. Étude linguistique du discours sur le langage. Paris: Le Robert.
- Rey-Debove, J.** (1979). Les logiciens et le métalangage naturel. *Histoire Épistémologie Langage*, 1(1), 15–22.
- Rey-Debove, J.** (1985). Le métalangage en perspective. *Revue de linguistique DRLAV*, (32 Métalangage, métadiscours, métacommunication), 21–32.
- Reyes, A., & Bonnin, J. E.** (2016). Negotiating use, norm and authority in online language forums. *Current Issues in Language Planning*, 1–25. <https://doi.org/10.1080/14664208.2016.1220280>
- Rezabek, R. J.** (2000). Online Focus Groups: Electronic Discussions for Research. *Forum: Qualitative Social Research*, 1(1), Art. 18.
- Rheault, A.-H.** (2004). De la description à la prescription analyse de la chronique «Corrigeons-nous» de la Société du parler français au Canada (1930–1945) (Mémoire présenté pour l’obtention du grade de maîtrise ès

- arts, Université de Sherbrooke). Retrieved from <http://savoirs.usherbrooke.ca/handle/11143/2390>
- Rheault, A.-H.** (2010). Analyse argumentative du discours épilinguistique au Québec: les lieux communs comme indicateurs de normes. Université de Sherbrooke/Université catholique de Louvain, Sherbrooke.
- Ricento, T.** (2000). Historical and theoretical perspectives in language policy and planning. *Journal of Sociolinguistics*, 4(2), 196–213.
- Ripfel, M.** (1990). Wörterbuchbenutzung bei Muttersprachlern. Untersuchungsbericht über eine Befragung erwachsener muttersprachlicher Sprecher zur Wörterbuchbenutzung. *Lexicographica*, 6, 237–251.
- Ritchie, J., & Lewis, J.** (2003). *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. SAGE.
- Rodríguez Penagos, C.** (2004a). *Metalinguistic Information Extraction for Terminology* (S. Ananiadou & P. Zweigenbaum, Ed.). Geneva.
- Rodríguez Penagos, C.** (2004b). *Metalinguistic Information Extraction from Specialized Texts to Enrich Computational Lexicons* (Tesi Doctoral). Universitat Pompeu Fabra, Barcelona.
- Rodríguez Penagos, C.** (2005). Explotación computacional del metalenguaje en corpus especializados para la generación de lexicones no convencionales. *Procesamiento del Lenguaje Natural*, (35), 137–144.
- Roelcke, T.** (2007). Effizienz sprachlicher Kommunikation. In J. A. Bär, T. Roelcke, & A. Steinhauer (Ed.), *Sprachliche Kürze. Konzeptuelle, strukturelle und pragmatische Aspekte* (pp. 8–26). Berlin: De Gruyter.
- Rogers, E. M.** (1983). *Diffusion of Innovations* (3rd édition). The Free Press, Collier Macmillan Publishers.
- Rondeau, G.** (1984). *Introduction à la terminologie* (2^e édition). Chicoutimi.
- Roser, T., Defillippi, R., & Samson, A.** (2012). Managing your co-creation mix: Co-creation ventures in distinctive contexts. *European Business Review*, 25(1).
- Rousseau, L.-J.** (2007). Élaboration et mise en œuvre des politiques linguistiques. *Cahiers du Rifal*, (26), 58–71.
- Sablayrolles, J.-F.** (1996). Néologismes: une typologie des typologies. *Cahiers du CIEL 1996-1997*, 11–48.
- Sablayrolles, J.-F.** (2000). La néologie en français contemporain. Examen du concept et analyse de productions néologiques récentes. Paris: H. Champion.
- Sablayrolles, J.-F., Jacquet-Pfau, C., & Humbley, J.** (2011). Emprunts, créations « sous influence » et équivalents. In T. Lino & R. Costa (Ed.), *Passeurs de mots, passeurs d'espoir. Lexicologie, terminologie et traduction*

- face au défi de la diversité* (pp. 325–339). Paris: Editions des Archives Contemporaines.
- Sager, J. C.** (1997). Term Formation. In S. E. Wright & G. Budin (Ed.), *Handbook of terminology management* (pp. 25–41). John Benjamins Publishing Company.
- Saint, E. C.** (2013a). La terminologie de référence au service de la langue française au Québec et son usage dans le domaine de l’informatique. *44*(2), 167–198.
- Saint, E. C.** (2013b). Les attitudes à l’égard de l’emprunt à l’anglais au Québec et en France : Le cas du domaine informatique. *Communication, lettres et sciences du langage*, *7*(1), 87–101.
- Saint, E. C.** (2016). Ce cuisinoman, il est fou, dis !? – Opinions sur un néologisme et anticipation de son implantation dans l’usage. *Cuizine*, *7*(1). Retrieved from <https://www.erudit.org/en/journals/cuizine/2016-v7-n1-cuizine02655/1037391ar/>
- Sakaguchi, A.** (1998). Interlinguistik. Gegenstand, Ziele, Aufgaben, Methoden. Frankfurt: Peter Lang.
- Samain, D.** (2010). Eugen Wüster. De l’esperanto à la terminologie. In M. Heinz (Ed.), *Cultures et lexicographie* (pp. 279–296). Berlin: Frank & Timme.
- Samuel, J.** (2005). Modernisation lexicale et politique terminologique : le cas de l’indonésien. Paris ; Louvain: Peeters.
- Sanz Vicente, L.** (2012). Approaching secondary term formation through the analysis of multiword units. *Terminology*, *18*(1), 105–127.
- Schapiro, R. E., & Singer, Y.** (2000). BoosTexter: A Boosting-based System for Text Categorization. *Machine Learning*, *39*(2/3), 135–168.
- Schmidt, J. B., & Spreng, R. A.** (1996). A Proposed Model of External Consumer Information Search. *Journal of the Academy of Marketing Science*, *24*(3), 246–256.
- Schmitz, K.-D., & Straub, D.** (2010). Erfolgreiches Terminologiemanagement im Unternehmen. Praxishilfe und Leitfaden: Grundlagen, Umsetzung, Kosten-Nutzen-Analyse, Systemübersicht. Stuttgart: TC and more GmbH.
- Schreier, M.** (2012). *Qualitative content analysis in practice*. Los Angeles: Sage.
- Schubert, K.** (1989). An unplanned development in planned languages. A study of word grammar. In *Interlinguistics. Aspects of the Science of Planned Languages Trends in Linguistics. Studies and Monographs: Vol. 42* (pp. 249–274). Berlin, New York: Mouton de Gruyter.
- Schubert, K.** (1993). Semantic compositionality: Esperanto word formation for language technology. *Linguistics*, *31*(2), 311–365.

- Schubert, K.** (2009). Kommunikationsoptimierung: Vorüberlegungen zu einem fachkommunikativen Forschungsfeld. *trans-kom*, 2(1), 109–150.
- Schubert, K.** (2010). Kiam Zamenhof perdis sian bastonon. Implicitaj premisoj en planlingvo. In D. Blanke & U. Lins (Ed.), *La arto labori kune* (pp. 355–362). Rotterdam: Universala Esperanto-Asocio.
- Schubert, K.** (2011). Zum bewussten Eingreifen in die Sprache. In Brosch C., & Fiedler, S. (Ed.), *«Florilegium Interlinguisticum»: Festschrift für Detlev Blanke zum 70. Geburtstag* (pp. 47–60). Frankfurt am Main: Peter Lang.
- Schubert, K.** (2014). Barrierefrei, reguliert, gelenkt. Prinzipien optimierenden Eingreifens in Sprache und Kommunikation. In S. J. Jekat, H. E. Jüngst, K. Schubert, & C. Villiger (Ed.), *Sprache barrierefrei gestalten. Perspektiven aus der Angewandten Linguistik* (pp. 201–220). Berlin: Frank & Timme.
- Schubert, K.** (2015). Word-formation and planned languages. In P. O. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Ed.), *Word-Formation. An International Handbook of the Languages of Europe* (Vol. 3, pp. 2210–2224). Berlin, Boston: De Gruyter Mouton.
- Schubert, K.** (2018). Apriorische und aposteriorische Plansprachen – eine Quellenrecherche. In *Jahrbuch der Gesellschaft für Interlinguistik 2018* (pp. 105–132). Retrieved from <http://www.interlinguistik-gil.de/wb/media/beihefte/JGI2018/JGI2018-Schubert.pdf>
- Schütte, C.** (2015). Zur Funktion von Emotionen in sprachkritischer Argumentation: Eine Analyse von Beiträgen aus der „deutschen Sprachwelt“. In J. Bückler, E. Diedrichsen, & C. Spieß (Ed.), *Perspektiven linguistischer Sprachkritik* (pp. 47–74). Stuttgart: Ibidem-Verl.
- Schwartz, H. A., & Ungar, L. H.** (2015). Data-Driven Content Analysis of Social Media: A Systematic Overview of Automated Methods. *The ANNALS of the American Academy of Political and Social Science*, 659(1), 78–94. <https://doi.org/10.1177/0002716215569197>
- Schweder, S.** (1999). *Entwicklung und Akzeptanz technischer Terminologien in der Plansprache Esperanto* (Diplomarbeit). Universität Hildesheim, Fachbereich III – Sprachen und Technik, Institut für Angewandte Sprachwissenschaft, Hildesheim.
- Seidlhofer, B.** (2016). ELF: English in a global context. In K. Murata (Ed.), *Exploring ELF in Japanese Academic and Business Contexts. Conceptualization, research and pedagogic implications* (pp. 17–28). London & New York: Routledge.

- Selten, R.** (2002). What Is Bounded Rationality. In G. Gigerenzer & R. Selten (Ed.), *Bounded Rationality: The Adaptive Toolbox* (pp. 13–36). Cambridge, MA: The MIT Press.
- Shamim, A., & Ghazali, Z.** (2014). A Conceptual Model for Developing Customer Value Co-Creation Behaviour in Retailing. *Global Business and Management Research: An International Journal*, 6(3), 185–196.
- Shanklin, W. L., & Ryans, J. K.** (1984). Organizing for High-Tech Marketing. *Harvard Business Review*. Retrieved from <https://hbr.org/1984/11/organizing-for-high-tech-marketing>
- Sherwood, B. A.** (1982). Variation in Esperanto. *Studies in the Linguistic Sciences*, 12, 183–196.
- Shin, H. B., & Rosalind, B.** (2000). *Language Use and English-Speaking Ability: 2000. Census 2000 Brief*. Retrieved from <http://files.eric.ed.gov/fulltext/ED482418.pdf>
- Siemund, P., Gogolin, I., Schulz, M. E., & Davydova, J.** (Ed.). (2013). *Multilingualism and Language Diversity in Urban Areas. Acquisition, identities, space, education*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Sikosek, M.** (2006). Die neutrale Sprache: eine politische Geschichte des Esperanto-Weltbundes. Universiteit Utrecht, Utrecht.
- Silverman, D.** (1998). Qualitative research: meanings or practices? *Information Systems Journal*, 8(1), 3–20. <https://doi.org/10.1046/j.1365-2575.1998.00002.x>
- Sim, J.** (1998). Collecting and analysing qualitative data: issues raised by the focus group. *Journal of Advanced Nursing*, 28(2), 345–352.
- Sinclair, J.** (1998). *The Lexical Item* (E. Weigand, Ed.). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Sitri, F.** (2003). L'autonymie dans la construction des objets de discours. In J. Authier-Revuz, M. Doury, & S. Reboul-Touré (Ed.), *Parler des mots. Le fait autonymique en discours* (pp. 205–216). Paris: Presses Sorbonne Nouvelle.
- Slodzian, M.** (1995). Transfert de connaissances nouvelles et aménagement terminologique. *Meta : journal des traducteurs / Meta: Translators' Journal*, 40(2), 238–243.
- Slodzian, M.** (2000). L'émergence d'une terminologie textuelle et le retour du sens. In H. Béjoint & P. Thoiron (Ed.), *Le sens en terminologie* (pp. 61–85). Presses universitaires de Lyon.
- Slodzian, M.** (2005). Rationalisation des langues : l'état de la question entre les années 1930 et 1950. In H. Béjoint & F. Maniez (Ed.), *De la mesure dans les*

- termes: *hommage à Philippe Thoiron* (pp. 227–240). Lyon: Presses universitaires de Lyon.
- Sokolova, M., & Lapalme, G.** (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, (45), 427–437.
- Song, Y., Sahoo, N., Srinivasan, S., & Dellarocas, C.** (2016). *Uncovering Characteristic Paths to Purchase of Consumers* (SSRN Scholarly Paper N° ID 2619674). Retrieved from Social Science Research Network website: <https://papers.ssrn.com/abstract=2619674>
- Soubrier, J.** (1998). Néologismes et termes d'emprunt dans le français des affaires. In *Actualité scientifique. Actes du Colloque La mémoire des mots* (pp. 403–418). AUPELF-UREF.
- Souza, M., Vieira, R., Buseti, D., Chishman, R., & Alves, I. M.** (2011). Construction of a Portuguese Opinion Lexicon from multiple resources. *Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology*, 59–66. Cuiabá.
- Sparks, B., & Pan, G. W.** (2009). Chinese Outbound tourists: Understanding their attitudes, constraints and use of information sources. *Tourism Management*, 30(4), 483–494.
- Speer, S. A.** (2002). “Natural” and “contrived” data: a sustainable distinction? *Discourse Studies*, 4(4), 511–525.
- Srinivasan, N.** (2012). Pre-Purchase External Search for Information. In V. A. Zeithaml (Ed.), *Review of Marketing 1990* (pp. 153–189). Marketing Classics Press.
- St. Clair, R.** (1978). Universals of Word Order in Esperanto. *Lektos*, III(2), 31–39.
- Stegu, M.** (2008). Linguistique populaire, language awareness, linguistique appliquée : interrelations et transitions. *Pratiques*, 81–92.
- Steinmüller, U.** (1978). Sprachveränderung – Politisch motiviert. Ein Beitrag zur Begriffsklärung im Bereich der Theorien sprachlichen Wandels und zu einigen Aspekten der Sprachlenkung. *Linguistische Arbeiten und Berichte*, (10), 1–47.
- Stewart, D. W., Shamdasani, P. N., & Rook, D. W.** (2007). *Focus Groups. Theory and Practice*. Thousand Oaks: SAGE Publications.
- Stewart, K., & Williams, M.** (2005). Researching online populations: the use of online focus groups for social research. *Qualitative Research*, 5(4), 395–416. <https://doi.org/10/15654.s>

- Stocker, F.** (1996). *Wer spricht Esperanto? Kiu parolas Esperanton?* (M. Haase, Ed.). München, Newcastle: Lincom Europa.
- Sturz, W.** (2014). Gut gerüstet. *Technische Kommunikation*, (5), 37–40.
- Suchowolec, K.** (2018). Sprachlenkung – Aspekte einer übergreifenden Theorie. Frank & Timme, Berlin.
- Šumarina, M. P.** (2014). Народная лексикография — феномен современной российской лингвокультуры [Folk lexicography is a phenomenon of modern russian linguistic culture]. *Проблемы истории, филологии, культуры. Научный рецензируемый журнал [Journal of historical, philological and cultural studies]*, 3(45), 291–293.
- Suonuuti, H.** (1998). *Terminologia Gvidilo* (W. Blanke, Ed.; S. Fiedler, Transl.). Rotterdam: Universala Esperanto-Asocio.
- Suonuuti, H.** (2001). *Guide to Terminology* (Revised edition). Finnish Centre for Technical Terminology / Nordterm.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M.** (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37(2), 267–307.
- Tamba, I.** (2003). Autonymie, dénomination et fonction métalinguistique : Quelques remarques. In J. Authier-Revuz, M. Doury, & S. Reboul-Touré (Ed.), *Parler des mots. Le fait autonymique en discours* (pp. 59–66). Paris: Presses Sorbonne Nouvelle.
- Tarp, S.** (2008). *Lexicography in the Borderland between Knowledge and Non-Knowledge*. Max Niemeyer Verlag.
- Tarp, S.** (2009). Reflections on lexicographical user research. *Lexikos*, 19, 275–296.
- Tates, K., Zwaanswijk, M., Otten, R., van Dulmen, S., Hoogerbrugge, P. M., Kamps, W. A., & Bensing, J. M.** (2009). Online focus groups as a tool to collect data in hard-to-include populations: examples from paediatric oncology. *BMC Medical Research Methodology*, 9(15).
- Tauli, V.** (1968). *Introduction to a theory of language planning*. Uppsala: Almqvist & Wiksell.
- Temmerman, R.** (2000). *Towards New Ways of Terminology Description. The Sociocognitive Approach* (H. Sonneveld & S. E. Wright, Ed.). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- TERMCAT, Centre de Terminologia.** (2006). *La normalització terminològica en català: criteris i termes 1986–2004* (primera edició). Barcelona: Publicacions de l'Abadia de Montserrat.

- The Unicode Consortium.** (n. d.). Unicode® 1.0. Accessed on 2018/4/8:
<https://www.unicode.org/versions/Unicode1.0.0/>
- Thier, K.** (2014). Das *Oxford English Dictionary* und seine Nutzer. *Online publizierte Arbeiten zur Linguistik (OPAL)*, 4(Der Nutzerbeitrag im Wörterbuchprozess. 3. Arbeitsbericht des wissenschaftlichen Netzwerk „Internetlexikographie“), 63–70.
- Thoiron, P., Iwaz, J., & Zaouche, N.** (1993). Étude d’implantation des arrêtés de terminologie. Domaines : santé et médecine. In L. Depecker (Ed.), *La mesure des mots. Cinq études d’implantation terminologique*. (pp. 47–95). Rouen: Publications de l’Université de Rouen.
- TNS Opinion & Social.** (2012, June). Special Eurobarometer 386. Europeans and their languages. Report of a survey conducted at the request of Directorate-General Education and Culture, Directorate-General for Translation and Directorate-General for Interpretation. Retrieved from https://web.archive.org/web/20150207220823/http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf.
- Tonkin, H.** (2007). Recent Studies in Esperanto and Interlinguistics: 2006. *Language Problems & Language Planning*, 31(2), 169–196.
- Tonkin, H.** (2015). Introduction: in Search of Esperanto. *Interdisciplinary Description of Complex Systems*, 13(2), 182–192.
- Tonkin, H., & Fettes, M.** (1996). *Esperanto Studies: An Overview*.
- Tournier, J.** (2007). Introduction descriptive à la lexicogénétique de l’anglais contemporain. Genève: Éditions Slatkine.
- Tran, M. K.** (2017). Four decades of engaging customers in product innovation. In E. Shiu (Ed.), *Research Handbook of Innovation and Creativity for Marketing Management* (pp. 91–121). Cheltenham (UK): Edward Elgar.
- Triano López, M.** (2007). The attitudinal factor in language planning: the Valencian situation. München: LINCOM Europa.
- Triano-López, M.** (2007). Languages attitudes and the lexical de-Castilianization of Valencian. Implications for language planning. In K. Potowski & R. Cameron (Ed.), *Spanish in Contact. Policy, Social and Linguistic Inquiries*.
- Tsekos, N.** (1996). Discours épilinguistique et construction identitaire : l’Imaginaire linguistique des locuteurs d’Athènes. *Travaux de Linguistique*, (7), 27–35.
- Tunmer, W. E., Pratt, C., & Herriman, M. L.** (Ed.). (1984). *Metalinguistic Awareness in Children : theory, research, and implications*. Berlin, Heidelberg: Springer Verlag.

- Valeontis, K., & Mantzari, E.** (2006). The linguistic dimension of terminology: principles and methods of term formation. *Translation: Between Art and Social Science*. Presentation at the 1st Athens International Conference on Translation and Interpretation. Retrieved from http://translation.hau.gr/telamon/files/HAU-speechValeontisMantzari_EN.pdf
- van der Sluijs, J.** (2005). Uncertainty as a monster in the science-policy interface: four coping strategies. *Water Science & Technology*, 52(6), 87–92.
- Vandermeeren, S.** (2008). Research on Language Attitudes. In U. Ammon, N. Dittmar, K. J. Mattheier, & P. Trudgill (Ed.), *Handbooks of Linguistics and Communication Science (HSK)* (2nd compl. rev. and extend. ed., Vol. 1, pp. 1318–1332). Retrieved from <https://www.degruyter.com/viewbooktoc/product/178201>
- Variano, A.** (2014). L'elemento amerindio nel lessico italiano. Analisi storico-etimologica e documentazione lessicografica di parole provenienti dall'America centrale e meridionale (Dissertation zur Erlangung des akademischen Grades eines Doktors der Philosophie der philosophischen Fakultäten der Universität des Saarlandes). Universität des Saarlandes, Saarbrücken.
- Verax, C.** (1911). Propono por Terminologiaj Fundamentaj Principoj por la scienca lingvo en Esperanto. *Oficiala Gazeto Esperantista*, IV, 378–382.
- Versteegh, K.** (1993). Esperanto as a first language: language acquisition with a restricted input. *Linguistics*, 31, 539–555.
- Vézina, R.** (2004). La prise en compte de la variation lexicale dans la production terminologique et linguistique de l'Office de la langue française. In P. Bouchard & R. Vézina (Ed.), *La variation dans la langue standard. Actes du colloque tenu les 13 et 14 mai 2002 à l'Université Laval dans le cadre du 70^e Congrès de l'ACFAS* (pp. 181–193). Montréal: Office québécois de la langue française.
- Vila i Moreno, F. X.** (2007a). Estudis d'implantació terminològica: una aproximació en l'àmbit dels esports. Vic : Barcelona: Eumo Editorial ; Termcat.
- Vila i Moreno, F. X.** (2007b). Introducció. Terminologia i sociolingüística: un distanciament innecessari? In F. X. Vila i Moreno, M. Nogué Pich, & I. Vila i Moreno (Ed.), *Estudis d'implantació terminològica: una aproximació en l'àmbit dels esports* (pp. 13–28). Vic : Barcelona: Eumo Editorial ; Termcat.
- Vila i Moreno, F. X., & Nogué Pich, M.** (2007). L'anàlisi de la implantació terminològica en comunitats d'usuaris reduïdes: alguns elements metodològics des de la sociolingüística. In F. X. Vila i Moreno, M. Nogué

- Pich, & I. Vila i Moreno (Ed.), *Estudis d'implantació terminològica: una aproximació en l'àmbit dels esports* (pp. 29–67). Vic: Barcelona: Eumo Editorial ; Termcat.
- Vila i Moreno, F. X., & Vila i Moreno, I.** (2007). *Touché?* La implantació dels neologismes terminològics en el camp de l'esgrima. In F. X. Vila i Moreno, M. Nogué Pich, & I. Vila i Moreno (Ed.), *Estudis d'implantació terminològica: una aproximació en l'àmbit dels esports* (pp. 69–89). Vic: Barcelona: Eumo Editorial; Termcat.
- Vincent, D.** (1994). Le recours à l'exemple dans la langue spontanée. *Travaux de linguistique*, (28), 93–111.
- von Ahn, L., & Hacker, S.** (2017, September). duolingo incubator. Accessed on 2017/9/13: <https://incubator.duolingo.com/>
- von Wunsch-Rolshoven, L. F.** (2013). Die Verbreitung von Ideen entsprechend dem Modell „Diffusion of Innovations“ (Everett Rogers), angewandt auf die Verbreitung des Esperanto. *Plansprachen – ideengeschichtliche Aspekte*, 20, 85–105.
- Waarts, E., van Everdingen, Y. M., & van Hillegersberg, J.** (2002). The dynamics of factors affecting the adoption of innovations. *The Journal of Product Innovation Management*, (19), 412–423.
- Walsh, O.** (2016). *Linguistic Purism. Language Attitudes in France and Quebec*. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Wandel, A.** (2015). How many people speak Esperanto? Or: Esperanto on the web. *Interdisciplinary Description of Complex System*, 13(2), 318–321.
- Wardhaugh, R.** (2006). *An Introduction to Sociolinguistics* (5th ed.). Malden, MA: Blackwell.
- Waringhien, G.** (1980). *1887 kaj la sekvo*. Antverpeno-La Laguna: Stafeto.
- Waringhien, G.** (2005). *Plena Ilustrita Vortaro de Esperanto* (M. Duc Goninaz & C. Roux, Ed.). Paris: Sennacieca Asocio Tutmonda (SAT).
- Warschauer, M., Donaghy, K., & Kuamojo, H.** (1997). Leoki: a Powerful Voice of Hawaiian Language Revitalization. *Computer Assisted Language Learning*, 10(4), 349–361.
- Welger, H.** (1999). *Kontribuoj al la norma esperantologio. Enkonduko, ekspertizetoj, kritikoj*. Retrieved from <https://web.archive.org/web/20070928121718/http://www.homaranismo.info/Kontribuoj.htm>
- Wellman, B.** (2002). Little Boxes, Glocalization, and Networked Individualism. In M. Tanabe, P. van den Besselaar, & T. Ishida (Ed.), *Digital Cities II: Computational and Sociological Approaches* (pp. 10–25). Berlin Heidelberg: Springer-Verlag.

- Wennergren, B.** (2005). *Plena Manlibro de Esperanta Gramatiko 14.0*. Retrieved from <http://bertilow.com/pmeg/>
- Wennergren, B.** (2008, mai 13). La Oficialaj Aldonoj. Accessed on 2016/12/7: <http://lingvakritiko.com/2008/05/13/la-oficialaj-aldonoj/>
- Wennergren, B.** (n. d.). Lingva Kritiko. Studoj kaj notoj pri la Internacia Lingvo. Esperantologio Interreta. Retrieved from Lingva Kritiko. Studoj kaj notoj pri la Internacia Lingvo. Esperantologio Interreta website: <http://lingvakritiko.com/>
- Werneck Dias, A. E.** (2007). *Renovação lexical do esperanto: mecanismos de formação de neologismos* (Mestre em lingüística, Universidade de Brasília (UnB)). Retrieved from <http://educado.net/ext/dip?d=31>
- Werner, J.** (1986). *Terminologia kurso*. Roudnice n.L.: Sdružený klub ROH.
- Werner, J.** (2004). *Terminologiaj konsideroj*. Dobřichovice (CZ): KAVA-PECH.
- Wiebe, J. M., Bruce, R. F., Bell, M., Martin, M., & Wilson, T.** (2001). A corpus study of evaluative and speculative language. *SIGDIAL '01 Proceedings of the Second SIGdial Workshop on Discourse and Dialogue*, 15. Aalborg.
- Wiegand, H. E.** (1998). Wörterbuchforschung : Untersuchung zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie (Vol. 1). Berlin/New York: De Gruyter.
- Wiegand, H. E., Beißwenger, M., Gouws, R. H., Kammerer, M., Storrer, A., & Wolski, W.** (2010). Systematische Einführung. In H. E. Wiegand, M. Beißwenger, R. H. Gouws, M. Kammerer, A. Storrer, & W. Wolski (Ed.), *Wörterbuch zur Lexikographie und Wörterbuchforschung: Vol. Bd. 1, A–C* (pp. 1–121). Berlin, New York: De Gruyter.
- Wilensky, U.** (2016). NetLogo. Accessed on 2019/3/31: <http://ccl.northwestern.edu/netlogo/index.shtml>
- Williams-van Klinken, C.** (2008). Developing Electoral Terminology for a New Official Language: Tetun in East Timor. *Current Issues in Language Planning*, 5(2), 142–150.
- Willkommen, D.** (2007). *Esperanto-Grammatik. Eine Lerner- und Referenzgrammatik* (2., durchges. Auflage). Hamburg: Buske.
- Wilton, A., & Stegu, M.** (2011). Bringing the « folk » into applied linguistics. *AILA Review*, 24 (*Applied Folk Linguistics*), 1–14.
- Witell, L., Kristensson, P., Gustafsson, A., & Löfgren, M.** (2009). Idea generation: customer co-creation versus traditional market research techniques. *Publication Cover Journal of Service Management*, 22(2).

- Witkam, T.** (2007). Automatische Morphemanalyse in Esperanto macht Komposita besser lesbar auf dem Bildschirm. In D. Blanke (Ed.), *Esperanto heute. Wie aus einem Projekt eine Sprache wurde* (Vol. 16, p. 51–58). Berlin: Gesellschaft für Interlinguistik.
- Wood, R. E.** (1979). A Voluntary, Non-Ethnic and Non-Territorial Speech Community. In W. F. Mackey & J. Ornstein (Ed.), *Sociolinguistic Studies in Language Contact* (pp. 433–450). The Hague: Mouton.
- Wüster, E.** (1931). Internationale Sprachnormung in der Technik, besonders in der Elektrotechnik (Die nationale Sprachnormung und ihre Verallgemeinerung) (3., abermals ergänzte Auflage). Bonn: Bouvier u. CO. Verlag.
- Wüster, E.** (1968). Die vier Dimensionen der Terminologiarbeit. In Bundesverband der Dolmetscher und Übersetzer e.V. (BDÜ) (Ed.), *Reprint 3: Terminologie und Lexikographie: Tagungsakte (Nachdruck)* (Reprint, p. 33–67). Bonn: Bundesverband der Dolmetscher und Übersetzer e.V. (BDÜ).
- Żadziłko, E.** (2011). *Web 2.0 jako narzędzie w środowisku użytkowników języka esperanto* (Praca magisterska (Master's thesis)). Uniwersytet Warszawski (University of Warsaw), Warsaw.
- Zaki, E. N.** (2015). *Quelle image pour l'espéranto?: Regard sur l'image à travers les principales positions idéologiques du mouvement en Europe* (Mémoire présenté pour l'obtention de la Maîtrise universitaire en traduction, mention traduction spécialisée). Université de Genève, Genève.
- Zamenhof, L. L.** (1905). *Fundamento de Esperanto*. Retrieved from <http://archive.org/details/fundamentodeesp00zamegoog>
- Zamenhof, L. L.** (1990). *Lingvaj Respondoj. Konsiloj kaj Opinioj pri Esperanto* (7a eldono; G. Waringhien, Ed.). Eldonejo Ludovikito.
- Zarnikhi, A.** (2014). *Towards a Model for Terminology Planning* (Tesi doctoral UPF). Universitat Pompeu Fabra, Barcelona.
- Zerkina, N., Lomakina, Y., & Kostina, N.** (2015). Axiological linguistics and foreign language teaching. *Procedia – Social and Behavioral Sciences*, 199, 254–260.
- Zesch, T., Müller, C., & Gurevych, I.** (2008). Extracting Lexical Semantic Knowledge from Wikipedia and Wiktionary. *Proceedings of the International Conference on Language Resources and Evaluation, LREC*. Presentation at LREC, Marrakech.
- Zhang, Y., & Wildemuth, B. M.** (2009). Qualitative analysis of content. In B. M. Wildemuth (Ed.), *Applications of Social Research Methods to Questions in*

Information and Library Science (pp. 308–319). Westport (CT): Libraries Unlimited.

Zwick, D., Bonsu, S. K., & Darmody Aron. (2008). Putting Consumers to Work. “Co-creation” and new marketing govern-mentality. *Journal of Consumer Culture*, 8(2), 163–196.

Endnotes

- i French: élaboration
- ii French: élaboration linguistique
- iii French: pratique dirigée de la néologie
- iv Original quote in French: “Quel peut être le sens d’une action positive dans le domaine de la néologie ? Et, en premier lieu, comment la dénommer ? Le terme de planification néologique ne nous paraît conforme ni à l’esprit ni aux démarches qu’il est actuellement possible de conduire, démunis comme nous le sommes sur le plan linguistique, psychologique et sociologique. Nous optons quant à nous pour la notion et le terme d’assistance néologique.”
- v French: néologie d’aménagement, néologie aménagiste / Spanish: neología planificada
- vi French: innovation lexicale planifiée
- vii French: changement terminologique planifié
- viii Spanish: terminología planificada
- ix Original quote in French: “Non seulement les typologies sont nombreuses et établissent des classes et sous-classes plus ou moins nombreuses [...], mais encore elles sont fondées sur des critères qui ne relèvent pas des mêmes domaines : ils peuvent être radicalement hétérogènes, ce qui interdit toute comparaison directe d’une typologie à l’autre.”
- x Original quote in French: “Dans la période de création d’une réalité nouvelle et de formation d’un vocabulaire adéquat, c’est une caractéristique de la situation linguistique qu’un certain foisonnement néologique transitoire se produise pour désigner un même concept.”
- xi Original quote in French: “Certains termes sont, d’après nos résultats, connus mais non employés. Une telle situation n’a rien de surprenant théoriquement. Elle correspond plus ou moins à l’opposition bien connue des enseignants de langues vivantes entre vocabulaire passif et vocabulaire actif.”
- xii German: Sprachgefühl
- xiii Original quote in German: “Die genormte technische Sprache, wie sie in den Normveröffentlichungen niedergelegt ist, bedeutet einen großen Fortschritt hinsichtlich der Systemgüte. Stehen ihre Neuerungen aber nicht etwa nur auf dem Papier? Wie ist die Forderung nach Wirtschaftlichkeit der Sprachentwicklung erfüllt?”
- xiv Original quote in German: “Über die Erfolge oder Mißerfolge genormter Terminologie ist fast gar nichts bekannt. Weder gibt es genügend exakte Untersuchungen darüber, ob genormte Terminologie in der wissenschaftlichen, halbwissenschaftlichen Literatur und in Werkprospekten tatsächlich wie vorgeschrieben verwandt wird, noch ob sich das genormte *Synonym* gegenüber anderen *Synonymen* durchsetzt.”

- xv** Original quote in German: “Den Makrobereich bildet die durch den Mikrobereich hervorgebrachte Struktur [...]ist die kausale Konsequenz einer Vielzahl individueller intentionaler Handlungen, die mindestens partiell ähnlichen Intentionen dienen.”
- xvi** Spanish: cambio de época sociolingüística
- xvii** Original quote in French: “il ne faut être ni linguiste ni Académicien pour juger sur le bon usage et les normes. Il suffit de se brancher sur Internet.”
- xviii** Original quote in German: “In Online-Diskussionsforen zu sprachlichen Fragen ist immer wieder zu beobachten, wie Teilnehmer das Internet insbesondere über kommerzielle Suchmaschinen wie Google auf das Vorkommen bestimmter Wörter, Wortverbindungen oder grammatischer Konstruktionen hin überprüfen. Mit Hilfe der so generierten Verwendungsbelege und/oder ihrer errechneten Häufigkeiten stellen sie Thesen auf oder überprüfen sie.”
- xix** Original quote in German: “Der erste, bedeutende Filter liegt auf der Ebene der Zugriffsmöglichkeit. Er umfaßt – mit nachlassender Wichtigkeit – die Kriterien Gebräuchlichkeit, Bekanntheit, Einfachheit und Verständlichkeit. Hat ein Neologimus diesen ersten Filter passiert, so ist seine “passive Akzeptanz” sehr wahrscheinlich.”
- xx** Original quote in German: “Der zweite Filter, die Benutzbarkeit, ist vor allem an die individuelle Nützlichkeits einschätzung seitens des Sprachbenutzers hinsichtlich der Verwendung eines Neologimus in einer konkreten Kommunikationssituation gebunden, mit der das Kriterium der Adäquatheit aufs engste gekoppelt ist. In dieser Hinsicht ist auch die Quelle, die den Neologismus hervorgebracht hat, von Bedeutung. Auf dieser zweiten Ebene wirken außerdem, mit mindere m Einfluß, die Kriterien Korrektheit, ästhetische Qualitäten und Normalität”
- xxi** Original quote in French: “Le processus d’implantation des innovations consiste essentiellement en une série de choix et d’actions qui conduisent un individu ou une organisation à prendre connaissance d’une innovation, à se former des attitudes positives ou négatives à son égard, à prendre la décision d’adopter ou de rejeter cette innovation, à donner suite à cette décision de façon concrète et, finalement, à maintenir ou à modifier cette décision.”
- xxii** Original quote in German: “Eine Sprachlenkungsmaßnahme ist genau dann erfolgreich, wenn eine kollektive nachhaltige Veränderung des Sprachgebrauchs sowohl das Ziel als auch das Ergebnis darstellt.”
- xxiii** Original quote in German: “Es erhebt sich auch die Frage, in welchen Maße sich die Entwicklung der Sprache spontan vollzieht und inwieweit sie vom Menschen zielgerichtet, also bewußt oder “künstlich”, gesteuert werden kann.”
- xxiv** Original quote in French: “[...] il n’est pas possible d’isoler les effets d’une action politique des autres facteurs, linguistiques et extralinguistiques, qui jouent sur l’évolution de la langue.”
- xxv** Original quote in French: “Malgré l’utilisation de grilles de critères d’implantation lors du travail terminologique, nul ne saurait prédire quel sera l’usage ou les usages réels.”
- xxvi** Original quote in Catalan: “ja que si coneixem les variables que influencien l’ús dels termes podem obtenir les condicions per garantir la implantació de la terminologia normalitzada.”

- xxvii** Original quote in French: “[...] un *terminogramme* devrait aider à situer les raisons d’une situation linguistique à un moment du temps (en synchronie), et d’expliquer et de prévoir les évolutions terminologiques sur une longue durée (en diachronie)”
- xxviii** Original quote in French: “Il est important que tout au long du processus, une évaluation des résultats obtenus soit effectuée. Est-ce que les termes adoptés « passent » auprès des utilisateurs? Quel est le sentiment des personnes visées par le changement? Le standard terminologique correspond-il aux attentes des futurs utilisateurs?”
- xxix** Original quote in French: “Comme le temps est un facteur crucial dans l’implantation de terminologies, l’évaluation ne peut prendre place que longtemps après la diffusion de celles-ci.”
- xxx** Original quote in Catalan: “Considerem que una denominació és coneguda *espontàniament* quand és produïda per l’informant en situació d’enquesta.”
- xxxi** Original quote in French: “Dans l’ensemble, comme nous l’avons déjà indiqué, les officialismes sont inégalement et peu connus des rédacteurs. La majorité des répondants (50 % et plus) ne peut reconnaître que 20 % des formes officialisées.”
- xxxii** Original quote in Catalan: “Les entrevistes fetes als amos o responsables dels establiments a propòsit de la difusió de les denominacions catalanes proposades pel TERMCAT ens permeten confirmar la nostra primera hipòtesi: la difusió de les denominacions catalanes proposades pel TERMCAT ha estat nul·la, ja que cap d’ells coneixia les formes proposades ni els recursos terminològics disponibles per consultar-les.”
- xxxiii** Original quote in French: “Beaucoup d’entrevues ont commencé par une hésitation de la part du sujet, qui se disait ignorant des termes français, des termes ‘correctes’ [sic] [...]”
- xxxiv** Original quote in French: “Pour la connaissance du terme, il faut envisager plusieurs niveaux. Il est simpliste de dire qu’un terme est connu ou non, même d’un spécialiste. Il existe des degrés et nous avons cherché ici à mettre en oeuvre diverses stratégies d’accès permettant d’appréhender le niveau de familiarité avec le terme.”
- xxxv** Original questions in Catalan: respectively “Com anomenes això?”, “Coneixes alguna altra denominació per a aquest concepte?”, “Coneixes aquesta denominació?”
- xxxvi** Original questions in Catalan:
 “Coneixes el diccionari d’hoquei?
 Has vist algun cop un cartell com aquest [ensenyar cartell]?
 Has vist alguna vegada un fullet com aquest [ensenyar fullet]?
 Saps què és el TERMCAT?”
- xxxvii** Original quote in French: “Par exemple, lorsque j’ai ouvert la porte d’un garage j’ai entendu quelqu’un dire ‘Va chercher les tires!’. Par la suite j’ai fait une entrevue avec cette personne (c’était le propriétaire), et il m’a dit qu’il n’utilisait que *pneu*.”
- xxxviii** Original quote in French: “[a]nalyse des degrés de connaissance des termes officiels et de leurs significations [...]”
- xxxix** Original quote in French: “estimer le degré de connaissance des termes recommandés”
- xl** Original quote in Catalan: “[...] copsar les opinions entorn de les denominacions i els esforços de normalització en l’àmbit de l’esport en general”
- xli** Original quote in Catalan: “Creus que es podria acabar estenent?”
- xlïi** French: discours épiterminologiques

- xlili** Original quote in Catalan: “A l’hora d’abordar un estudi d’implantació terminològica, és a dir, de l’impacte que ha tingut la difusió d’una determinades propostes de normalització terminològica [...]”
- xliv** Original quote in French: “[...] il s’agirait de découvrir où les termes sont appris, d’où ils viennent, et comment ils se diffusent à l’intérieur d’une génération”
- xlv** Original quote in Catalan: “Així, per exemple, el fet de conèixer a priori les vies de difusió reals per a les innovacions proporciona una perspectiva molt valuosa a l’hora de proposar formes innovadores [...]”
- xlvi** Original French quote: “[...] la branche de la terminologie qui étudie les terminologies selon un point de vue ethnographique (étude de terrain) et ethnologique (généralisation des faits observés et comparaison des groupes humains).”
- xlvii** Original French quote: “Utiliser telle ou telle langue ou tel ou mot est un événement linguistique (ou langagier) constitué par l’interaction de plusieurs composantes, dont la langue n’en est qu’une.”
- xlviii** Original quote in French: “Ce discours sur les langues peut être explicitement sollicité, à partir de questionnaires qui placent les locuteurs en situation de réagir à des productions, de produire des jugements sur les langues parlées ou écrites dans une communauté déterminée. Mais il peut être repéré dans de nombreuses productions discursives spontanées ou non (discours politiques, syndicaux, textes divers, journalistiques, littéraires, pédagogiques...), écrites ou orales, en particulier dans les situations conflictuelles où les langues constituent un enjeu de pouvoir.”
- xliv** Original quote in French: “L’étude du corpus nous fait cependant constater, comme nous le verrons, que des facteurs subjectifs ou des considérations d’ordre métalinguistique interviennent dans le choix et l’emploi de certains vocables. Considérations métalinguistiques, dans la mesure où les utilisateurs du vocabulaire prennent une certaine distance avec celui-ci et où ils ne l’assument pas encore complètement; dans la mesure aussi où ils prennent certaines précautions lorsqu’ils emploient le vocabulaire anglo-saxon, qu’ils prennent soint de marquer par rapport au terme français correspondant, ou lorsqu’ils commentent la validité de l’emploi de certains termes. Et facteurs subjectifs, qui font intervenir dans le choix des termes certaines connotations sous-jacentes.”
- I** Original quote in German: “Ein Zugang zu einem breiten Spektrum and sprachkritischen Kommentaren von Laien ist über Forenkommunikation möglich. In den sprachkritischen Kommentaren werden Normen der linguistischen Laien sichtbar.”
- li** Original quote in Esperanto: “‘Selektiva’ estas malbona vorto, kvankam gxi certe altentas iujn uzantojn, kaj cxiuj inhib-vortoj (inhibi, inhibicio, inhibitoro) estas tute nenecesaj kaj nur malklarigas la aferon.”
- lii** Original quotes in Esperanto: “Kiel oni diras ‘Walkie-talkie’ aŭ ‘CB Radio’
[...] Mi nur trovis la vorton ‘portebla radiotelefono’ rilate al ‘talkie-walkie’ en dulingva (Franca-Esperanto) vortaro eldonita de SAT-Amikaro en 2000.
Pri walkie-talkie mi eltrovis ankaŭ la vorton ‘promenradio’ (Minnaja).[...]”
- liii** Original quote in Esperanto: “En julio 1887 eliris el presejo la unua eldonaĵo, la rusa fundamenta lernolibro, por kiu estis ricevita permeso de la rusa cenzuro antaŭan

monaton [...] Saman jaron D-ro Zamenhof ankaŭ eldonis polan, francan kaj germanan tradukon de tiu unua broŝuro, ĉiam laŭ la sama plano [...]"

- liv** German: Plansprache
- lv** Original quote in German: "Die Frage nach der Datierung von neuen Spracherscheinungen ist im Esperanto wie in den Ethnosprachen meistens schwer zu beantworten. Jede Neuerung setzt sich nur sukzessiv durch und beginnt mit dem Sprechakt eines einzelnen Individuums oder, im Falle einer ursprünglichen Interferenz, mit dem spezifischen Sprachgebrauch einer bestimmten Sprechergruppe gleicher Primärsprache. [...] Bei welchen Sprechern und vor allem zu welchem Zeitpunkt die Neuerung ihren Ursprung hat, läßt sich nur ganz selten erschließen, denn sie läßt sich gewöhnlich erst dann feststellen, wenn sie bereits allgemein übernommen worden ist."
- lvi** German: einheitliches Schema
French: schéma préalable.
- lvii** In Italian and Esperanto respectively: simpatizzante/simpatianto and attivista/aktivulo.
- lviii** In Italian and Esperanto respectively: il principiante/komencanto, l'abile/lertulo, il fluente/fluulo and il denaska/denaskulo.
- lix** Italian: esperantofono
- lx** Italian: respectively esperantofono and esperantista
- lxi** Original quote in French: "Mais ce n'est pas Zamenhof qui fait évoluer l'espéranto, et tant qu'on s'appuiera sur des écrits de Zamenhof, on ne sera pas armé pour étudier l'évolution de l'espéranto."
- lxii** Original quote in Italian: "L'esperanto può essere visto come un laboratorio linguistico veramente straordinario, perché molte variabili fondamentali della lingua – quali data di nascita, vocabolario di base, numero di parlanti a una certa data – possono essere se non certi almeno più delineati rispetto alle lingue storico-naturali."
- lxiii** Original quote in Esperanto: "[...] la ĝenerala principo de Zamenhof estas doni informon pri la sistemo, sed ne pri la uzo."
- lxiv** Original quote in Esperanto: "Ĉu do Zamenhof estas nura konstruinto de la lingvosistemo, lasante la uzon al hazardo? Ne. Li metis en sian lingvosistemon la potencialon por evoluo kaj li ja donis al ĝia evolukapablo difinitan direkton. Tiu direkto implicate kaŝiĝas en la aglutina vortfarado kaj en la malpreciza konekto al la etnolingvaj modeloj [...]"
- lxv** Original quote in Esperanto: "Mi akceptas la decidon de la Akademio, kiu per la BRO ŝanĝis la kategorion de tondr-. Persone mi dubas, ke ili faris tion per formala decido pri ĉiu ŝanĝita radiko, tamen ni devas akcepti la rezulton."
- lxvi** Original quote in Esperanto: "Grave estas solvi konsiston de la kolektivo. Ĝia gvidanto estu homo fake kaj lingve kvalifikita, lia karaktero plenumu premison de la demokrata traktado en la kolektivo. La membroj apartenus al diversaj etnaj komunumoj kun malsamaj lingvoj."
- lxvii** Original quote in Esperanto: "La ĉho pri la revizia laboro istigis multajn fakulojn aŭ ordinarajn uzantojn de la vortaro sendi propraincate siajn kontribuojn, proponojn, rimarkojn."

- lxxviii** Original quote in Esperanto: “[...] la tri TEC-organizantoj estis kaj restis la solaj kun siaj iom superdimensioniaj planoj: daŭre mankis iu kunordiga gvidanto kaj ĉefe mankis amaso da kunlaborpretaj fakuloj, kiuj ne nur okupigis pri siaj specialaj fakoj, sed pretus kun-organizi en la planitaj strukturoj.”
- lxxix** Original quote in German: “Aufgrund der Art und Weise, wie Termini in Esperanto geprägt werden (s. Kap. 5.1), ist es in den meisten Fällen nicht eindeutig zu sagen, ob der Sprachgebrauch in bezug auf Fachtermini den Wörterbüchern folgt oder ob umgekehrt die Wörterbücher bei der Registrierung von Fachtermini den Fachtexten folgen; wahrscheinlich spielt beides eine Rolle. Gewisse Einflüsse von Wörterbüchern auf den Sprachgebrauch in Fachtexten sind aber durchaus zu erkennen”
- lxxx** Original quote in Esperanto: “Lingva Konsultejo estas forumo por interparolo kaj konsilado pri la strukturo kaj uzado de Esperanto. Ĝi celas trakti pli maloftajn, nekutimajn, komplikajn kaj neklarajn flankojn de la lingvo, kies respondoj povas esti malfacile troveblaj. Cetere, ĝi celas informon kaj klarigojn de homoj kiuj bone scias Esperanton kaj havas sperton pri la lingvo diversmaniere (verkado, instruado, legado, ktp). La celo ne estas klarigi simplajn demandojn, kiel en lernolibro aŭ baza kurso, nek difini vortojn kiuj troviĝas en vortaro.”
- lxxxi** Tial ni petas
1. ke vi ne demandu pri tio, kion vi facile trovos en reta vortaro, kiel vortaro.net aŭ reta-vortaro.de;
 2. kaj ke vi ne aŭŝu en la grupo se vi ne havas lingvo-rilatajn demandojn.”
- lxxxii** German: Ausbau eines retrodigitalisierten Printwörterbuches
- lxxxiii** German: aktuelle Neologismenlexikographie
- lxxxiv** Original quote in Esperanto: “ViVo estas celita por la vortoj kiuj ‘vivas’ kaj do ankoraŭ povas ‘morti’, malaperi. Multaj el ili meritis la eternan vivon kaj eniros la paradizon por la vortoj, tio estas la ReVo (Reta Vortaro) de Esperanto. La ViVo-vikio estas diskutejo, antaŭpreparo por aliaj vortaroj kaj terminaroj kiel ReVo, Vikivortaro kaj la vortaro de Lernu”
- lxxxv** Original quote in Esperanto: “Problemo tamen restas parolaj situacioj, kiam ne eblas simpe konsulti Vikipedion, enciklopediojn, interreton, ...”
- lxxxvi** Original quote in Esperanto: “En miaj spertoj, foje mi kaj miaj korespondamikoj, sen koni la taŭgan esprimon en ia situacio, simple skribas la signifon de tiu esprimo kaj la vorton en nia lingvo, aŭ en alia lingvo.”
- lxxxvii** Original quote in French: “[...] on a remarqué que les étrangers apprenant le français [...] n’hésitent pas à créer les unités lexicales françaises dont ils ont besoin en appliquant les règles de création des mots qu’ils ont intégrées. La maîtrise de plusieurs langues a sans doute des incidences sur les mécanismes intellectuels en action dans les activités langagières et la gymnastique mentale liée aux passages d’un lexique à un autre facilite probablement l’activation des procédés de formation des unités lexicales, et ce dans toutes les langues.”
- lxxxviii** Original quote in Esperanto: “Se temas pri inventi novajn radikojn, oni hezitu. Sed se temas pri novaj kunmetitaĵojn, estas cxiam amuza kaj bone komprenebla.”
- lxxxix** Original quote in Esperanto: “jes, kelkfoje ili mankas, sed en la esperantlingva Vikipedio aŭ en la Hejma Vortaro aŭ eĉ en PIV povas troviĝi interesaj versioj. Se ne,

mi serĉas en aliaj eŭropaj lingvoj kaj, eble, kalkas la esprimon kiun mi ŝatis. Padonu, 'kalki' ne estas la ĝusta vorto. Paŭsas.”

- lxxx** Original quote in Esperanto: “Ĝis nun, mi serĉis ekzemplojn en Esperanto pri arto (tre malgranda aro, laŭ mia sperto) kaj, se mi ne trovis ion, mi kreis mian propran terminon, kaj aldonis eksplikon post la artikolo.”
- lxxxix** Original quote in Esperanto: “Alfredo, bone uzi jam-estantajn estas pli bone ol krei novajn vortojn kaj fari la vortaron eĉ pli grandan.”
- lxxxix** Original quote in Esperanto: “Vi pravas. Ni ne enkonduku novajn vortojn, kiam ili jam ekzistas kaj taŭgas. Sed kiam ili ne ekzistas aŭ ne taŭgas?”
- lxxxix** Original quote in Esperanto: “Dum mi renkontis la problemon, mi povas konsulti vortaron paperan aŭ simple interretan.”
- lxxxix** Original quote in Esperanto: “Mi neniam uzas retvortaron. Sed kial ne.”
- lxxxix** Original quote in Esperanto: “Mi agnoskas nur uzi retan vortaron...”
- lxxxix** Original quote in Esperanto: “Por modernaj aŭ strangaj vortoj, aŭ stranga uzado, mi iras al Google serĉilo. Ĝenerale en la unua paĝo mi trovas kion mi serĉas.”
- lxxxix** Original quote in Esperanto: “Mi ofte parolas pri komputiloj, aŭtomobiloj, kaj tiel plu. Ekz. mi ne trovis la tradukon de ‘Smartphone’ en la PIV. Se mi serĉas modernaj vortoj mi serĉas tion unue en la germana aŭ angla Vikipedio, kaj poste provas sxalti al esperanta Vikipedio por vidi tekstojn pri la sama temo – ofte mi trovis vortojn kiuj ne ekzistas en la PIV.”
- lxxxix** Original quote in Esperanto: “Paradokse Vikipedio estas sufiĉe bona plurlingva vortaro kaj sufiĉe malbona enciklopedio laŭ mia opinio. Alivorte ĝi prezentas sufiĉe kontentigajn kapvortojn (krom propraj nomoj, kiujn ĝi tre malmulte tradukas) eĉ se la enhavo de la artikoloj lamas.”
- lxxxix** Original quote in Esperanto: “[...] en la reto oni povas trovi verajn ekzemplojn de uzado, ne nur tiujn cititajn en PIV.”
- xc** Original quote in Esperanto: “Ĝenerale oni ne timu demandi samideanojn. Kun Interreto tio tre facilas!”
- xc** Original quote in Esperanto: “Foje mi aktive serĉas en la reto kaj babilejo kaj foje ricevas bonegan helpon.”
- xc** Original quote in Esperanto: “Mi pensas ke la rezulto de la kunlaborado de multaj profesiaj kunlaborantoj kaj kelkaj nesciuloj povas esti fusxa – vortaro estas nur bona, se estas fidinda.”
- xc** Original quote in Esperanto: “Oni ne scias ke la verkinto estas nefakulo. Sed oni povas tion malkovri cxe iu vorto kiun oni mem komprenas pli bone ol la verkinto. Se oni tiam povas certi ke ne temas pri tajp- aux preseraro.”
- xc** Original quote in Esperanto: “Ni fakte ne scias kiajn spertojn havas la volontuloj kiuj kontribuas al tiuj enretaj vortaroj; eble ili estas tre sperta kaj klera.”
- xc** Original quote in Esperanto: “Sed, fakte, mi nur uzas du vortarojn. Unu el ili estas la vortaro trovita je lernu.com. La alia estas vortareto esperanto-(mia nacia lingvo) kaj mi neniam pensis pri la uloj kiu laboris en la farado de la vortarojn. [...] Mi neniam pensis pri tio! Laux mi ekpensis nun, eble la vortaro esperanto-(mia nacia lingvo) estus

- pli profisia, cxar gxi estas papera vortaro. Sed mi neniam pensis pri la scipovo de la ulo(j) kiu skribis gxin.”
- xcvi** Original quote in Esperanto: “Mi agnoskas ne atenti ankaŭ: mi kontrolas la vorton mem ĉu ĝi taŭgas al mi”
- xcvii** Original quote in Esperanto: “Dum mi elektas vortaron por mi, mi pli atentis, cxu la vortaro povas kontigi mian bezonon. [...] Tamen ne ĉiu vortaro povas kontentigi mian bezonon, do ĉe mi estas ankaŭ pluraj vortaroj.”
- xcviii** Original quote in Esperanto: “al tiuj vortaroj oni rilatu speciale kritike (aŭ malfide, kiel iu ĵus diris). Gravas bone scii la lingvon por ne kaŭpi el vortaro eraron. ‘Por doni bonan demandon endas scii pejparton de la respondo.’”
- xcix** Original quote in Esperanto: “Vikivortaro povas esti bona, mi ofte uzas ĝin (por aliaj lingvoj ol Esperanto), sed ja indas trakti ĝin kritikeme.”
- c** Original quote in Esperanto: “La avantaĝo de tiaj ‘vortaroj’ (aŭ de serĉo tra Google de iu vorto) estas la moderneco kaj akurateco de tiuj vortoj”
- ci** Original quote in Esperanto: “Foje oni rimarkas, ke iu vortare trovita esperantigo estas simple elturniĝo de la verkinto. Tiam mi sentas min rajtigita pripensi, ĉu alia formo eventuale pli taugas, t.e. io, kion mi me elpensas.”
- cii** Original quote in Esperanto: “PIV ne estas absoluta dogmo. Oni kontrolu laŭ ĝi, sed konsideru ankaŭ aliajn fontojn (ne nur vortarojn, sed realan uzon)”
- ciiii** Original quote in Esperanto: “mi ne ezitas neuzi vorton de la vortaro se mi trovas ke ĝi ne taŭgas. [...] Do mi ne ‘fidas’ entute pri la vortaroj.”
- civ** Original quote in French: “Entendu à Radio-Canada pour remplacer ‘cuisinmane’ : ‘gastronaute’. Un peu long, mais piste intéressante! #foodie”
Original quote in Esperanto: “Mi ege bedaŭras tiun senutilan disputon kiun proponis S-ro [Name] [...]”
- cv** Original quote in Esperanto: “mi ege bedaŭras tiun senutilan disputon kiun proponis S-ro [Name] [...]”
- cvii** Original quote in Esperanto: “Mi ankaŭ ne ŝatas radikojn kun duoblaj vokaloj kaj pensas, ke ili ne estu en E-o.”
- cviii** Original quote in Esperanto: “neniam uzis la vorton ‘brodkasti’, ankaŭ por mi gxi sonas kiel evitinda anglismo”
- cxix** Original quote in Esperanto: “Ruse kaj mongole la meza konsonanto estas voĉa”
- cix** Original quote in Esperanto: “En la portugala ne ekzistas unikaj vortoj por NERD kaj GEEK.”
- cx** Original quote in Esperanto: “[Name], kiu longe vivis en Taŝkento, plej probable eĉ proksime vidis tie tiajn malpli ekzotikajn varanojn.”
- cxii** Original quote in Esperanto: “The letter ‘x’ exists in the Polish alphabet, but currently it is not used in words.”
- cxii** Translation in English: “I prefer vertical, because it goes well with horizontal”

Appendices

The Appendices are available on Frank & Timme's website:

https://www.frank-timme.de/fileadmin/docs/Maradan_Appendices.pdf



SPRACHWISSENSCHAFT

- Bd. 1 Sergio Viaggio: *A General Theory of interlingual Mediation*.
408 Seiten. ISBN 978-3-86596-063-4
- Bd. 2 Karl-Ernst Sommerfeldt: *Regionalgeschichte im Spiegel der Sprache*.
188 Seiten. ISBN 978-3-86596-077-1
- Bd. 3 Ulla Fix: *Stil – ein sprachliches und soziales Phänomen. Beiträge zur Stilistik*.
Herausgegeben von Irmhild Barz, Hannelore Poethe, Gabriele Yos.
462 Seiten. ISBN 978-3-86596-138-9
- Bd. 4 Anja Seiffert: *Autonomie und Isonomie fremder und indigener Wortbildung
am Beispiel ausgewählter numerativer Wortbildungseinheiten*.
418 Seiten. ISBN 978-3-86596-141-9
- Bd. 5 Ulla Fix: *Texte und Textsorten – sprachliche, kommunikative und
kulturelle Phänomene*. 510 Seiten. ISBN 978-3-86596-179-2
- Bd. 6 Juliana Goschler: *Metaphern für das Gehirn. Eine kognitiv-linguistische
Untersuchung*. 244 Seiten. ISBN 978-3-86596-188-4
- Bd. 7 Maria Wirf Naro: *La composition en allemand : structure sémantique et fonction
littéraire*. Mit einer ausführlichen Zusammenfassung in deutscher Sprache.
416 Seiten. ISBN 978-3-86596-284-3
- Bd. 8 Bettina Bock/Ulla Fix/Steffen Pappert (Hg.): *Politische Wechsel –
sprachliche Umbrüche*. 404 Seiten. ISBN 978-3-86596-363-5
- Bd. 9 Kerstin Preiwuß: *Ortsnamen in Zeit, Raum und Kultur. Die Städte Allenstein/Olsztyn
und Breslau/Wrocław*. 470 Seiten. ISBN 978-3-86596-368-0
- Bd. 10 Werner Holly: *Sprache und Politik. Pragma- und medienlinguistische Grundlagen
und Analysen*, herausgegeben von Sonja Ruda und Christine Domke.
394 Seiten. ISBN 978-3-86596-438-0
- Bd. 11 Hatice Deniz Canoğlu: *Kanak Sprak versus Kiezdeutsch – Sprachverfall
oder sprachlicher Spezialfall? Eine ethnolinguistische Untersuchung*.
138 Seiten. ISBN 978-3-86596-483-0
- Bd. 12 Ulla Fix: *Sprache in der Literatur und im Alltag. Ausgewählte Aufsätze*.
452 Seiten. ISBN 978-3-86596-369-7
- Bd. 13 Anja Hennemann/Claudia Schlaak (Hg.): *Korpuslinguistische Untersuchungen.
Analysen einzelsprachlicher Phänomene*. 202 Seiten. ISBN 978-3-86596-519-6
- Bd. 14 Joachim Born/Wolfgang Pöckl (Hg.): *„Wenn die Ränder ins Zentrum drängen ...“.
Außenseiter in der Wortbildung(sforschung)*. 304 Seiten. ISBN 978-3-86596-449-6

SPRACHWISSENSCHAFT

- Bd. 15 Ulla Fix: Sprache, Sprachgebrauch und Diskurse in der DDR. Ausgewählte Aufsätze. 744 Seiten. ISBN 978-3-86596-549-3
- Bd. 16 Jenny Brumme/Sandra Falbe (ed.): The Spoken Language in a Multimodal Context. Description, Teaching, Translation. 316 Seiten. ISBN 978-3-7329-0021-3
- Bd. 17 Inna Kozlova: De la búsqueda a la consulta. Resolución de problemas léxicos en producción textual en lengua extranjera. 374 Seiten. ISBN 978-3-7329-0049-7
- Bd. 18 Gerd Antos/Ulla Fix/Bettina Radeiski (Hg.): Rhetorik der Selbsttäuschung. 262 Seiten. ISBN 978-3-86596-513-4
- Bd. 19 Carmela Pérez-Salazar/Inés Olza (eds.): Del discurso de los medios de comunicación a la lingüística del discurso. Estudios en honor de la profesora María Victoria Romero. 598 Seiten. ISBN 978-3-7329-0058-9
- Bd. 20 Cordula Schwarze/Carmen Konzett (Hg.): Interaktionsforschung: Gesprächsanalytische Fallstudien und Forschungspraxis. 234 Seiten. ISBN 978-3-7329-0073-2
- Bd. 21 Ingrid Wotschke: How Educated English Speak English. Pronunciation as Social Behaviour. 282 Seiten. ISBN 978-3-7329-0062-6
- Bd. 22 Rudolf von Raumer: Geschichtliche Grammatik der Deutschen Sprache (1851). Faksimile der Handschrift und Transkription, herausgegeben von Hannelore Schwartze-Köhler. Mit einer Einleitung von Horst Haider Munske. 326 Seiten. ISBN 978-3-7329-0079-4
- Bd. 23 Josef Klein: Grundlagen der Politolinguistik. Ausgewählte Aufsätze. 392 Seiten. ISBN 978-3-7329-0110-4
- Bd. 24 Joanna Szczek: Absageschreiben auf Bewerbungen. Eine pragma-linguistische Studie. 420 Seiten. ISBN 978-3-7329-0146-3
- Bd. 25 Hartmut E.H. Lenk/Ulrike Richter-Vapaatalo (Hg.): Sie leben nicht vom Verb allein. Beiträge zur historischen Textanalyse, Valenz- und Phraseologieforschung. 228 Seiten. ISBN 978-3-7329-0098-5
- Bd. 26 Anja Hennemann/Claudia Schlaak (Hg.): Politische und mediale Diskurse. Fallstudien aus der Romania. 296 Seiten. ISBN 978-3-7329-0074-9
- Bd. 27 Alejandro González Villar: Un análisis funcional y descriptivo de los marcadores pragmáticos y su traducción como herramienta en la construcción del diálogo ficticio. Estudio contrastivo alemán-catalán-español en base a tres novelas de Hans Fallada. 436 Seiten. ISBN 978-3-7329-0183-8
- Bd. 28 Josef Klein: Von Gandhi und al-Qaida bis Schröder und Merkel. Politolinguistische Analysen, Expertisen und Kritik. 432 Seiten. ISBN 978-3-7329-0123-4

SPRACHWISSENSCHAFT

- Bd. 29 Sorin Gadeanu: Intelligente Laute: Ein System mit eigener Ordnung. Dynamisch-funktionale Kategorien zur „Phonetologie“ des Deutschen für Rumänischsprachige / Sunete inteligente: Un sistem cu ordine proprie. Categoriile dinamice-funcționale ale „fonetologiei“ limbii germane pentru vorbitorii de limba română. 344 Seiten. ISBN 978-3-7329-0229-3
- Bd. 30 Juan Cuartero Otal/Juan Pablo Larreta Zulategui/Christoph Ehlers (Hg.): Querschnitt durch die deutsche Sprache aus spanischer Sicht. Perspektiven der Kontrastiven Linguistik. 272 Seiten. ISBN 978-3-7329-0244-6
- Bd. 31 Uwe Grund: Orthographische Regelwerke im Praxistest. Schulische Rechtschreibleistungen vor und nach der Rechtschreibreform. 248 Seiten. ISBN 978-3-7329-0279-8
- Bd. 32 Enrico Garavelli/Hartmut E. H. Lenk (Hg.): Verhüllender Sprachgebrauch. Textsorten- und diskurstypische Euphemismen. 174 Seiten. ISBN 978-3-7329-0213-2
- Bd. 33 Bettina Radeiski: Denkstil, Sprache und Diskurse. Überlegungen zur Wiederaaneignung Ludwik Flecks für die Diskurswissenschaft nach Foucault. 210 Seiten. ISBN 978-3-7329-0325-2
- Bd. 34 Coline Baechler/Eva Martha Eckkrammer/Johannes Müller-Lancé/Verena Thaler (Hg.): Medienlinguistik 3.0 – Formen und Wirkung von Textsorten im Zeitalter des Social Web. 354 Seiten. ISBN 978-3-7329-0078-7
- Bd. 35 Joanna Szczek: Phraseologie der Farben. Phraseologisches Wörterbuch Deutsch–Polnisch/Polnisch–Deutsch. 130 Seiten. ISBN 978-3-7329-0337-5
- Bd. 36 Dieter Cherubim: Sprachliche Aneignung der Wirklichkeit. Studien zur Sprachgeschichte des neueren Deutsch. 636 Seiten. ISBN 978-3-7329-0250-7
- Bd. 37 Bernd Spillner: Error Analysis in the World. A Bibliography. 586 Seiten. ISBN 978-3-7329-0356-6
- Bd. 38 Karolina Suchowolec: Sprachenkung – Aspekte einer übergreifenden Theorie. 458 Seiten. ISBN 978-3-7329-0339-9
- Bd. 39 Anna Gondek/Joanna Szczek (Hg.): Kulinarische Phraseologie. Intra- und interlinguale Einblicke. 250 Seiten. ISBN 978-3-7329-0389-4
- Bd. 40 Marcelina Kałasznik: Bezeichnungen für bekannte Personen im Polnischen und im Deutschen. Eine lexikologische Analyse. 512 Seiten. ISBN 978-3-7329-0398-6
- Bd. 41 Larissa Naiditsch/Anna Pavlova: Prädikatives Attribut. Eine Vergleichsstudie für Deutsch und Russisch. ISBN 978-3-7329-0426-6. 180 Seiten

SPRACHWISSENSCHAFT

- Bd. 42 Irmtraud Behr/Florence Lefeuve (éds): Le genre bref. Des contraintes grammaticales, lexicales et énonciatives à une exploitation ludique et esthétique. 238 Seiten. ISBN 978-3-7329-0412-9
- Bd. 43 Hana Menclová: Diminutiva im Deutschen und Tschechischen. Eine Studie am Beispiel von Märchen des 19. Jahrhunderts und der Gegenwart. 292 Seiten. ISBN 978-3-7329-0494-5
- Bd. 44 Gerd Antos: Wissenskommunikation. Ausgewählte Aufsätze. 514 Seiten. ISBN 978-3-7329-0602-4
- Bd. 45 Cinzia Gigliani: Linguistic and Rhetorical Perspectives on Congressional Hearings. 150 Seiten. ISBN 978-3-7329-0646-8
- Bd. 46 Sylvia Weber: Computergesteuerte etymologische Wörterbücher. Konzeption und Umsetzung anhand des Französischen. 388 Seiten. ISBN 978-3-7329-0710-6
- Bd. 47 Hartmut E.H. Lenk/Hans W. Giessen (Hg.): Persuasionsstile in Europa. Strategien und Mittel des Überzeugens in Zeitungskommentaren aus kulturkontrastiver Sicht. 238 Seiten. ISBN 978-3-7329-0517-1
- Bd. 48 Mélanie Maradan: Uncertainty in deliberate lexical interventions. Exploring Esperanto speakers' opinions through corpora. 372 Seiten. ISBN 978-3-7329-0690-1

