

THOMAS JURCZYK

THE NOTION
OF "HOLY"
IN ANCIENT
ARMENIAN TEXTS
FROM THE
FIFTH CENTURY CE

A COMPARATIVE APPROACH USING
DIGITAL TOOLS AND METHODS

DIGITAL HUMANITIES RESEARCH
BIELEFELD UNIVERSITY PRESS

Thomas Jurczyk
The Notion of »holy« in Ancient Armenian Texts from the Fifth Century CE

Editorial

Digital Humanities is an evolving, cross cutting field within the humanities employing computer based methods. Research in this field, therefore, is an interdisciplinary endeavor that often involves researchers from the humanities as well as from computer science. This collaboration influences the methods applied as well as the theories underlying and informing research within those different fields. These implications need to be addressed according to the traditions of different humanities' disciplines. Therefore, the edition addresses all humanities disciplines in which digital methods are employed. **Digital Humanities Research** furthers publications from all those disciplines addressing the methodological and theoretical implications of the application of digital research in the humanities. The series is edited by Silke Schwandt, Anne Baillot, Andreas Fickers, Tobias Hodel and Peter Stadler.

Thomas Jurczyk (Dr.), born in 1985, studied history and religious studies in Bochum. He finished his Ph.D. in September 2020. Since 2014, he has worked as a research assistant at the »Center for Religious Studies« (CERES) at Ruhr-Universität Bochum. His principal areas of interest are Greco-Roman religions, early Christianity, an/iconism, and the application of computational methods and tools.

Thomas Jurczyk

The Notion of »holy« in Ancient Armenian Texts from the Fifth Century CE

A Comparative Approach Using Digital Tools and Methods

[transcript]

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>



This work is licensed under the Creative Commons Attribution 4.0 (BY) license, which means that the text may be remixed, transformed and built upon and be copied and re-distributed in any medium or format even commercially, provided credit is given to the author. For details go to <http://creativecommons.org/licenses/by/4.0/>

Creative Commons license terms for re-use do not apply to any content (such as graphs, figures, photos, excerpts, etc.) not original to the Open Access publication and further permission may be required from the rights holder. The obligation to research and clear permission lies solely with the party re-using the material.

© 2022 Thomas Jurczyk

Published by Bielefeld University Press, an Imprint of transcript Verlag.

<http://www.bielefeld-university-press.de>

Cover layout: Maria Arndt, Bielefeld

Printed by Majuskel Medienproduktion GmbH, Wetzlar

Print-ISBN 978-3-8376-6181-1

PDF-ISBN 978-3-8394-6181-5

<https://doi.org/10.14361/9783839461815>

ISSN of series: 2747-5476

eISSN of series: 2749-1986

Printed on permanent acid-free text paper.

Contents

Acknowledgements	17
1. Introduction	19
1.1 The Importance of Comparative Notions for the Study of Religion	22
1.2 Overall Structure of This Book	22
2. State of Research	25
2.2 History of Armenia in the Fifth Century CE	25
2.3 The Notion of Holy in the Study of Religion	36
2.4 Historical Notions of Holy	50
2.5 Corpus Linguistics and Distributional Semantics	63
3. Methodology	65
3.2 Text Statistics	66
3.3 Synoptic Tables and Sentence Analysis	71
3.4 Annotation Scheme	74
4. Data	79
4.1 Holy/Sacred English Corpus (HSEC)	79
4.2 Ancient Armenian Full Text Corpus (AAFTC) and Ancient Armenian Subr Corpus (AASC)	95
5. Examination	107
5.1 Armenian Corpora	107
5.2 English Corpora	203
6. Comparison	305
6.1 Examination of the Individual Semantic Fields	306
6.2 The Comparison	340

7. Conclusion	363
7.1 Results of the Individual Examinations	363
7.2 Overlappings between <i>Surb, Holy</i> , and the Comparative Notion of <i>Holy</i>	365
7.3 Contributions to the History of Religion and the Comparative Notion of <i>Holy</i>	367
7.4 Critical Review of the Methodology	370
Bibliography	371

List of Figures and Tables

Figure 1: Potential overlappings between the three semantic fields examined in this book.	21
Figure 2: Holy and its subfields in the Study of Religion.	50
Figure 3: Pipeline of methods used in this book.	66
Table 1: Exemplary row in a synoptic table.	71
Figure 4: Number of tokens (running words) in absolute numbers of the texts included in the HSEC.	81
Table 2: Exemplary row after the processing of the "Tweet objects" with Python.	86
Figure 5: F ₁ scores of the trained classifiers.	90
Figure 6: Top 20 most frequent user locations in the Twitter data (large sample).	92
Figure 7: Top 20 most frequent locations from where the tweets in the Twitter data (large sample) were sent (countries).	93
Figure 8: Top 20 most frequent associated locations (countries) in the Twitter data (large sample).	94
Figure 9: Distribution of tokens (running words) in the Ancient Armenian Full Text Corpus (AAFTC).	102
Figure 10: Token (running words) distribution in the Ancient Armenian Surb Corpus (AASC).	104
Figure 11: Keywords and lockwords in the AASC (+) and AAFTC (-).	109
Table 3: English translation keywords and lockwords in the AASC (+) and AAFTC (-).	110
Table 4: Most frequent words (top 15) in AASC Agatangelos.	111
Figure 12: Most frequent words (top 15) and their absolute frequencies in AASC Agatangelos.	112
Figure 13: All inflected forms of lexical units based on the root surb/srb- and their absolute frequencies in AASC Agatangelos (bar plot).	113
Table 5: All inflected forms of lexical units based on the root surb/srb- and their absolute frequencies in AASC Agatangelos.	114
Table 6: Most frequent lemmas (top 15) in AASC Agatangelos.	115
Figure 14: Most frequent lemmas (top 15) in AASC Agatangelos (bar plot).	116
Table 7: Top 15 words collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb/srb- in AAFTC Agatangelos.	117
Figure 15: Top 15 types collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb/srb- in AAFTC Agatangelos (graph) (c) LancsBox.	118

Table 8: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb in AASC Agatangelos.	119
Figure 16: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3, no function words) surb in AASC Agatangelos (graph) (c) LancsBox.	120
Table 9: Synoptic table of the AASC and AAFTC Agatangelos.	121
Table 10: Domains in the AAFTC/AASC Agatangelos in a descending order according to their "sum" values in the synoptic table.	123
Figure 17: Keywords and lockwords in the AASC (+) and AAFTC (-).	132
Table 11: English translation keywords and lockwords in the AASC (+) and AAFTC (-).	133
Table 12: Most frequent words (top 15) in AASC Epic Histories.	134
Figure 18: Most frequent words (top 15) and their absolute frequencies in AASC Epic Histories.	135
Figure 19: All inflected forms of lexical units based on the root surb/srb- in AASC Epic Histories (bar plot).	137
Table 13: All inflected forms of lexical units based on the root surb/srb- in AASC Epic Histories.	137
Table 14: Most frequent lemmas (top 15) in AASC Epic Histories.	138
Figure 20: Most frequent lemmas (top 15) in AASC Epic Histories (bar plot).	139
Figure 21: Top 15 types collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb/srb- in AAFTC Epic Histories (graph) (c) LancsBox.	140
Table 15: Top 15 words collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb/srb- in AAFTC Epic Histories.	141
Table 16: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb in AASC Epic Histories.	142
Figure 22: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb in AASC Epic Histories (graph) (c) LancsBox.	143
Table 17a: Synoptic table of the AASC and AAFTC Epic Histories.	144
Table 17b: Synoptic table of the AASC and AAFTC Epic Histories.	145
Table 18: Domains in the AAFTC/AASC Epic Histories in a descending order according to their "sum" values in the synoptic table.	146
Figure 23: Keywords and lockwords in the AASC (+) and AAFTC (-).	154
Table 19: English translation keywords and lockwords in the AASC (+) and AAFTC (-).	154
Table 20: Most frequent words (top 15) in AASC Koriwn.	155
Figure 24: Most frequent words (top 15) and their absolute frequencies in AASC Koriwn.	156
Figure 25: All inflected forms of lexical units based on the root surb/srb- in AASC Koriwn (bar plot).	157
Table 21: All inflected forms of lexical units based on the root surb/srb- in AASC Koriwn. ...	157
Table 22: Most frequent lemmas (top 15) in AASC Koriwn.	158
Figure 26: Most frequent lemmas (top 15) in AASC Koriwn (bar plot).	159

Table 23: Top 15 words collocation analysis (6-LL, L2-R2, C2-NC2, no function words) surb/srb- in AAFTC Koriwn.....	160
Figure 27: Top 15 types collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb/srb- in AAFTC Koriwn (graph) (c) LancsBox.	161
Table 24: Top 15 lemmas collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb in AASC Koriwn.....	162
Figure 28: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb in AASC Koriwn (graph) (c) LancsBox.	163
Table 25: Synoptic table of the AASC and AAFTC Koriwn.	164
Table 26: Domains in the AAFTC/AASC Koriwn in a descending order according to their “sum” values in the synoptic table.	166
Figure 29: Keywords and lockwords (“types”) in the AASC (+) and AAFTC (-).	171
Table 27: English translation keywords and lockwords in the AASC (+) and AAFTC (-).	171
Table 28: English translation keywords and lockwords in the AASC (+) and AAFTC (-).	172
Figure 30: Keywords and lockwords (“lemmas”) in the AASC (+) and AAFTC (-).	173
Table 29: Most frequent words (top 15) in AASC Gospels.	174
Figure 31: Most frequent words (top 15) and their absolute frequencies in AASC Gospels.	175
Table 30: All inflected forms of lexical units based on the root surb/srb- in AASC Gospels. .	176
Figure 32: All inflected forms of lexical units based on the root surb/srb- in AASC Gospels (bar plot).	177
Table 31: Most frequent lemmas (top 15) in AASC Gospels.	178
Figure 33: Most frequent lemmas (top 15) in AASC Gospels (bar plot).	179
Table 32: Top 15 words collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb/srb- in AAFTC Gospels.	180
Figure 34: Top 15 types collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb/srb- in AAFTC Gospels (graph) (c) LancsBox.	181
Table 33: Top 15 lemmas collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb in AAFTC Gospels.	181
Figure 35: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb in AAFTC Gospels (graph) (c) LancsBox.	182
Table 34a: Synoptic table of the AASC and AAFTC Gospels.	184
Table 34b: Synoptic table of the AASC and AAFTC Gospels.	185
Table 35: Domains in the AAFTC/AASC Gospels in a descending order according to their “sum” values in the synoptic table.	186
Table 36: Surb and the corresponding Greek terminology in the Gospel of Matthew.	191
Table 37: Surb and the corresponding Greek terminology in the Gospel of Luke.	192
Table 38: Surb and the corresponding Greek terminology in the Gospel of Mark.	193
Table 39: Surb and the corresponding Greek terminology in the Gospel of John.	193
Figure 36: 15 most common #hashtags in tweets including “holy/holiness.”	205

Figure 37: Absolute frequency of words (no stop words) in tweets that include "holy" or "holiness" (top 20 excluding "holy"). 206

Table 40: Absolute frequency (AF) table based on tweets that include "holy/holiness" (top 20 excluding "holy"). 207

Table 41: Top 15 words collocation analysis (6-LL, 3L-3R, C100-NC100; no function words) HSEC Twitter Data large ("holy"). 208

Figure 38: Top 15 words collocation analysis (6-LL, L3-R3, C100-NC100; no function words) HSEC Twitter Data large ("holy") (graph) (c) LancsBox. 209

Figure 39: 15 most common #hashtags in tweets including "sacred." 210

Figure 40: Absolute frequency of words (no stop words) in tweets that include "sacred" (top 20). 211

Table 42: Absolute frequency (AF) table based on tweets that include "sacred" (top 20). 212

Table 43: Top 15 words collocation analysis (6-LL, 3L-3R, C5-NC5; no function words) HSEC Twitter Data large ("sacred"). 213

Figure 41: Top 15 words collocation analysis (6-LL, L3-R3, C5-NC5; no function words) HSEC Twitter Data large ("sacred") (graph) (c) LancsBox. 214

Table 44: Sketch Engine "Word Sketch" search for "holy" (as noun) top results (selection; top 10 collocates; no scores/frequencies displayed). 219

Table 45: Sketch Engine "Word Sketch" search for "holy" (as adjective) top results (selection; top 10 collocates; no scores/frequencies displayed). 219

Table 46: Sketch Engine "Word Sketch" search for "sacred" (as noun) top results (selection; top 10 collocates; no scores/frequencies displayed). 220

Table 47: Sketch Engine "Word Sketch" search for "sacred" (as adjective) top results (selection; top 10 collocates; no scores/frequencies displayed). 220

Table 48: Most frequent words (top 15) in HSEC Twitter Data (full). 224

Figure 42: Most frequent words (top 15) and their absolute frequencies in HSEC Twitter Data (full). 225

Table 49: Top 15 words collocation analysis "holy" (6-LL, L3-R3, C45-NC45; function words removed) in HSEC Twitter Data (full). 226

Figure 43: Top 15 types collocation analysis "holy" (6-LL, L3-R3, C45-NC45; function words removed) in HSEC Twitter Data (full) (graph) (c) LancsBox. 227

Table 50: Synoptic table of the HSEC Twitter Data (full). 228

Table 51: Most frequent words (top 15) in HSEC Twitter Data (religious). 229

Figure 44: Most frequent words (top 15) and their absolute frequencies in HSEC Twitter Data (religious). 230

Table 52: Top 15 words collocation analysis "holy" (6-LL, L3-R3, C15-NC15; function words removed) in HSEC Twitter Data (religious). 231

Figure 45: Top 15 types collocation analysis "holy" (6-LL, L3-R3, C15-NC15; function words removed) in HSEC Twitter Data (religious) (graph) (c) LancsBox. 232

Table 53: Synoptic table of HSEC Twitter Data (religious). 233

Table 54: Categories in the HSEC Twitter Data (religious) in descending order according to their "sum" values in the synoptic table.....	234
Table 55: Most frequent words (top 15) in HSEC Twitter Data (non-religious).	241
Figure 46: Most frequent words (top 15) and their absolute frequencies in HSEC Twitter Data (non-religious).	242
Table 56: Top 15 words collocation analysis "holy" (6-LL, L3-R3, C35-NC35; function words removed) in HSEC Twitter Data (non-religious).	243
Table 57: Synoptic table of HSEC Twitter Data (non-religious).	244
Table 58: Categories in the HSEC Twitter Data (non-religious) in descending order according to their "sum" values in the synoptic table.	245
Figure 48: Keywords and lockwords in the HSEC Canon Law (only sentences incl. "holy") (+) and HSEC Canon Law (full) (-).	248
Table 59: Most frequent words (top 15) in HSEC Canon Law (full).	250
Figure 49: Most frequent words (top 15) and their absolute frequencies in HSEC Canon Law (full).	251
Table 60: Most frequent words (top 15) in HSEC Canon Law (holy).	251
Figure 50: Most frequent words (top 15) and their absolute frequencies in HSEC Canon Law (holy).	252
Table 61: Top 15 words collocation analysis "holy" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full).	254
Figure 51: Top 15 types collocation analysis "holy" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full) (graph) (c) LancsBox.	255
Table 62: Top 15 words collocation analysis "sacred" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full).	256
Figure 52: Top 15 types collocation analysis "sacred" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full) and its overlappings with "holy" (graph) (c) LancsBox.	257
Table 63: Synoptic table of the HSEC Canon Law.	258
Table 64: Categories in the HSEC Canon Law in descending order according to their "sum" values in the synoptic table.	259
Table 65: Most frequent words (top 15) in HSEC Encyclopaedia Britannica.....	264
Figure 53: Most frequent words (top 15) and their absolute frequencies in HSEC Encyclopaedia Britannica.	265
Figure 54: Top 15 types collocation analysis "holy" (6-LL, L4-R4, C1-NC1; function words removed) in HSEC Encyclopaedia Britannica (graph) (c) LancsBox.	266
Table 66: Top 15 words collocation analysis "holy" (6-LL, L4-R4, C1-NC1; function words removed) in HSEC Encyclopaedia Britannica.	267
Table 67: Top 15 words collocation analysis "sacred" (6-LL, L2-R2, C2-NC2; function words removed) in HSEC Encyclopaedia Britannica.	268
Figure 55: Top 15 types collocation analysis "sacred" (6-LL, L2-R2, C2-NC2; function words removed) in HSEC Encyclopaedia Britannica (graph) (c) LancsBox.	269

Table 68: Synoptic table of the HSEC Encyclopaedia Britannica.	270
Table 69: Categories in the HSEC Encyclopaedia Britannica in descending order according to their "sum" values in the synoptic table.	272
Table 70: Most frequent words (top 15) in HSEC Eliade.	278
Figure 56: Most frequent words (top 15) and their absolute frequencies in HSEC Eliade.	279
Table 71: Top 15 words collocation analysis "holy" (6-LL, L2-R2, C1-NC1; function words removed) in HSEC Eliade.	279
Figure 57: Top 15 types collocation analysis "holy" (6-LL, L2-R2, C1-NC1; function words removed) in HSEC Eliade (graph) (c) LancsBox.	280
Table 72: Top 15 words collocation analysis "sacred" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Eliade.	282
Figure 58: Top 15 types collocation analysis "sacred" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Eliade (graph) (c) LancsBox.	283
Table 73: Synoptic table of the HSEC Eliade.	284
Table 74: Categories in the HSEC Eliade in descending order according to their "sum" values in the synoptic table.	285
Table 75: Most frequent words (top 15) in HSEC Otto.	291
Figure 59: Most frequent words (top 15) and their absolute frequencies in HSEC Otto.	292
Table 76: Top 15 words collocation analysis "holy" and "holiness" (6-LL; L2-R2; C3-NC3; function words removed) in HSEC Otto.	294
Figure 60: Top 15 types collocation analysis "holy" and "holiness" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Otto (graph) (c) LancsBox.	295
Table 77: Top 15 words collocation analysis "numinous" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Otto.	295
Figure 61: Top 15 types collocation analysis "numinous" (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Otto (graph) (c) LancsBox.	296
Table 78: Synoptic table of the HSEC Otto.	298
Table 79: Categories in the HSEC Otto in descending order according to their "sum" values in the synoptic table.	299
Table 80: Table of the "domain" section in the AAFTC and AASC (min-max normalized).	307
Table 81: Table of the "domains" with the number of corpora in which the domains appear (AAFTC/AASC).	308
Figure 62: Visualization of the "domain" graph of the Armenian texts in the AAFTC and AASC; Fruchterman Reingold Graph rendered with (c) Gephi.	308
Table 82: Overall table of the "domain" section of the HSEC (min-max normalized).	309
Table 83: Table of the "domains" with the number of corpora in which the domains appear (HSEC).	309
Figure 63: Visualization of the domain graph of the texts in the HSEC; Fruchterman Reingold Graph rendered with (c) Gephi.	310

Table 84a: Table including the top 15 words plus their corresponding “domains” in the AAFTC and AASC data.	310
Table 84b: Table including the top 15 words plus their corresponding “domains” in the AAFTC and AASC data.	311
Figure 64: Visualization of the “domain” graph of the Armenian texts in the AAFTC and AASC including all words and domains; Fruchterman Reingold Graph rendered with (c) Gephi.	311
Table 85: Table including the top 15 words plus their corresponding domains in the HSEC data.	312
Figure 65: Visualization of the “domain” graph of the English texts in the HSEC; Fruchterman Reingold Graph rendered with (c) Gephi.	313
Figure 66: Visualization of the “domain” graph of the Armenian texts in the AAFTC/AASC, including all words and “domains” that appear in all subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi.	314
Table 86a: Table including the top 15 words plus their corresponding “domains” in the AAFTC/AASC data that appear in only one subcorpus.	314
Figure 67: Visualization of the “domain” graph of the English texts in the HSEC, including all words and “domains” that appear in all subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi.	315
Table 86b: Table including the top 15 words plus their corresponding “domains” in the AAFTC/AASC data that appear in only one subcorpus.	315
Figure 68: Visualization of the “domain” graph of the Armenian texts in the AAFTC/AASC, including all words that appear in only one subcorpus; Fruchtermann Reingold Graph rendered with (c) Gephi.	316
Table 87a: Table including the top 15 words plus their corresponding “domains” in the HSEC data that appear in only one subcorpus.	316
Table 87b: Table including the top 15 words plus their corresponding “domains” in the HSEC data that appear in only one subcorpus.	317
Figure 69: Visualization of the “domain” graph of the HSEC, including all words that appear in only one subcorpus; Fruchtermann Reingold Graph rendered with (c) Gephi.	317
Table 88: Table with top 15 entries of the “semantic field” section in the AAFTC/AASC data (min-max normalized).	319
Table 89: Table with top 15 entries of the “semantic field” section and the numbers of subcorpora they appear in.	320
Figure 70: Visualization of the “semantic field” graph (no words) of the Armenian texts in the AAFTC/AASC; Fruchterman Reingold Graph rendered with (c) Gephi.	321
Table 90: Table with top 15 entries of the “semantic field” section of the HSEC (min-max normalized).	322

Table 91: Table with top 15 entries of the “semantic field” section and the numbers of subcorpora they appear in. Note that the Z9 semantic field “not defined” was dropped from the table. 323

Figure 71: Visualization of the “semantic field” graph (no words) of the English texts in the HSEC; Fruchterman Reingold Graph rendered with (c) Gephi. 324

Table 92: Table including the top 15 words plus their corresponding “semantic fields” in the AAFTC/AASC data. 325

Figure 72: Visualization of the “semantic field” graph of the Armenian texts in the AAFTC/AASC including all words and “semantic fields;” Fruchtermann Reingold Graph rendered with (c) Gephi. 326

Table 93a: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data. 327

Table 93b: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data. 327

Figure 73: Visualization of the “semantic field” graph of the English texts in the HSEC including all words and “semantic fields;” Fruchtermann Reingold Graph rendered with (c) Gephi. 328

Figure 74: Visualization of the “semantic field” graph of the Armenian texts in the AAFTC/AASC including all words and “semantic fields” that appear in all subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi. 329

Table 94: Table including the top 15 words plus their corresponding “semantic fields” in the AAFTC/AASC data. 329

Figure 75: Visualization of the “semantic field” graph of the English texts in the HSEC including all words and “semantic fields” that appear in all three subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi. 330

Table 95a: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data. 330

Figure 76: Visualization of the “semantic field” graph of the Armenian texts in the AAFTC/AASC including all words and “semantic fields” that appear in only one subcorpus; Fruchtermann Reingold Graph rendered with (c) Gephi. 331

Table 95b: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data. 331

Figure 77: Visualization of the “semantic field” graph of the English texts in the HSEC including all words and “semantic fields” that appear in only one subcorpus; Fruchtermann Reingold Graph rendered with (c) Gephi. 332

Figure 78: Intersection of “domains” between the HSEC and the AAFTC/AASC. (c) Gephi. ... 342

Table 96: Comparison of “domain” values between the HSEC and AAFTC/AASC according to their normalized “sum” values. 342

Table 97: Comparison of “semantic field” values between the HSEC and AAFTC/AASC according to their normalized “sum” values. 343

Figure 79: Intersection of “semantic fields” between the HSEC and the AAFTC/AASC. (c) Gephi.	344
Table 98: Intersection of words and their corresponding “domains” and “semantic fields” between the HSEC and the AAFTC/AASC.	345
Figure 80: Intersection of words and their “semantic fields” between the HSEC and the AAFTC/AASC. Different colors indicate different “domains.” (c) Gephi.	346
Figure 81: Visualization of the connections between surb (AAFTC/AASC) and holy (HSEC) when compared to the notion of holy in the study of religion as a tertium comparationis (see Chapter 2).	360
Figure 82: Visualization of the few but central overlappings between the three semantic fields from a quantitative perspective.	367

Acknowledgements

I would like to thank my two supervisors, Professor Dr. Volkhard Krech and Dr. Zara Pogossian, for the continuous and intensive support they have shown during my doctoral project. I am particularly grateful to them for their openness with regard to my methodological approaches, which they have always accepted and engaged with regardless of their diverging preferences. Generally, it has been a privilege to be able to develop my project within the Center for Religious Studies (CERES), which has provided a stimulating environment and where my position as a research assistant allowed me to dedicate sufficient time to my research. Its financial support, as well as that of the Research School PLUS (Ruhr-University Bochum), has enabled this project to come to fruition.

I am also indebted to Dr. Marion Steinicke and Professor Dr. Heinz-Georg Held, who have largely contributed to my decision to take up a doctoral project. The intensive discussions with them over the past few years have repeatedly opened up unexpected, sometimes controversial, but always enriching perspectives.

I would also like to thank my parents, Ms. Birgit Jurczyk and Mr. Jürgen Jurczyk, who have always supported and accompanied me throughout the long years of my studies. Thanks are also due to Jessie Pons, who proofread my book trying to improve both its language and content. This was certainly not always easy, and any mistakes remain my own.

Finally, thank you, Charlotte, for making sure that the process of writing my dissertation always remained exciting.

Bochum, December 2021

Thomas Jurczyk

1. Introduction

This book “The Notion of ‘holy’ in Ancient Armenian Texts from the Fifth Century CE”¹ aims at contributing to the broader discussion in the study of religion on the comparative notion of *holy*² beyond religious, cultural, and linguistic boundaries. Notions such as *holy* are of crucial importance for the study of religion as they enable the comparison and hence study of different religious traditions based on common *tertia comparationis*. With this in mind, this book compares the meaning and application of two notions and their related word fields that are commonly associated with a broader comparative notion of *holy*. Both notions stem from different linguistic and cultural backgrounds. The first notion is represented by the Ancient Armenian term *surb*, a word that is often rendered as “holy” in English translations of Ancient Armenian texts. The second notion is represented by the English term “holy.”³

-
- 1 Which is a slightly adopted version of my dissertation “The Notion of *surb* in Ancient Armenian Texts from the Fifth Century CE” (defended in September 2020 at the Center for Religious Studies, Ruhr-University Bochum).
 - 2 Note that the use of italics indicates a focus on the general notion and its related semantic field rather than on the actual word. The general notion might include derivatives, inflected forms, as well as synonyms. For instance, in the case of the general notion of *holy*, different words such as “sacred,” “holiness,” or “sanctified” are included. When talking about the particular word “holy” as it appears in a text, quotation marks are used. Note that italics are also used for non-English (transliterated) words (such as *sacer* or *hâgios* or *surb*), proper nouns (such as *LancsBox*), and encyclopedia or publication series titles (such as *Religion in Past & Present*).
 - 3 In this book, the term **notion** refers to an abstract idea or concept. The term **semantic field** is used interchangeably with **lexical field** and **word field** and denotes “(...) a semantic relationship between lexemes which relates to a particular subject or domain rather than to the denotation of the lexeme” (entry “lexical field” in *The Oxford Companion to English Language* (2017)). For example, the words “sacred,” “holy,” and “saint” are usually assumed to belong to the semantic field of the notion of *holy*. The word **terminology** is used with its general meaning of denoting “(...) a set of (technical) terms on a specific topic or in a specific field” (entry “terminology” in *International Encyclopedia of Linguistics* (Frawley and Bright 2003)). Thus, the words “holy,” “sacred,” and “saint” can be regarded as belonging to the terminology related to *holy*. The expressions **term** and **word** are used interchangeably. The mentioning of words

The choice of Ancient Armenian and English notions was guided by the following reasons. Firstly, the examination of Ancient Armenian sources and the use of *surb* therein are still a *desideratum* in the broader discussions about a comparative notion of *holy* in the study of religion, at least compared to other ancient notions such as Latin *sacer/sanctus* or Greek *hágios*.⁴ Consequently, a detailed analysis of the Armenian terminology adds new aspects to both the broader discussion of the comparative notion of *holy* and the application and meaning of words related to *holy* in ancient societies.

Secondly, the two notions of *surb* and *holy* stem from different languages with different cultural and historical backgrounds. Therefore, both notions and their respective semantic fields are relatively independent of each other. Consequently, potential overlappings between these semantic fields can rightfully be interpreted as an indication of an existing comparative notion or semantic field of *holy* beyond cultural, historical, and linguistic boundaries.

Lastly, despite the long time gap between the languages and the corresponding cultural backgrounds, the notions of *surb* and *holy* are related to one another. The relation between these two notions is notably perceivable in their common application in the translation of the Christian terminology of *holy* in the Bible. Furthermore, contemporary English translations of Ancient Armenian texts tend to translate the notion of *surb* and its related word field with words from the semantic field of the English notion of *holy*. These links between *surb* and *holy* illustrate that a comparison of both concepts is not entirely arbitrary but based on empirical data. Nevertheless, the assumption implicitly made by the translators that both notions correspond to each other still needs to be examined. This is one major aim of this book.

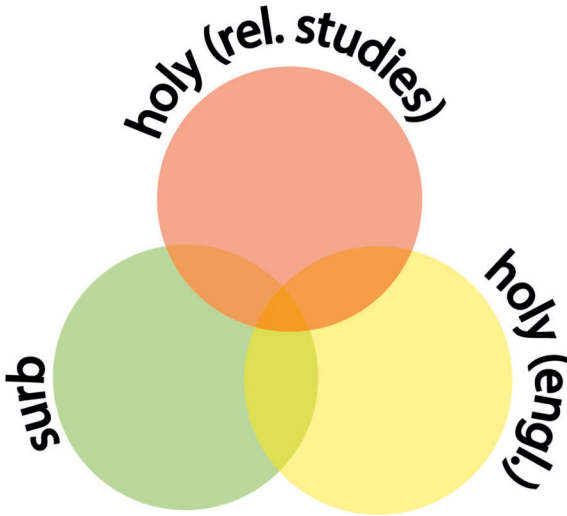
The overall research question is divided into two interrelated research foci. The first focuses on the specific semantic problem of the meaning of the Armenian notion of *surb* in early Armenian texts from the fifth century CE. This first part exclusively concentrates on the examination of the Armenian texts and does not take into account the use and meaning of *holy* in the English sources. The second research focus undertakes a comparative study by relating the word field around *surb* to the contemporary English notion of *holy*. Furthermore, the two notions of Armenian *surb* and English *holy* are evaluated in view of the comparative notion of *holy* in the study of religion. The inquiries are based on macro- and microanalyses

such as “holy” usually includes all derivatives and words that are based on the (root-)word. Consequently, when referring to “holy,” other words such as “holiness” or “holies” are implied.

4 For instance, the following overviews of the application and meaning of *holy* ignore the Armenian terminology: Gemeinhardt and Heyden (2012); Colpe (1988a); Dihle (1988); Lanczkowski (1980); Wokart (2018). Other articles that also include a discussion of the Armenian terminology usually do this from an etymological or very general point of view: Morani (2001); Stempel (1988).

of Ancient Armenian and contemporary English text corpora that apply digital tools and methods from the field of corpus linguistics.⁵

Figure 1: Potential overlappings between the three semantic fields examined in this book.



The methodological approach in this book operates on the underlying principle that the meaning of words and notions can only be derived from their semantic, syntactical, and pragmatic contexts. Accordingly, a lexicographical⁶ approach based on fixed meanings, synonyms, and etymological explanations is rejected in favor of an approach which seeks to avoid any presuppositions in meaning. One such premise is as that *surb* means “holy, sacred; pure, clean, exempt, spotless, stainless (...)” simply because it is commonly found that way in dictionaries such as Petrosian’s *Նոր Բառգիրք Հայ-անգլիարէն* (Venice, 1875). Instead, the meaning of both the Ancient Armenian terminology around *surb* and the contemporary English notion of *holy* should be derived from their use cases. The application of tools and methods from the field of corpus linguistics is an adequate choice to obtain a thorough overview of the syntactical and semantic contexts of both notions as they allow the handling and examination of large text corpora.

5 For a detailed overview of the corpora and methods, see the corresponding Chapter 2 and Chapter 3.

6 Adler (1992). Others would even speak of a “Lexikokratie” to underline the authoritative character of these approaches, see the corresponding entry “Lexikokratie” in Fuchs-Heinritz (2013).

1.1 The Importance of Comparative Notions for the Study of Religion

The importance of comparative notions such as *holy* for the study of religion, to which this book attempts to contribute, has been stressed by Oliver Freiberger in his book “Considering Comparison. A Method for Religious Studies” (2019). Besides Freiberger’s convincing plea for the need for comparative notions in the study of religion, his book also provides a concise framework of how to evaluate and locate comparative studies based on their **scale**, **mode**, and **scope**.

In accordance with the methodological framework that Freiberger presents in his fourth chapter (Freiberger 2019, 111ff.), the aim of the comparative study in this book corresponds to the taxonomic mode (Freiberger 2019, 126ff.). This **mode** is preoccupied with “classify[ing] religious items and thus contribut[ing] to the taxonomic effort in the study of religion” (Freiberger 2019, 127). The **scale** of the study in this book differs depending on the size of the texts in the respective corpora.⁷ Yet, the overall focus lies on the meso- or even macro-level (Freiberger 2019, 131ff.). Regarding the **scope** of a study, Freiberger proposes three different categories.

Generally put, the category “scope” reflects the distance between the items compared in a study. Studies with a contextual scope compare items within one historical context or cultural milieu that can be delineated both spatially and temporally—for example, the Mediterranean world in late antiquity, northeast India in the fifth century BCE, or contemporary Brazil. Studies with a cross-cultural scope go beyond postulated cultural boundaries, such as in a comparison of ancient Chinese and ancient Greek religion. Studies with a trans-historical scope are comparisons across time and always appear in conjunction with one of the other two scopes. (Freiberger 2019, 143)

Taking up these three scopes, the comparative approach in this book has a cross-cultural and trans-historical scope. Yet, despite their different cultural and historical backgrounds, the contemporary English notion of *holy* is still linked to the Ancient Armenian notion due to translations, among others.

1.2 Overall Structure of This Book

Following the introduction, Chapter 2 provides a brief overview of the historical background of Armenia in the fifth century CE, the comparative notion of *holy* in the study of religion, and selected examples of the notion of *holy* in other ancient

7 See Chapters 3 and 4 for more information.

languages. Chapter 3 offers a more detailed articulation of the methodological approach of this book that is largely based on corpus linguistics and influenced by ideas of distributional semantics. Chapter 4 provides an overview of the texts used in the examination. The texts range from collections of short texts (tweets) to large-sized text corpora (*English Web corpus 2015* [*enTenTen2015*]) and can be grouped into three overall corpora.

The first corpus is the Armenian text corpus, including the *Ancient Armenian Surb Corpus* (AASC) and the *Ancient Armenian Full Text Corpus* (AAFTC). Both corpora are based on the same texts, namely “The History of the Armenians” by Agat’angelos, the “Life of Maštoc’” by Koriwn, the “Epic Histories,” and the Gospels in their Armenian translation according to the so-called Zohrab Bible. The English corpus is divided into two subcorpora of very different sizes. Firstly, there is the *Holy/Sacred English Corpus* (HSEC), which includes text samples from different genres (academic texts, religious texts, encyclopedia entries, tweets). Secondly, this book uses the already existing *English Web corpus 2015* (*enTenTen2015*) corpus with the help of *Sketch Engine*. The *English Web corpus 2015* (*enTenTen2015*) includes more than 15 billion words of contemporary English texts that have been scraped from the Internet.

The examination part of this book starts in Chapter 5. This chapter analyzes the semantic fields of the English notion of *holy* and the Armenian notion of *surb*. The examination is based on the text corpora and the methods introduced in Chapters 3 and 4.

The results of the examination in Chapter 5 are used in Chapter 6 for the overall comparison of the different semantic fields. This chapter also discusses the final results of the comparisons in view of the two previously mentioned research foci. Thus, Chapter 6 will propose a concise answer to the question: What is the meaning of *surb* in Ancient Armenian texts from the fifth century CE? Furthermore, it will also evaluate whether the semantic fields of the English notion of *holy* and the Ancient Armenian notion of *surb* are linked to each other and if so, what implications this might have for the comparative notion of *holy* in the study of religion (Freiberger 2019, 158ff.).

2. State of Research

This second chapter serves as the prolegomena to my study. It sketches the historical background of Armenia in the fifth century CE, the meaning and use of the notion of *holy* in the study of religion, the meaning of selected examples of ancient terminologies related to *holy*, and the theoretical background of distributional semantics. These preliminary remarks will focus on aspects deemed essential for understanding the examinations conducted in Chapters 5 and 6 due to the following reasons.

The historical overview of Armenia in the fifth century CE allows contextualizing the Armenian text corpora. The outline of the meaning and use of the notion of *holy* in the study of religion and ancient societies is important for the analysis of the Armenian vocabulary and contemporary layers of meaning. This section will help to link the results of the analyses in Chapters 5 and 6 to the debates on *holy* as a comparative notion in the study of religion as well as to broader discussions about *holy* in ancient societies.

The last overview reviews the general ideas underlying corpus linguistics and distributional semantics. This part will address the methods and the approaches used in this book from a broader methodological perspective. The detailed discussion of the methods, however, follows in Chapter 3.

2.2 History of Armenia in the Fifth Century CE

The history of Armenia in the fifth century CE has to be examined in a slightly larger context, one that also takes into account the centuries shortly before and after the fifth century CE.¹ In this historical context, “Armenia” is used as a collective term for several regions in the (Sub-)Caucasian area and thus covers an area that is much larger than the territory of the current Republic of Armenia.

1 The entire chapter on the “History of Armenia in the Fifth Century CE” is heavily based on the works by Nina Garsoïan, Robert Thomson, and Jean-Pierre Mahé, see Mahé and Mahé (2012); Thomson (1994); Garsoïan and Mahé (1997); Garsoïan, Mathews, and Thomson (1982); Garsoïan (1999).

The history of Armenia goes back to, at least, the first millennium BCE. For the purpose of this book, however, the time frame from the fourth to sixth centuries CE is of particular importance. During the fourth and fifth centuries CE, Armenia underwent a long process of Christianization that also resulted in the invention of the Armenian script with its own alphabet (early fifth century CE). The invention of the Armenian alphabet triggered both extensive translation processes from other languages into Armenian, most prominently the translation of the Bible, as well as the formation of original Armenian literature.² Selected sources from this early period of Armenian literature of the so-called “Golden Age” (fifth to sixth centuries CE) also build the empirical basis for the examination of *surb* in this book.³ As has been stated in the introduction, the research question of this book is not primarily concerned with a typical historical or text-critical question. Rather, it deals with a semantic question that compares notions from different languages and time periods. In order to be able to explore the semantic field of *surb* and to better understand its cultural and religious embedding, the following sections will provide an overview of the crucial political, military, and religious developments during the period in question. Therefore, the overview will introduce the circumstances under which the Armenian texts and the terminology therein emerged.

The following overview is split into two parts. The first part deals with the political and military events. The second part sketches the formation of the (later autocephalic) Armenian Church and her internal developments. Even though both parts are discussed separately, they were closely intertwined.

A telling example of this entanglement are the most likely legendary events leading to the official Christianization of Armenia in the early fourth century CE. The inception of the official Christianization in Armenia is usually linked to the events leading to the conversion of the Armenian king of Parthian origin Trdat IV (early fourth century CE). According to Agat'angelos' “History of the Armenians” (the primary source for these legendary events), Trdat IV was converted to Christianity by Gregory the Illuminator, the first head of the Armenian Church, after

2 In this context, the term “literature” is restricted to written works and does not include oral traditions that existed long before the invention of the Armenian script in the fifth century CE.

3 Even though the earliest surviving testimonies and manuscripts are much later, the composition of the Armenian text sources used in this book is commonly dated to the fifth century CE. However, as the earliest surviving Armenian manuscripts date to the ninth century CE, a definite dating of early Armenian texts is always tricky. As an example, the earliest complete witnesses of historical accounts such as Agat'angelos' “History of the Armenians” are often much later (mostly after 1600 CE). For a concise overview of Armenian manuscripts and text witnesses, see Kouymjian (2015). For an overview of the manuscripts and text witnesses of Agat'angelos' “History of the Armenians,” see Thomson (1976, 504ff.).

a long chain of events. These include many topoi typically found in early Christian conversion narratives such as refusal, persecution, sin, discernment, healing, conversion, etc. A more detailed account of the content of the “History of the Armenians” by Agat’angelos follows in Chapter 4 of this book. The turning of the Armenian king towards Christianity is supposedly the first official conversion of an ancient pagan king, preceding that of Constantine the Great (whose conversion is still a much-debated topic). It is also an illustrative example of the entanglement of politics, religion, and the military in the missionary activities that successively led to the Christianization of Armenia. The conversion to Christianity not only resulted in the self-perception of ancient Armenia as a Christian state in the eyes of ancient Armenian writers such as Ehiše,⁴ but it also serves until today as an important religious and a political identification marker for Armenians.

The strong impact of this Christianization process on the political and ethnic identity of Armenians is summarized by Nina Garsoïan and Jean-Pierre Mahé as follows:

En Arménie et dans les pays de diaspora arménienne, la religion devient d’autant plus facilement le support de la culture nationale que l’invention de l’alphabet et l’éclosion d’une littérature sont liées à la christianisation. L’Église est aussi le noyau dur de l’organisation sociale; elle acquiert tout naturellement un rôle structurant dans un pays de grande instabilité politique et dans une société qui connaît très tôt une sorte d’éclatement féodal. (Garsoïan and Mahé 1997, 10:6)

2.2.1 Political and Military History

Due to the geographical location of Armenia, its history was often influenced by events in the Roman and in the Persian world. For a long time, scholars had mostly concentrated on the occidental Roman influence on the history of Armenia. Since the important contributions by Nina Garsoïan, however, much more attention has been placed on the South-Eastern territories and their influence on the history of Armenia and the Armenian Church.⁵

4 For the transliteration of Armenian terms and names, I applied the Hübschmann-Meillet transliteration system available under this link to the *Hübschmann-Meillet transliteration system*: <https://www.transliteration.com/transliteration/en/armenian-eastern-classical/hubscmann-meillet/> (last accessed: 10/28/2021) (the full link is also provided in the link section “Internet Resources”). Yet, I have decided to make small changes; for instance, instead of transliterating diphthongs such as “nl” as “ow,” I usually use “u.”

5 Among others, see Garsoïan, Mathews, and Thomson (1982); Garsoïan and Mahé (1997); Garsoïan (1999; 2004a).

2.2.1.1 Geographical Overview

The region of Armenia in late antiquity did not form an administrative unit and was divided into two different sectors since the fourth century CE (Garsoïan 2004a, 80). The smaller part was called *Armenia Minor* and was situated West of the Euphrat. Throughout the fourth and fifth centuries CE, this area was mostly part of the Roman administration and later, during the reign of Theodosius the Great, further subdivided into *Armenia I* and *Armenia II*. The administrative and religious center of *Armenia I* was Sebaste and that of *Armenia II* Melitene. According to Nina Garsoïan, *Armenia Minor* remained outside of the jurisdiction of the Armenian Church. That being said, church officials from *Armenia Minor* also participated in important councils such as that in Chalcedon (451 CE),⁶ whereas the Armenian Church in *Greater Armenia* received the *kanones* of Chalcedon much later and eventually rejected them in the seventh century CE.⁷

From a political and religious standpoint, the Southern territories of Armenia posed a problem in themselves. Some of the Southern territories such as Sophene, Sophanene, Ingilene, or Anzitene were closely connected to the Greco-Roman world, whereas others such as Ajjnik', Korduk', Cawdek, Mokk', Mahk'ertun, and Dasn maintained close ties with the Mesopotamian realm (Garsoïan 2004a, 81).

The Eastern part of Armenia was called *Greater Armenia* (or *Persian Armenia*) and included almost 80% of the overall territory of ancient Armenia. The territory of *Greater Armenia* can rightfully be regarded as the center of most developments in the ancient Armenian history and particularly the ancient Armenian Church history. *Greater Armenia* was heavily influenced by the Parthian and later Sasanian world due to centuries-old religious, commercial, political, and military exchange processes. Among others, the Persian influence is evident in the long reign of Armenian kings of Parthian origin in *Greater Armenia*, who ruled until 428 CE and belonged to the Parthian Arsacid dynasty. *Greater Armenia* even became a temporary Sasanian province after the deposition of the last Armenian Arsacid king Aršak in 428 CE.

2.2.1.2 Historical Overview

The first historical event that had a great influence on the history of Armenia from the fourth to sixth centuries CE was the coming into power of the Sasanians in the early third century CE. This led to continuing conflicts with the Roman empire. In 279 CE, the Sasanian Empire had to agree to a peace treaty with the Roman empire. After the so-called Peace of Nisibis in 298 CE (that went in favor of the Romans), Trdat IV became king of Armenia. According to Agat'angelos, Trdat IV had beforehand survived the assassination of his family and had also sought for refuge

6 The decisions made in Chalcedon caused long-lasting controversies about the nature of Christ, among others in the Armenian Church.

7 For the details of this controversy, see the section on the "Armenian Church History" below.

in the Roman empire. Trdat IV was later converted to Christianity by Gregory the Illuminator, with whom he initiated the Christianization process of Armenia.⁸

The coming into power of the Sasanians in Iran was of central importance for the previously mentioned developments in Armenia due to several reasons. Firstly, the kings in Arsacid Armenia were of Parthian origin and thus belonged to the dynasty that had been overthrown by the Sasanians in the Persian realm. However, the Arsacid kings in Armenia remained in power until 389 CE in *Armenia Minor* and until 428 CE in *Greater Armenia*. Secondly, the rise to power of the Sasanians went hand in hand with a strong focus on a religious reformation program that attempted to re-establish the traditional Persian belief of Zoroastrianism. This reformation program was not only executed in the Persian heartland but also in Armenia (Mahé and Mahé 2012, 76). The pressure exerted by the new Sasanian rulers in Armenia lasted throughout the fourth and fifth centuries CE and led to several military conflicts, the invasion of Armenia by the Sasanians between 450–451 CE, and the famous battle of Aravayr in 451 CE. It is during this period that Mesrop Maštoc' is said to have invented the Armenian script. This invention instigated important translation processes from other languages into Armenian and the creation of original Armenian literature.⁹ The conflicts with the Sasanian empire also led to the absence of bishops from *Greater Armenia* in the famous Council of Chalcedon in 451 CE. This, in turn, was one central reason for the later schism between the Armenian and Chalcedonian Churches such as the Church in Constantinople. Even though the Armenians lost the battle at Aravayr, the lasting resistance and constant uprisings in Armenia eventually led to the political and religious independency of *Greater Armenia* in 485 CE. However, political and often also military conflicts with both the Sasanian and Eastern Roman empires remained a crucial factor during the following centuries. These conflicts were not only often driven by religious issues, among other motivations, but they also impacted on religious developments in Armenia.

2.2.1.3 Political Structure

Two factors dominated the political structure of Arsacid Armenia during the centuries in question. On the one hand, traditional kings from the Arsacid dynasty continued ruling as kings in *Greater Armenia* until 428 CE. On the other hand, powerful local rulers and lords with their respective families, clans, and armed contingents exerted a significant influence, too. These clans continued to have an important impact on both the political and religious history of Armenia and often stood

8 Whether it was 301 CE or 314 CE is of no importance for the topic of this book. For a critical overview of the proposed dates, see Chaumont (1969).

9 For more information, see Chapter 4 and the description of Koriwn's "The Life of Maštoc'" in particular. For the overall importance of the invention of the Armenian script for the "Armenian identity," see Calzolari (2014, 373ff.).

in opposition to the ruling Arsacids. A case in point are the Mamikoneans, a very influential clan that plays a major role in the sources discussed in this book, such as the “Epic Histories.” Consequently, Armenia in late antiquity cannot be regarded as a state with one particular center. Instead, ancient Armenia was a mosaic of different regions with their respective local rulers who could, in an ideal case, unite under the rulership of a single king, but who were usually in conflict against each other (Mahé and Mahé 2012, 79ff.).

2.2.2 Church History

Today, the Armenian Church counts among the Oriental Orthodox Churches.¹⁰ The historical process of the gradual separation of the Armenian Church from other Churches resulting in her eventual autocephaly lasted for several centuries (Garsoïan and Mahé 1997, 10:6).

The following outline of the Armenian Church history is relevant to understand the complex historical and religious context of the Armenian sources in this book. This section will focus on the overall Church history and particularly highlight the pivotal position of the Armenian Church between the Eastern Roman Church and the Syriac/Persian Churches. It will also briefly discuss the Armenian Church's reception of important early Christian councils such as that of Chalcedon in 451 CE and their impact on the overall development of the Armenian Church.

2.2.2.1 Pre-Christian Cults in Armenia

Even before the official Christianization of Armenia in the fourth century CE, the ancient non-Christian cults and deities in Armenia revealed many overlappings with both Persian as well as Roman/Eastern Mediterranean religious traditions (Garsoïan and Mahé 1997, 10:14; Russell 1990). Even though only little is known about pre-Christian cults and myths in Armenia, the Armenian pantheon included several important gods such as Tir/Apollon, Anahita/Artemis, Vahagn/Herakles, Barshamin, Aramazd/Zeus and many more that were mostly also known to other regions in the neighborhood of Armenia (Mahé and Mahé 2012, 77). Most information on these pre-Christian cults, myths, and deities stem from later Armenian sources that were all written down after or during the Christianization of Armenia,

¹⁰ When speaking of the Armenian Church, the Armenian Apostolic Church of Armenia, with her see in Ējmiacin, is meant. However, there are other Armenian Churches as well, for instance, a Church Union with the Catholic Church. For an overview of the different Armenian Churches, see Oeldemann (2011).

such as “The History of the Armenians” by Agat’angelos or “The History of Armenia” by Movsēs Xorenac’i.¹¹

A crucial factor for the religious and political developments in Armenia was the reform Zoroastrianism of the Sasanian Empire. In terms of religious changes, the often aggressive Sasanian attempts to reform Zoroastrianism in the Persian and in the Armenian realm had important effects on the overall religious landscape of these regions. In Armenia, the Sasanian aggressions led to a growing focus on a trinity from the wider Armenian pantheon consisting of Aramazd, Anahita, and Vahagn. According to Annie and Jean-Pierre Mahé, the hostile conversion policy of the Sasanian Empire may have even caused the official Christianization of Armenia:

Tiridate avait plus d'une raison pour faire du christianisme la religion officielle de son royaume. Même simplifié, le mazdéisme syncrétique des Arsacides restait vulnérable aux critiques du monothéisme zoroastrien des Sassanides, qui avaient gagné à leur cause toute la noblesse arménienne. Le seul moyen de s'y opposer était d'instituer un monothéisme encore plus exigeant, où le Dieu unique domine non seulement toute la création, qui est intrinsèquement bonne, mais convertit au bien la liberté humaine, qui est à l'origine du mal. (Mahé and Mahé 2012, 79)

2.2.2.2 Christianization of Armenia

The account of the Christianization of Armenia is usually twofold. It considers both the Southern Syriac and the North-Western Greco-Roman influences on the Christianization of Armenia. The first series of narratives is concerned with the likely legendary apostles Thaddeus and Bartholomew, who initiated the early evangelization of the Southern provinces of Armenia in the second century CE.¹² This line of tradition represents the Syriac influence on the Christianization of Armenia and is usually dated much earlier than the later conversion of the Arsacid rulers of Armenia. These stories are, to a great extent, legendary accounts that refer to other Christian narratives, for instance, to the legendary story of the missionary activities of Thaddeus in Edessa (Illert 2007). However legendary these accounts are, Christian missionary activities in Southern Armenia are likely to have taken place at such an early point.

The second wave of missionary activities in Armenia is connected to Trdat IV and his conversion by Gregory the Illuminator in the early fourth century CE (see also the introduction of this subchapter). These accounts play a major role in the “History of the Armenians” by Agat’angelos, which is one of the primary sources

11 The latter is another important historical account written in Ancient Armenian that most likely stems from the ninth or tenth century CE. Among others, see Thomson (2006); Garsoïan (2004b; 2000).

12 For the legendary appearance of Thaddeus and Bartholomew in Armenia, see Calzolari (1997); van Esbroeck (1983).

in this book.¹³ Similarly to the missionary activities by Thaddeus and Bartholomew in Armenia, the conversion plot around Trdat IV also includes many unhistorical and legendary elements. Yet, despite the legendary and hagiographical character of this story, most scholars agree on the historical veracity of the conversion and on its early date (Mahé and Mahé 2012, 79). The conversion story around Gregory the Illuminator and Trdat IV is often regarded as deriving from the Roman sphere and from Cappadocia in particular (Garsoïan 2004a, 82). According to the legend told by Agat'angelos, Gregory the Illuminator became the first bishop (later called *katholikos*) of Armenia. He established a lineage of bishops in Armenia that lasted until 435 CE. His first successor was his son Aristakes who also participated in the Council of Nicea in 325 CE.

2.2.2.3 Armenia in Christian Times (Fourth to Seventh Centuries CE)

Most of the patriarchs of *Greater Armenia* during the fourth century CE were consecrated in Cappadocia (among them the first patriarch of Armenia, Gregory the Illuminator). The consecration of Armenian bishops in Cappadocia became a problem after the official division of Armenia in 387 CE. Regardless of the ongoing process of a developing autocephaly, the Armenian Church remained in communion with the Eastern Roman Church in Constantinople throughout the fourth and fifth centuries CE and accepted the first three Oecumenical Councils (Garsoïan 2004a, 84).

Despite the Eastern Greco-Roman influence on the emergence of the early Armenian Church, the Persian and Syriac Christian traditions remained important throughout the early Armenian Church history as well. Their continuing impact is notably evident in the participation of many of the South-Eastern satrapies of *Greater Armenia* in the Council of the East in Ktesiphon/Seleucia in 410 CE. Bishops from the Southern districts of Armenia are listed among the participants of the Council in Ktesiphon and of other Persian councils. Furthermore, Nina Garsoïan has demonstrated that the deportation of Armenians to the province of Xuzastân near the Persian Gulf by the Sasanian ruler Šāpūr II during the fourth century CE fostered ongoing contact processes between these two regions. As tendencies of so-called dyophysitism had heavily influenced parts of the Syriac and Persian Christian traditions during this period, the “heretical missionaries [from these regions] (...) had a major (...) mostly negative significance in the formulation of Armenian doctrine” (Garsoïan 2004a, 85f.). These two examples demonstrate the overall close contact between parts of the Armenian Church and the Persian/Syriac Churches (Garsoïan 2004a, 82), which is of crucial importance for the discussions

13 The details of these accounts will be discussed in Chapter 4 and are thus not repeated in this historical outline. For a concise overview, see Mahé and Mahé (2012, 78f.) and the introduction by Thomson (1976).

of the Christological tendencies in the early Armenian Church that led to the later schism.¹⁴

Before Garsoïan's influential works, the overall orientation of the Armenian Church had traditionally been regarded as miaphysite. Yet, when considering the close contacts between the Armenian and the Syriac/Persian realm, it is implausible that the early Armenian Church was uniformly miaphysite. Another case in point are the Sasanian emperors who supported the dyophysite Christians in their realm and officially ruled the territory of *Greater Armenia* during parts of the fifth century CE. Thus, Persian kings also temporarily exercised a direct influence on the decisions made by the Armenian Church and the election of the *katholikos*.

Nina Garsoïan has convincingly demonstrated that the influence of the Persian/Syriac and the Eastern Greco-Roman world is reflected in the entire early Armenian Church history. She argues that the Syriac tradition was, at least throughout the fourth and significant parts of the fifth centuries CE, even more important for the early Armenian Church than the Greek tradition that was mainly fostered by the Gregorid house and the Mamikoneans (Garsoïan 2004a, 86). Therefore, authors and church members of the early Armenian Church such as Koriwn, Maštoc', and the famous patriarch Sahak the Great were most likely influenced by dyophysite tendencies as well.

The dyophysite position was rejected by the Armenian Church shortly after the Council of Ephesos in 433 CE.¹⁵ The fact that *katholikos* Sahak (first half of the fifth century CE) also seems to have converted to the anti-dyophysite party led to conflicts with the Persian Church.¹⁶ In the following decades and centuries, the Armenian Church, as well as many other Churches in the Eastern Mediterranean world, turned towards a more and more rigorous rejection of the formerly tolerated or even supported dyophysitism. The rising miaphysite orientation successively led to conflicts with the Roman and Greek Churches as well, particularly after the Council of Chalcedon in 451 CE (Sarkissian 1965). From the viewpoint of the Armenian Church in the late fifth and sixth centuries CE, the decisions made in Chalcedon diluted the miaphysite tendencies that had been prevailing since the Council of Ephesos in 433 CE and thus needed to be rejected.

Yet, the Christological debates and the emerging miaphysite and anti-Chalcedonian orientation of the Armenian Church did not immediately result in

14 These conflicts can very roughly be summarized under the labels of miaphysitism vs. dyophysitism. An in-depth discussion of the Christological debates and the respective councils in the fifth and partially fourth centuries CE goes far beyond the scope of this historical introduction. For a concise overview, see Karrer, Williams, Hauschild et al. (2019).

15 The acts of Nicaea, as well as those of Ephesos from 433 CE, were both known and accepted in Armenia in the first half of the fifth century CE (Garsoïan and Mahé 1997, 10:42ff.).

16 Regarding the possible reasons for this turnaround, see Garsoïan and Mahé (1997, 10:45ff.).

an official separation between the Armenian Church and other Churches such as the Church of Constantinople. Instead, the separation process that led to the eventual schism lasted for several centuries.

One reason for this long process was the promulgation of the famous *Henotikon* by the Roman emperor Zenon in 482 CE. The *Henotikon* was an attempt to avoid the separation of several Eastern Churches in, among others, Egypt and Syria from the Church in Constantinople by re-evaluating the decisions made in Chalcedon from a more pro-Cyrrillian and thus pro-miaphysite position (Brennecke 1998). The *Henotikon* was, therefore, most likely known and accepted in Armenia as well. Due to the *Henotikon* and the late reception of Chalcedon in Armenia, the Armenian Church stayed in communion with the Church of Constantinople for most parts of the fifth century CE (Garsoïan and Mahé 1997, 10:46ff.), at least until Justin I dismissed the *Henotikon* in 518 CE.

Eventually, two Armenian councils taking place during the sixth century CE led to the consecutive autocephaly of the Armenian Church and initiated the schism between the Armenian Church and the Church in Constantinople.¹⁷ The first council took place in Duin in 505/6 CE. However, it is unlikely that the first Council of Duin sanctioned a separation of the Armenian Church from the Churches that had adopted the Chalcedonian *kanones*. Instead, it officially adopted the *Henotikon* and can thus perhaps be regarded as an attempt of rapprochement. The second Council of Duin in 555 CE, however, is often seen as the council during which the official separation took place.

Yet, even this date is still debated and the question of the circumstances of the schism remains uncertain (Garsoïan and Mahé 1997, 10:47). According to Nina Garsoïan, the eventual rejection of the Chalcedonian dogmata and the schism between the Armenian Church and the Church of Constantinople occurred during the sixth and seventh centuries CE or even later, and were likely unrelated to the second Council in Duin as well (Garsoïan 2004a, 93ff.). Nina Garsoïan believes that the two councils in Duin “se préoccupent en premier lieu [de (...) forte recrudescence de dyophysisme] (...) et non de la question de Chalcedoine” (Garsoïan and Mahé 1997, 10:50). Thus, according to the author, the official separation of the Armenian Church from the Church in Constantinople did not take place before the seventh century CE (Garsoïan and Mahé 1997, 10:53).

The exact date of the schism is of minor importance for the topic of this book. Yet, it is important to note that the Armenian Church in the fifth century CE cannot be regarded as a monolithic institution with fixed dogmata. Instead, the fifth century CE was a time of continuous religious and political conflicts that initiated

17 The schism, however, did not only exist between the Armenian Church and the Church in Constantinople but also between the Armenian Church and other Churches which had accepted the decisions made in Chalcedon, such as the Georgian Church.

important shifts in both the political landscape in Armenia and in the dogmatics and status of the Armenian Church.

These processes are also mirrored in the Armenian text basis of this book and in the strong influence of both the Greco-Roman as well as Syriac/Persian Christian traditions on the ongoing translation processes during the fifth and sixth centuries CE in Armenia which were based on both Syriac and Greek sources (Thomson 1994; 1995). Nina Garsoïan attributes works such as the “History of the Armenians” by Agat’angelos to the Northern part of *Greater Armenia* with its center in Ējmiacin and its close ties to Cappadocia and the Church of Constantinople. On the other hand, accounts such as the famous “Epic Histories” by Pseudo-Faustus of Byzantium reflect the influence of Syriac/Persian Christianity on the Southern provinces of Armenia such as Taron with its center in Aštišat. Nina Garsoïan argues that this proves the “initial north-south cultural dissimilarity” that can be regarded as a parallel to the geographical division of Armenia in the Eastern *Greater Armenia* and the Western *Armenia Minor* (Garsoïan 2004a, 83). In sum, “in the light of this duality, a unilateral Hellenocentric approach to the character and evolution of the Armenian Church no longer seems warranted and these can be understood only through an appreciation of the twofold influences upon it” (Garsoïan 2004a, 86).

2.2.3 Summary

This overview of the political and religious history of Armenia during the fourth to sixth centuries CE highlighted key processes in the formative phase of Armenian literature from the fifth century CE onwards.

The first central development that needs to be addressed was the official Christianization of Armenia starting in the fourth century CE. The Christianization of Armenia had a strong influence on the internal and the external constitution of the (Sub-)Caucasian region. For the context of this book, the invention of the Armenian alphabet in the first half of the fifth century CE is of particular importance. This invention resulted from the Christianization process and the need to translate the Bible and other Christian texts. The creation of the alphabet was followed by vast translation processes and the emergence of original Armenian literature. Some of the central texts and translations that stem from this early period of Armenian literature build the basis of the Armenian text corpus used in this book.

The second relevant characteristic of the period under scrutiny is the position of Armenia between the Eastern Roman and the Parthian and later Sasanian empires. The pivotal position of Armenia between two powerful empires had implications on both political and religious developments. In the political realm, the official division of Armenia towards the end of the fourth century CE is probably the most important and most obvious consequence. Regarding the religious developments, the equal importance of Christian texts and influences deriving from

different neighboring regions and languages, such as Syriac or Greek, is also an outcome of the geographical localization of Armenia.

The third and last aspect is related to the above-mentioned phenomenon, too. The history of the Armenian Church during the fourth to sixth centuries CE is characterized by a constant shift between the growing importance of the Greek Church of Constantinople and other Christian traditions in the East, particularly the Syriac and Persian Churches. The differences between these regions and their corresponding Churches also developed during the period between the fourth to sixth centuries CE. Located at the center of these developments, the Armenian Church continuously strove to adapt to the ever-changing religious landscape. While the details of the process which resulted in the autocephaly of the Armenian Church go beyond the scope of this research, it must be remembered that the texts which constitute the corpus of this book were shaped within this context of religious demarcation and inclusion.

2.3 The Notion of Holy in the Study of Religion

This preliminary examination of the notion of *holy* in the study of religion is articulated along two main lines. The first one (Chapter 2.3) deals with the scholarly discourse within the discipline of the study of religion and seeks to highlight diverging understandings of the notion of *holy* therein. While also addressing developments in the history of the notion in the eighteenth and nineteenth centuries CE, it will essentially concentrate on its contemporary academic use. The overview of the main tendencies accepted in the scholarship presented in the first section will shape the working definition of the comparative notion of *holy* in this book.

The second line (Chapter 2.4) focuses on the historical lexical fields commonly related to notions of *holy* that are of particular relevance for the examination of the ancient Armenian terminology around *surb*. These are Latin *sanctus*, Greek *hágios*, and parts of the Persian terminology. Since the modern Western vocabulary that is usually related to a more abstract field of *holy* has its etymological roots in the ancient terminology, the two sections are related to each other. However, this subchapter does not seek to trace possible links between the ancient and the modern terminology throughout the centuries. Instead, the goal is to reconstruct the different semantic contexts in which ancient and contemporary terminologies related to *holy* appear in order to create a framework against which the corpora will be assessed.

This first section concentrates on the contemporary application and understanding of the notion of *holy* in the study of religion. Nevertheless, since the forthcoming discussions are also pertinent to other academic fields such as po-

litical science, psychology, or sociology, this section will also touch on the use of *holy* therein.¹⁸

The main objective is to identify central tendencies that allow sketching the semantic boundaries of the notion of *holy* in the study of religion. Indeed, there is no single definition of the notion but rather a multiplicity of coexisting definitions. A comprehensive overview of central aspects of the notion of *holy* and the *holy/profane* distinction, in particular, is found in William Paden's entry in *Religion in Past & Present* on "Holy & Profane":

The concept of the holy is irritatingly ambiguous: what is pure and what is impure are both "set apart" (pure and impure). In addition, holy can refer to order as well as to transcendence, to protection of boundaries as well as transgression of boundaries or freedom from boundaries, the incontrovertibility of faith as well as a feeling of numinous terror. As an analytic, secular category in the study of religion, the development of the sacred/profane duality is an unfinished process. (Paden 2007)

Discussions in the study of religion and academia have a continuing influence on the meaning and understanding of notions such as *holy* outside of scholarly discourses. For instance, the *Encyclopaedia Britannica* entry on "Sacred" — included in the corpus of this book — targets a broader audience and draws upon the work of Mircea Eliade and Rudolf Otto, two prominent scholars in the study of religion. Another example is the reception of academic notions and discussions in the media. For instance, the film "Medea" by Pier Paolo Pasolini was influenced by debates and works in the study of religion. As argued by Shapiro (2013), it refers to Eliade's distinction in "The Sacred and the Profane" between religious/primitive and modern/profane man who are represented by Pasolini in the persons of Medea and Jason.¹⁹

2.3.1 Carsten Colpe

Carsten Colpe, one of the leading scholars in Germany who worked on the notion of *holy* from an academic perspective, is of particular relevance within the context of this book (Colpe 1977; 1988b; 1990). Carsten Colpe's research interest lies in a potential cross-religious and cross-cultural field of *holy*. His main approach focuses on the comparison of word fields between different languages that are potentially related to the comparative inter-religious notion of *holy*. According to Colpe's

18 For the modern (partially) non-religious use of *holy* in other societal and particularly academic fields, see Canal et al. (2013).

19 I would like to thank Alexander Schröder, a student at the *Center for Religious Studies* in Bochum during the time I was writing my Ph.D., who wrote an excellent paper on this topic.

approach, the examination of the individual semantic fields and their respective overlappings and differences needs to include the discussion of two layers.

The first one is the so-called horizontal layer, which can also be described as representing the synchronic dimension. Conducting an analysis on the horizontal layer implies the comparison of words or notions to other words or notions from the same language that have a similar meaning or are even used as synonyms. The examination of Colpe's horizontal layer of comparison is also part of the examinations in this book, whose quantitative analysis is based on concepts of distributional semantics. The core idea behind distributional semantics is that the joint use of words is central for their individual and their shared meaning.²⁰ Consequently, the examination of word fields and notions that frequently appear in the proximity of *holy* plays a central role in the analysis and includes the discussion of words that are closely related to *holy* within the same language as well.²¹

The second layer of analysis in Carsten Colpe's model is called the vertical dimension and is characterized by a diachronic approach. The vertical dimension is mainly based on translation processes. It traces links between semantic fields of notions across different languages via translation or other linguistic exchange processes such as loan words. This layer is also considered in this book by looking into translations from Ancient Greek to Ancient Armenian. Furthermore, the empirical starting point of this book — namely that the term *surb* in Armenian texts is often translated as *holy* in contemporary English translations — is another example of the vertical dimension. The comparison of the English word fields related to *holy* with the Armenian terminology around *surb* is thus part of the vertical layer of comparison.

The next sections also deal with the vertical layer by examining the contemporary notion of *holy* in the study of religion and selected historical notions in Greek, Latin, and Persian languages. As a result, readers will gain an understanding of the overall semantic field of *holy* in the study of religion, including some of its diverse subfields, as well as the application and meaning of selected notions related to *holy* in ancient languages.

2.3.2 Holy as an Aesthetic Category (Feelings)

The connection between the notion of *holy* and specific aesthetic categories such as the sublime has been particularly present in literature and art. The following quote from Goethe's "Die Leiden des jungen Werther" illustrates how *holy* can be applied in the context of nature experience and gives a good example of the connection

20 See also Chapter 3.

21 A good example is the question of semantic overlappings/differences between "holy," "pure," and "blessed."

between *holy* and aesthetics in the eighteenth and nineteenth centuries CE (Colpe 1990, 50).

(...) ich dann die Vögel um mich den Wald beleben hörte, (...) und das Moos (...) und das Geniste (...) mir das innere, glühende, heilige Leben der Natur eröffnete; wie faßte ich das alles in mein warmes Herz, fühlte mich in der überfließenden Fülle wie vergöttert und die herrlichen Gestalten der unendlichen Welt bewegten sich allbelebend in meiner Seele. (Johann Wolfgang von Goethe, Die Leiden des jungen Werther, in Colpe 1990, 10f.).

In this example, the terminology clustered around *holy* (German “heilig”) is not used as a marker to separate a religious from a profane space or to describe certain persons as *holy*. Instead, *holy* is used as an adjective to characterize external objects that cause impressions, feelings, and emotions of a particular kind (via an aesthetic input). The close relationship between feelings, aesthetics, and the *holy* is, however, not only observable in German literature or philosophy, where it often oscillates between a religio²² and an explicit religious meaning. It is also visible in the study of religion, where the examples of Rudolf Otto’s “The Idea of the Holy” and Friedrich Schleiermacher’s “Über die Religion” illustrate that this relationship is also assumed to be present in explicit religious situations. The examination of “The Idea of the Holy” will also be part of the macro- and microanalyses in this book. Otto’s category of “the numinous”²³ in the early twentieth century CE demonstrates the aesthetic dimension of *holy* as a religious category that triggers feelings and emotions of awe or the sublime through aesthetic impacts in extraordinary situations.

2.3.3 Holy as a Synonym for Religion

Holiness is the great word in religion; it is even more essential than the notion of God. Real religion may exist without a definite conception of divinity, but there is no real religion without a distinction between holy and profane. (Söderblom 1913, 731ff.)

Since the first quarter of the 20th century many historians of religions have accepted the notion of the sacred and of sacred events, places, people, and acts as being central in religious life if not indeed the essential reality in religious life. For example, phenomenologists of religion such as Gerardus van der Leeuw and W. Brede Kristensen have considered the sacred (holy) as central and have organized

22 For the discussion of the term religio^{id}, see the following section on “Holy and Sacralization Processes.”

23 “(...) das Numinose (...) das Heilige minus seines sittlichen Momentes und (...) minus seines rationalen Momentes” in Otto (2004 [1917], 6).

the material in their systematic works around the (transcendent) object and (human) subject of sacred (cultic) activity, together with a consideration of the forms and symbols of the sacred. Such historians of religions as Friedrich Heiler and Gustav Mensching organized their material according to the nature of the sacred, its forms and structural types. Significant contributions to the analysis and elaboration of the sacred have been made by Roger Caillois, a sociologist, and by Mircea Eliade, an eminent historian of religions. (Streng 2018)

Söderblom's entry in the encyclopedia *Religions and Ethics* and Frederick Streng's entry "Sacred" in the *Encyclopedia Britannica* illustrate how some scholars in the study of religion regard *holy* as an essential part of religion(s) or even as a synonym for religion.

This tendency emerged from comparative studies and discoveries in the study of religion during the nineteenth and early twentieth centuries CE. Back then, scholars drew their attention to the many non-Christian and non-European phenomena that seemed intuitively related to religion, but that lacked familiar religious characteristics such as deities or dogmas. As a result, alternative concepts became necessary in order to enable a meaningful comparison between different religious traditions. The attempt to define a comparative notion of *holy* is one part of this undertaking to establish *tertia comparationis* in the study of religion.

Yet, replacing the concept of religion with the concept of *holy* does not make the matter less problematic. The first issue is that the notion of *holy* is anything but well defined. Secondly, the notion itself is deeply indebted to the history of Western religions such as Christianity. Consequently, most scholars nowadays seem to have abandoned the approach which replaces religion with *holy*. I agree with this tendency, hence this book does not seek a comparative notion of *holy* that replaces the notion of religion. Instead, the examinations in this book attempt to contribute to a comparative notion of *holy* that considers not only its Western background but that also addresses other cultural contexts, historical periods, and religious traditions.²⁴

24 As mentioned in the introduction of this book, the importance of comparative research and thus the need for reliable *tertia comparationis* for the study of religion as a discipline was recently underlined by Oliver Freiberger (2019). A good example of an attempt that takes into account non-Western/non-European religions, although related to the notion of *religion* instead of *holy*, is the volume "Religion in Asien? Studien zur Anwendbarkeit des Religionsbegriffs" (Schalk 2013).

2.3.4 Holy and Sacralization Processes

The idea of *holy* in the context of sacralization processes is closely linked to the concept of so-called religioids.²⁵

In this book, a religioid is understood as a religious term (such as *holy*) used in non-religious contexts.²⁶ An example of a religioid is the use of terms such as “sacred” or “holy” in everyday language or other non-religious contexts such as politics. However, although these religious notions may be used in non-religious contexts, their potential to switch back to a religious meaning is still present. An illustrative example is the use of “holy” in nationalist movements where “holy” can stand for “the unified nation” (*heiliges Vaterland*) devoid of any religious connotations. Another example is “(...) la sainteté du contrat social (...)” in Rousseau’s “Contrat Social,” chapter eight. In these cases, the use of notions such as *holy* as religioids means that they maintain the potential to fall back into a more religious interpretation in the sense of “God’s chosen country” or “a contract ratified by God.”

The use of notions such as *holy* as religioids is closely related to the social phenomenon of sacralization processes (as opposed to secularization processes):

Unter Sakralisierung ist in der Religionsforschung ganz allgemein zu verstehen, dass sich die kollektive Verständigung über das Heilige vom religiösen in den nicht-religiösen Bereich ausweitet oder verlagert. (Schlette and Krech 2017, 437)

In accordance with Carsten Colpe, Magnus Schlette, and Volkhard Krech (Colpe 1990, 16f.; Schlette and Krech 2017, 437), the notion of *holy* (or German *heilig*) in this context is defined as a “stabil gewordene Übersetzungsterminologie,” which is “(...) ein Ambivalenzphänomen im Schnittfeld von Immanenz und Transzendenz. Es oszilliert zwischen der Immanentisierung von Transzendenz und der Transzendierung von Immanenz” (Schlette and Krech 2017, 439). According to Schlette and Krech (based on Burkert), the notion of *holy* is defined as:

(...) was über alle Maßen werthaft ist. Unter allem Werthaften wird als heilig nur das Wertvollste bezeichnet, 'letzte' Werte, also dasjenige, was selbst im Ernstfall seiner Bedrohung oder Infragestellung über alle Maßen wichtig genommen wird, was nicht nur Disposition steht (Burkert 1981: 114). (Schlette and Krech 2017, 441)

25 A concept that originally stems from Georg Simmel. See also Bräunlein (2017, 12ff.); Tyrell (2018).

26 In the case of Simmel’s use of religioid as “religiöse Halbprodukte,” this relationship was meant to be the other way round. Simmel tried to identify pre-religious notions and social situations in particular which had the potential to become or better create religion. Among others, see Tyrell (2018, 353ff.).

In this case, the religioid use of *holy* is thus defined by its function as a *Diskurstopper*. Whether in a religious or non-religious context, what is *holy* cannot be put into question. In a religious context, this absolute status of *holy* is represented and guaranteed by a transcendent power such as god. In a non-religious context, however, the reason for the *Unhinterfragbarkeit* is more difficult to establish and often absent.

Thus, the potential to fall back or, from the viewpoint of Simmel, move to a transcendent and thus religious understanding remains an ultimate option in order to legitimize the *Unhinterfragbarkeit*.

2.3.5 Holy and Profane

The ambivalent character of *holy*²⁷ is repeatedly taken up in the discussions about this term in the study of religion. In this context, the notion of *holy* is often expressed in the form of the term “sacred.” The common distinction between *holy* (that includes terms such as “sacred”) and *profane* in a religious context presupposes a separation between two spheres that might be regarded as “religious” on the one hand and “worldly” on the other hand. One of the most influential works for the study of religion in this context was Emil Durkheim’s “The Elemental Forms of Religious Life,” originally published in the early twentieth century CE (Durkheim 2017). In this book, Durkheim states that “sacrality is not a theological concept indicating divine origin but a category of human social behavior” (Paden 2007). Therefore, the distinction between *holy* and *profane* has an intrinsic religioid potential as well.

The implicit and the explicit dichotomy between *holy* and *profane* is often seen as a central indication of the existence of a social field of religion. A case in point is the question of whether the difference between *holy* and *profane* was made in Greco-Roman antiquity.²⁸ Due to the Latin origin of the vocabulary, at least regarding the etymology, one is tempted to assume that such a distinction existed. However, the dichotomy expressed by the contemporary understanding of *holy* and *profane* is not identical to the distinction made between *sacer* and *profanes* in Latin texts but may be better expressed by *fas* and *nefas* (Colpe 1990, 34). This example underlines the difficulty and the pitfalls of applying modern terminology (often based on ancient words) to describe historical phenomena, even if they appear to be closely related to each other.

In addition to the observation made by Durkheim and others that the distinction between *holy* and *profane* functions as a categorical marker demarcating dif-

27 Either within the notion itself or in the form of opposing notions such as *profane*.

28 Something that, in turn, could be taken as an indication that a separate field of religion existed in Roman antiquity.

ferent (social) spheres, there are other approaches underlining the hybrid nature of *holy* and *profane* or even stressing its unifying nature. As Dietmar Kamper stated:

Das Profane ist vom Heiligen überlagert und das Heilige vom Profanen durchzogen. In solcher Ununterscheidbarkeit liegt die Präsenz und die Macht des Heiligen heute. (Kamper 1997, 5)

Another example of the unifying character of *holy* is Paul Tillich's "Systematische Theologie." It assumes that the notion of *holy* is used as "Wiedervereinigung des Getrennten in allen Dimensionen" therefore overcoming the distinctions mentioned above instead of creating them (Bräunlein 2017, 20).

2.3.6 Holy as Power

The view that *holy* is closely related to power or even thought of as power was first articulated in scholarship in the nineteenth century CE on concepts of mana (and its opposite taboo) in Polynesian cultures.²⁹ The concept of power in this context resembles ideas of energy, force, or might, and represents *holy* as a mighty autonomous entity or energy that has its own presence and impact. The idea of a relation between *holy* and such forms of power also influenced Durkheim and his theory of *holy* ("sacred") and *profane*. According to Durkheim, the power of *holy* lies in its potential unification of societies. The power of *holy* can, among others, reveal itself in extraordinary situations such as feasts. In these extraordinary situations, the perceivable force of *holy* triggers extraordinary emotional and physical reactions in those present. The experience of *holy* as power is a so-called *Grenzerfahrung* and an important concept in the works by Rudolf Otto.

Das Heilige zeigt sich in Grenzerfahrungen von Opferhandlungen, Krieg, Gewalt, Tod, Erotik, sexuellen Ausschweifungen, Ritual, Spiel, Fest. Die ontologische Tatsache des Heiligen als vitale Kraft, die Möglichkeit seiner „aktiven Präsenz“, wird als gegeben angenommen. Diese intellektuellen Impulse, das Heilige als das „wilde Heilige“ zu denken, es aus dem „Geist der Gewalt“ zu entfalten, zeitigten nachhaltig Wirkung in Frankreich, und nicht nur dort. (Bräunlein 2017, 20)

The close relationship between *holy* and power, included in the very physical sense of force, violence, and authority, is also observable in certain Hebrew words and notions in the Old Testament that are commonly attributed to the field of *holy*.³⁰ In particular, the Judeo-Christian idea that *holy* derives directly from god underlines the close relationship between *holy* and power. The god of the Old Testament is often

29 See Codrington and the influence of the concepts of mana and totem on important scholars such as Söderblom, Durkheim, Wundt, and others.

30 For more details, see the following section on the historical word fields related to *holy*.

characterized as powerful, mighty, and not least violent.³¹ Another good example is the powerful agency of the “Holy Spirit” both in the New Testament and in many Christian texts and prayers as represented in the corpora analyzed in this book. Yet, the power of *holy* is not only visible in a religious context but also coincides well with its religioid function as a *Diskurstopper*.

2.3.7 Holy as Perfection

The interpretation of *holy* as an attribute of perfection in a moral, ethical, or aesthetic sense is somewhat contrary to the idea of *holy* as a physical or spiritual power mentioned above. In the first case, the *holy* is defined as a (often wrathful) power that, in some instances, might even cause death and destruction. In the second case, however, it is used as a very abstract attribute to address a type of transcendental perfection. An excellent example of the use of *holy* in the sense of perfection³² is Immanuel Kant’s “Kritik der praktischen Vernunft” that also had a great influence on Rudolf Otto (Otto and Harvey 1923, 5).

Die völlige Angemessenheit des Willens aber zum moralisch Guten ist Heiligkeit, eine Vollkommenheit, deren kein vernünftiges Wesen der Sinnenwelt in keinem Zeitpunkt seines Daseins fähig ist. (KdpV, 122)

(...) dem moralischen Gesetz völlig angemessene Gesinnung... (KdpV, 128)³³

Comparable to the idea of *holy* as perfection are Wilhelm Windelband’s and Max Scheler’s assumptions of *holy* as a personalized category of love and compassion (Wokart 2018).

Even though this interpretation of *holy* as an abstract (ethical or moral) perfection seems to be a rather modern development, the idea of *holy* as perfection is also echoed in some of the historical texts analyzed in this book. Used as an attribute for specific persons along with other attributes such as “worthy,” “chaste,” or “virtuous,” *holy* often serves as a marker of (attempted) perfection.³⁴ Furthermore, the (Judeo-Christian) interpretation that the *holy* entirely derives from god, who itself is assumed to be perfect, also strengthens the idea of the *holy* as (transcendent) perfection.³⁵ As Gunsinger writes:

31 As an example, see Jes 12:1-6 where the god of the Old Testament is described as mighty, powerful, as well as holy.

32 For earlier examples such as René Descartes, see Paden (2007).

33 Both quotes are cited after Colpe (1990, 3:21).

34 Frequently in a moral, ethical, and sexual sense.

35 For further details, see the discussions below.

Holiness is the perfection that distinguishes God from the world as the only true object of worship, adoration, and reverence. (Paden 2007)

2.3.8 Holy and Persons

A widespread use of the notion of *holy* in both historical and contemporary religious contexts is its attribution to persons. In its current English use, these persons are usually called “saints.”

The attribution of terms commonly related to the field of *holy* to persons, such as *sacer* and *sanctus*, has ancient and even pre-Christian origins.³⁶ It can already be observed in the old Roman *Zwölftafelgesetz*, albeit with a different meaning than the standard modern or Christian usage.

patronus si clienti fraudem fecerit, sacer esto. (*Zwölftafelgesetz* 8,21; quoted after Servius ad Aen. 6, 609)

In this example from the *Zwölftafelgesetz*, the term *sacer* characterizes a person (the *patronus*) as outside the human jurisdiction and thus belonging to a deity. This is telling of the ambivalent nature of the semantic field of *holy* previously mentioned (see the sections on “Holy and Profane” or “Holy as Power”). Other instances of the ancient personalized use of terms such as *sacer* or *sacrosanctus* are found in Sextus Pompeius Festus’ “De verborum significatu.” They also allude to the fact that persons called *sacer* were regarded as being outside of society.³⁷ Also worth mentioning in this context is the Roman institution of the *tribunus plebis* who was also called *sacrosanctus*.

Another important pre-Christian predecessor of the later Christian saints were the so-called “men of God” (Latin *vir dei*, Greek *theíos anér*). These represent mortals who are close to the gods, which is perceivable by their extraordinary strength and virtue, among others. The notion of “a man of God” is also used for the apostles in the New Testament (Angenendt 2007, 68ff.; Brown 1986, 44ff.). Yet, despite the ancient etymological roots of the terminology, the Christian attribution of *holy* in the context of persons (“saints”) mainly goes back to late antiquity and was influenced by the biblical terminology. The Christian use of *holy* was not based on the term *sacer* but instead on the Latin term *sanctus* (Greek *hágios*). More details are provided in the subsequent section on the “Historical Notions of Holy.”

From early Christian times until today, the connection between *holy* and persons (“saints”) has become both a marker for exceptional people in a religious and

36 For a contemporary use of *sacer* in the context of persons, see Agamben (2015).

37 “At homo sacer is est, quem populus iudicavit ob maleficium; neque fas est eum immolari, sed, qui occidit, parricidi non damnatur (...)” (Thewrewk and Festus 1889, 466).

non-religious or religioid sense.³⁸ What the attribution of *holy* to a person means, however, remains a much-debated question.³⁹ The ascription of the Armenian and English terminology to persons is common in the selected texts in this book as well.

Regarding the early Christian perspective, Arnold Angenendt offers a compelling overview of the characteristics of holy persons. According to his classification, there are two different types of saints. Firstly, every baptized Christian can be called "holy" (and Christianity is, therefore, the communion of holy people). Secondly, according to Angenendt, a particular type of "saints" emerges from late antiquity onwards. This group includes people who either confessed their belief through extraordinary actions (like martyrs) or through words (Angenendt 2007, 34ff.). For Angenendt, the following aspects are typical for holy people. Firstly, holy people may use prayers as a potential intervention before god. They seldom invoke god on behalf of themselves, but mostly for other people. They will also act as advocates before god on behalf of the ordinary people during revelation (2007, 43ff. and 80ff.). Secondly, they exercise fasting and are deeply committed to charity, social dedication, and (to some extent) poverty as well (2007, 48ff.). Thirdly, they often practice an ascetic lifestyle as eremits (*Weltabgewandtheit*), including celibacy and pacifism (2007, 58ff.). Lastly, they can work miracles and fight against demons and the devil (2007, 85ff.).

2.3.9 Holy as a Distinct Religious State and Category of Research (Phenomenology)

Auf das „Phänomen des Heiligen“ stößt man überhaupt nur, wenn man sich irgendwie phänomenologisch einstellt. Stellt man sich wissenschaftlich anders ein, etwa funktionalistisch, sozialanthropologisch oder logisch-analytisch, dann kommt etwas ganz anderes heraus. Die dergestalt angelegten Wissenschaften, sofern sie sich auch mit Religion befassen, beweisen es, indem sie zum Heiligen nichts zu sagen haben, selbst wenn sie konventioneller Weise mit „heiligen“ Einzelheiten operieren. Wenn man sich phänomenologisch einstellt, dann erscheint das Heilige als eine extrem zusammengesetzte Kategorie. (...) Rudolf Otto hatte (...) den kategorialen und als solchen synthetischen Charakter des Heiligkeitsbegriffes bereits richtig erkannt (...) (Colpe 1990, 78)

This quote by Carsten Colpe exemplifies the significant influence that *Religion-sphänomenologie* (as a school of thought in the study of religion) had and partially

38 "Er verhält sich wie ein Heiliger" as an expression to underline that someone behaves in an exceptional (perfect) way.

39 See Beck, Herbers, and Nehring (2017); Angenendt (2007); Beck and Berndt (2013); Brown (2015); Gantke and Serikov (2015); Gemeinhardt and Heyden (2012).

still has on the study of *holy*.⁴⁰ Particularly Rudolf Otto's work "Das Heilige" was crucial for the reception of notions such as *holy* and *numinous* as distinct religious experiences. Even though Otto is usually not regarded as a phenomenologist, he strongly underlined the importance of experiencing religious phenomena such as *holy* in order to understand them. This openness towards religious phenomena that often goes hand in hand with the deliberate exclusion of other states of mind, such as rationality or historicity, is also typical for phenomenologists.

An important phenomenologist was Mircea Eliade, a scholar who focused less on individual experiences of *holy* (as Rudolf Otto did) and more on tracking manifestations of *holy* in the history of religion (Eliade and Trask 1959; Berner 2010). The central importance of *holy* in the phenomenology of religion is also underlined in the following quote from the corresponding lemma on "Phenomenology" in *Religion in Past & Present*.

The phenomenology of religion is the study of manifestations of religious phenomena (Phenomenon, Phenomenology), or their comparative morphology. The prime task is the recording of elements constituting religion, for example, sacred objects, sacred sites, sacred times, actions, holy scripts or people and groups, and also forms of religiosity. But the phenomenology of religion was usually less interested in finding analogies and parallels than in "grasping the essence" (Wesenserfassung, Lanczkowski) of religious phenomena. (Michaels and Bergunder 2019)

The process of "grasping" religious phenomena such as *holy* by directly experiencing them was also underlined by Gerardus Van der Leeuw.

Van der Leeuw did not, however, seek an eidetic vision (eidetic reduction) of the religious phenomenon, but rather the "switching on" of the religious phenomenon "in one's own life," and the consequent "clarification of what is viewed," which feeds into a process of understanding. Dependence, however selective, on Husserl, means that historical and social contexts are consciously blanked out. (Michaels and Bergunder 2019)

The importance of *Religionsphänomenologie*, together with the importance of directly experiencing religious phenomena such as *holy* and understanding them as central for religious experiences in general, decisively shaped the contemporary meaning of *holy* within and beyond academic discourses.

Due to its assumed exceptional significance, ahistorical character, and overall otherness (see also Otto and Harvey 1923, 25ff.), scholars thought for a long time that religious phenomena such as the *holy* called for particular methods to approach them (Pylajew and Antonow 2015, 113ff.). However, this call for "special approaches"

40 See also Colpe (1988b); Gantke and Serikov (2015); Pylajew and Antonow (2015); Schröder (2012).

or an overall exceptional treatment of *holy* further strengthened its perception as “wholly other” and made it difficult to treat it as one social phenomenon among others. This exaggerated emphasis on *holy* as the “wholly other” (Rudolf Otto) is problematic from an academic perspective as it assumes that *holy* cannot be studied with established academic methods but can only be accessed from an inner-religious perspective. This, in turn, would demand the scholar to *be* religious in order to study religious phenomena such as *holy*. Besides the methodological issues that arise with such assumptions, the entire discussion on the centrality of *holy* in religion in this field of the study seems to partly exceed the actual empirically observable importance of this terminology in religions. Even though these tendencies have recently decreased in significance, the initial quote by Carsten Colpe in this section shows that they are still present in the contemporary study of religion.⁴¹

2.3.10 Holy as Purity

Das Heilige ist aus dem Unreinen entstanden. Das Tabu, die Einflößung einer Berührungsscheu, war beiden gemeinsam (und bleibt es auch), der Bereich des Alltäglichen ist profan und rein. Darauf differenziert sich die Scheu in Ehrfurcht vor dem Heiligen und Abscheu vor dem Dämonischen; nun gilt als unrein das, was der heiligen Gottheit mißfällt, das ist das Profane, und das Heilige ist rein. Das Unreine kann dann noch dem Heiligen oppositionell entgegentreten, und zwischen beiden liegt das Reine und Gewöhnliche oder das Profane als Bereich des Erlaubten (Wilhelm Wundt und viele andere). (Colpe 1990, 62)

Two states can be present simultaneously: things that are pure can be either holy or profane; things that are profane can be either pure or impure (...) But what is holy must never come into contact with what is impure. These last two categories are mutually exclusive. And moreover, they possess a dynamic: they try to exert their influence and control over the other two categories, the profane and the pure. Unlike the sacred and the impure, the profane and the pure are static. They cannot share their state; they are not infectious. That makes them secondary categories; they derive their identity from their opposites. Purity is the absence of impurity, profaneness the absence of holiness. Therefore the boundaries between the sacred and the profane as well as between the pure and the impure are porous, so that there are no fixed boundaries. (Paden 2007)

41 See also Gantke and Serikov (2015); Colpe (1988b).

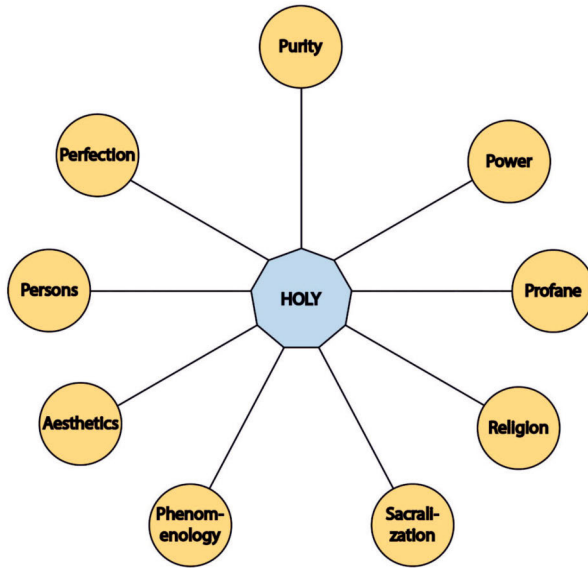
These two quotes by Colpe and Paden underline the necessarily challenging attempts to set the two immensely complex social fields of purity and impurity as well as *holy* and *profane* into a relation.⁴² That both fields were and maybe still are closely related to each other is hardly deniable, not least because both are often understood as dichotomies that operate on the same basis. Yet, whether the *holy* and *profane* distinction derives from the field of purity and impurity or whether they both fall together is another question. In the contemporary academic discourse, however, the dichotomies of *purity/impurity* and *holy/profane* are usually discussed separately, although the overlappings between them are generally acknowledged. Even though the topic of religious purity is frequently discussed in the study of religion (Frevel and Nihan 2013; Malinar 2009), the importance and widespread reception of scholars such as Mircea Eliade and Rudolf Otto and the idea of *holy* as an ethical and moral marker or an autonomous agent that manifests itself in the history of religion have shifted the overall focus towards the notion of *holy* (see also the HSEC Encyclopaedia Britannica below).

2.3.11 Summary

The purpose of this chapter was to sketch the semantic field ascribed to the notion of *holy* in the (contemporary) study of religion. Before turning to the discussion of historical lexical fields related to the notion of *holy*, the results of the current subchapter are summarized in the form of a visualized graph (see figure 2). This outline is meant to serve as a *tertium comparationis* in the comparative part of this book, where both the Armenian notion of *surb* and the contemporary English notion of *holy* will be set into relation with one another. This graph does not display any potential dominance of any of the fields related to the discussion of *holy*. Instead, they are all treated equally, although certain fields such as *holy* in *Religionsphänomenologie* were certainly more influential than others. Yet, it is difficult to measure the importance of each individual subfield precisely. Not least because their relevance often changes between the different (sub-)disciplines within the study of religion. For instance, historical case studies in the study of religion are often concerned with “holy persons,” whereas sociological studies of religious impacts on contemporary societies might be more interested in sacralization processes.

42 For the social and religious importance of purity and impurity, see, among others, Douglas (2001); Frevel and Nihan (2013); Malinar (2009).

Figure 2: Holy and its subfields in the Study of Religion.



2.4 Historical Notions of Holy

The search for an “original” notion or term of *holy* underlying the diverse terminologies in ancient languages such as Greek, Latin, Armenian, Persian, or Hebrew has attracted many scholars in different disciplines, among them the study of religion and Indo-European studies. Harriet Lutzky, a scholar in Indo-European studies, expressed the difficulties of such an undertaking by stating that “no Proto-Indo-European term for the concept ‘sacred’ has been posited, the religious vocabulary varying greatly from one Indo-European language to another” (Lutzky 1993, 283).

Nevertheless, Lutzky observed that different European languages have concepts of two opposing notions to describe a process of separation/binding, which she deems essential for the existence of a field of *holy* (see the previous section on *holy* and *profane*). According to Lutzky, the bipolar concept of separation/binding can be found among different cultures in the form of a great variety of concepts that do not need to rely on a single etymological root or word field (Lutzky 1993, 292).

It is evident that Lutzky’s thoughts are heavily influenced by Emil Durkheim’s works and are, to a great extent, based on Durkheim’s concept of the *sacred* and the *profane* as social categories that structure and stabilize human societies. Lutzky,

however, calls this phenomenon according to its function *separation-binding*, which she also sees at play in Rudolf Otto's famous book "The Idea of the Holy" (Otto 2004 [1917]). The concept of *separation-binding* has the character of dichotomy while also standing for the (potential) unifying aspect of every separation.

(...) the 'sacred' signifies not only division or separation but also stands for wholeness, integration, and binding at the same time. (Lutzky 1993, 284)

Another excellent overview of the use and formation of the notion of *holy* in ancient languages is found in the entry on "Heilig/Heiligkeit" by Wokart in the *Historisches Wörterbuch der Philosophie* (Wokart 2018). In this lemma, the author agrees, at least partially, with some of the assumptions made by Harriet Lutzky and other scholars of Indo-European languages.

Heilig, Heiligkeit. Das griechische ἅγιος bezeichnet wie das lateinische <sanctus> (von sancire: umschließen, umgrenzen) einen abgegrenzten Bezirk, womit alles vor diesem Bezirk (fanum) pro-fanus ist. Dieselbe Herkunft hat auch das hebräische, aus dem Kanaanäischen übernommene Wort <qadōš> für <heilig>, dessen Wurzel <scheiden, absondern> bedeutet. Zur Übersetzung in germanische Sprachen lagen zwei Wörter vor: <hailagaz> mit der Grundbedeutung <eigen, zueigen> und <wihaz>, <geweiht>. Verwendete Wulfila noch das Wort <weihs>, so setzte sich doch unter dem Einfluß der angelsächsischen Missionierung <heilig> als Übersetzung von <sanctus> durch. (Wokart 2018)

According to Wokart, the use of *holy* (mainly the Greek term *hágios*) in ancient pre-Christian texts is connected to the spatial field of separation and often used in tight relation with objects or places, such as temples or sanctuaries (Greek *témenos*).

Wokart argues that the growing influence of the Old and New Testament through Christianity caused a shift in the meaning of *holy*. According to Wokart, the notion changed its focus from the object level (meaning its primary attribution to physical objects and places) to that of persons and god. However, the lexical fields around *holy* could still be attributed to objects just as before. Yet, they now derive their "holy" or "sacred" status from their closeness to god and no longer because they are separated from other types of (non-religious) space. The growing Christian and Jewish influence in (late) antiquity thus caused a shift in the overall notion of *holy* from a spatial-religious marker to a more personalized concept that was very closely related to the Judeo-Christian god.

The following three sections are dedicated to the examination of the concrete vocabulary related to the notion of *holy* in different ancient languages. I will focus on languages and the lexical fields related to *holy* therein, which were important for both the history of the modern notion of *holy* and the Armenian language, namely Greek, Latin, Hebrew, and Persian.

Due to the geographical position of Armenia and the historical events outlined in

the previous part of this chapter, the concept of *holy* in these languages likely had a significant impact on the shaping of the semantic field of Armenian *surb*, for instance, through the translation of the Bible from other ancient languages such as Ancient Greek. Outlining the semantic content of these historical lexical fields of *holy* will help to understand what possible influence other languages might have had on the shaping of the semantic field of *surb*.

This section first considers the Greek terminology, proceeds with an overview of the Latin vocabulary, and closes with a brief review of the Persian/Zoroastrian and Hebrew terminology. While I have examined parts of the original texts of the sources in Greek, Latin, and Armenian, the examinations in the sections on Hebrew and Persian are based on secondary literature rather than my own original research.

Furthermore, other important languages such as Syriac are not considered in this book, even though their influence on the Armenian terminology is widely acknowledged (P'awstos and Garsoïan 1989; Garsoïan, Mathews, and Thomson 1982; Garsoïan 2010). The importance of Syriac for the emerging Armenian literature during the fifth century CE is demonstrated in the historical account by Koriwn and by Nina Garsoïan, who continuously underlined the significance of the Southern and South-Eastern Armenian provinces and their neighbors. Therefore, due to the lack of language skills and the complexity of the Southern relations of Armenia to the Syriac world (to which a short *excursus* without knowing the language would not do any justice), the question of the influence of Syriac on the emergence of the Armenian notion of *surb* and its corresponding semantic field will not be covered.

2.4.1 The Greco-Roman Context

According to Walter Burkert and Albrecht Dihle, the Greek language had several words commonly associated with the notion of *holy* (Burkert, Schröder, and Antes 2011; Dihle 1988). This chapter focuses on two of these terms, namely *hierós* and *hágios* (*hágnos*). Other terms such as *bébelos* or *hósios* will also be touched upon in the course of this chapter. Walter Burkert describes the major differences in the use of *hierós* and *hágios* as follows:

Hierós zieht Grenzen; hagiós schafft ein Kraftfeld, fordert Aufblick und Distanz. (Burkert, Schröder, and Antes 2011, 15:406)

2.4.1.1 Hierós

According to Walter Burkert, *hierós* was the most important term used by ancient Greeks to define a religious sphere since Mycenaean times (Burkert, Schröder, and Antes 2011, 15:402). *Hierós* qualified sanctuaries, material objects such as votive offerings to a deity, as well as abstract categories such as “holy days” or a “holy sick-

ness.”⁴³ *Hierós* was also applied to places such as mountains or entire cities and was used as a noun for sanctuaries (*hierón*) and priests (*hiereús*). According to Burkert, the term *hierós* was never used for the gods themselves but should, instead, be regarded as the “Schatten, den die Gottheit wirft” (Burkert, Schröder, and Antes 2011, 15:403).

Furthermore, while also connected to prohibitions, the term *hiéros* differs from concepts such as a taboo or modern connotations of *holy* in that it is barely connected to emotions such as fear or awe (Otto 2004 [1917], 42ff.). Opposed to *hierós* are words such as *bébelos* (“profane”) and *hósios* (“allowed” in the sense that what does not belong to the deity is available for man) (Burkert, Schröder, and Antes 2011, 15:403f.).

According to some scholars of Indo-European languages and cultures, neither the linguistic nor the conceptual roots of the term “*hierós*” are apparent. Some scholars have suggested that the term might not even be of Indo-European origin (see Gamkrelidze-Ivanov). Others proposed that *hierós* stems from **eis-* which means “vital, fierce, quick,” thereby connecting the conceptual roots of *hierós* to the field of power (Lutzky 1993, 292; Burkert, Schröder, and Antes 2011, 15:403).

2.4.1.2 Hágios/Hágnos

Compared to *hierós*, the term *hágios* is only rarely used in pre-Christian Greek texts. The root of *hágios* is *hag-* from which the verb *hazesthai* is also derived.⁴⁴ *Hazesthai* is often replaced by *aideísthai* (“to fear”) and *sébesthai* (“to worship”). Unlike *hierós*, the term *hágios* mostly describes the inner attitude of the worshipper rather than an object or a place dedicated to a deity (Burkert, Schröder, and Antes 2011, 15:404). Another branch in Indo-European linguistics (represented by Meillet and Lutzky) hypothesizes that the term *hágios* was derived from the Indo-European root **sak-*, therefore meaning “to cut, separate.” This would potentially relate it to the Latin terms *sacer* or *sanctus* (Lutzky 1993, 292).

This hypothesis, however, does not seem to be very likely and appears to be motivated by the attempt of many Indo-Europeanists to establish an abstract cross-cultural field of the (contemporary) notion of *holy* that is linked to functional aspects such as separation and unification. The Greek terms *ágos* and *hágnos* stem from a different root *ag-*. Both are connected to the sphere of an inner taboo and the field of purity (Burkert, Schröder, and Antes 2011, 15:404). Consequently, *hagnízein* and *kat’agízein* are often translated as “to purify.” As opposed to *hierós*, the

43 The “holy sickness” was epilepsy. See, for instance, the text Περὶ ἱερῆς νοῦσου in the *Corpus Hippocraticum*.

44 ἄζομαι, 1. stand in awe of, esp. gods and one’s parents 2. reverently, in holy fear 3. to be angry (LSJ).

word *hágno*s can also be used for gods (for instance, Artemis) as well as for humans, objects, and places. *Hágno*s is closely linked to blood, sexuality, and death. *Hagnà thýmata*, for instance, are bloodless sacrifices. The opposite term of *hágno*s is *miarós* ("stained").

2.4.1.3 Sacer

In ancient Rome the word *sacer* could mean that which would pollute someone or something that came into contact with it, as well as that which was restricted for divine use. (Streng 2018)

Harriet Lutzky, whose assumptions are based on the arguments by Ernout-Meillet and others, has proposed in her article that the Latin term *sacer* is an isolated word in the Italian-Latin context, therefore making it challenging to find an underlying meaning through comparison to terms in other languages (Lutzky 1993, 285). From an etymological and a conceptual perspective, Lutzky suggests that *sacer* might be based on the Indo-European root **sek-* (as in Latin *secare*, "to cut") instead of **sak-* (which might mean "real") (Lutzky 1993, 286). According to Lutzky, this assumption that an aspect of (spatial) separation is expressed in *sacer* perfectly aligns with Durkheim's idea that the *sacred/profane* distinction is a crucial phenomenon in every society. Therefore, the bipolarity of the conceptual phenomenon of *separation-binding* (see above) is, again according to Lutzky, also observable in the case of *sacer*. *Sacer* can mean both the separation of particular objects, concepts, persons, places, etc. (see the *Zwölftafelgesetz*) as well as the ritualistic, political, and religious meaning of "making a contract" (unification) (Lutzky 1993, 288).

2.4.1.4 Sanctus

The term *sanctus* perfectly meets the criteria established by Indo-Europeanists such as Lutzky who posit a general conceptual category of the *sacred* in Indo-European and non-Indo-European societies with an overarching function of *separation-binding*. While the connection of *sacer* to **sak-* ("to cut") is still debatable, the etymological origins of the term *sanctus* from Latin *sancire* (meaning "to separate, cut") are much more evident. Benveniste and other Indo-Europeanists recognized *sanctus/sancire* as a derivative of *sacer*, therefore assuming that they are associated with each other (Lutzky 1993, 290). *Sanctus* was and still is an important Latin term in the context of Christian texts and was used as a major term to translate notions of *holy* from Greek/Hebrew in the Latin translations of the Bible. It was most likely chosen as an alternative to the already frequently used terms such as *sacer* (or *hierós*) in ancient pagan cults. Therefore, due to its importance for the shaping of early Christian thought and literature, the word *sanctus* and its Greek equivalent *hágios* had and still have a significant influence on the contemporary (religious) notion of *holy*.

2.4.2 The Judeo-Christian Context

2.4.2.1 Qadoš and Other Notions

The most pertinent Hebrew notion in the Old Testament that is commonly attributed to a field of *holy* is the word *qadoš*. As the Christian tradition has its roots in the Old Testament, among others, the related terminology is of particular interest. The root of *qadoš*, *qd-* (or *qdš*) meaning “(to) separate,” derives from Akk. *qadāšu* meaning “(to) become clean” (Colpe 1990, 42).

The connotation of separation and the deriving of *qadoš* from *q-d* meaning “to cut, divide” is also strengthened by Lutzky (based on Baudissin). Carsten Colpe states that unlike the Greco-Roman terminology, the word fields related to a notion of *holy* in the Bible and particularly in the Old Testament are attributions of god and only in a second step of holy places or objects.

These objects or places derive their quality of holiness from their relation to god (Colpe 1990, 43). Therefore, there is a significant difference between word fields in the Bible related to a notion of *holy*, particularly in the Old Testament, and the Greek and Latin semantic field of *holy* as described above.

Der Gebrauch der geschaffenen Dinge, die Gott heiligt, (...) ist nicht dasselbe, als wenn bei Griechen und Römern eine Sache dem profanen Gebrauch entzogen wird. Mit letzterem Tatbestand kann man in Israel allenfalls die Bannung vergleichen. (Colpe 1990, 44)

Friedrich Avemarie's entry on the non-biblical Jewish use of notions of *holy* underlines this assumption and shares the Old Testament understanding of *holy/sacred* mentioned above. However, he adds that specific objects and places can be “sanctified” (שִׁדְּדֵשׁ/*qiddeš*) by man in Mishnaic traditions (as well as in the Old Testament).

In the Mishnah, the piel שִׁדְּדֵשׁ/*qiddeš* means “sanctify” in the context of preparing water for purification, ritual bathing, betrothal, and determining the beginning of the month, while the hiphil שִׁדְּדֵשׁ/*hiqdish* means “sanctify” in the context of dedicating votive offerings. Other rabbinic works contain additional usages. The common semantic burden is setting something apart, dedicating it, and defining its purpose. Unlike objects, which become holy by being sanctified, God is holy per se, as his common title “the Holy One, blessed be he” implies. (Paden 2007)

William Paden's list of categories that can become (or are) *holy*⁴⁵ underlines the diversity of the notion of *holy* in the Old Testament and its close relation to other notions such as purity (Paden 2007). The categories mentioned in this list can be sanctified (and profaned) by specific human actions (i.e., rituals) and through the

45 Such as places, objects, etc.

use of objects (such as oil). However, some of them derive their holiness from their close contact with god, other holy objects, or their unique nature (as an example, Paden mentions firstfruits or firstborn children) (Paden 2007). Particularly noteworthy is Paden's focus on the fact that the status of holiness in the Old Testament is "contagious" in that the contact with holy objects can transfer the status of holiness. This idea of contamination is very close to concepts and ideas that are commonly attributed to a field of purity. The close relation between the notion of *holy* and that of purity in particular word fields in the Old Testament is visible in the so-called "Code of Holiness" in Lev 17–26, among others.

In the New Testament and the early Christian tradition, god is relatively seldom called "holy." Yet, things that are close to god (such as persons, objects, and the spirit) are commonly called *holy* (Greek *hágios* or Latin *sanctus*; see the previous sections on the Greek and Latin terminology). They thus derive their holiness from their closeness to god or from actions that are appreciated by god. The following quote from Paden's entry underlines this assumption that all the critical (early) Christian rituals, writings, and persons derive their respective holiness from their relation to god.

The category of the holy is reconstituted by the election of grace, by holiness that manifests itself in mercy. After his resurrection from the dead, the holy one who died on the cross, Jesus Christ, makes all who believe in him holy, "saints." The writings that bear witness to these saints, the testimony of the prophets and apostles, are collectively called Holy script. The community of faith that lives through this sacred witness is called the holy catholic church. The sacred rite of initiation of this community is called holy baptism, and its rite of renewal is called the Holy Communion (Eucharist). Even though the term holy is used in a different way in this context, all its uses have one thing in common: neither the saints, Holy script, the holy church, the holy sacraments, nor anything else is holy by nature, in itself, or per se, but only by God's free grace. (Paden 2007)

Furthermore, from the perspective of early Christians, the early Christian community is holy per se in a world of profaneness due to its elected and separated status.⁴⁶ The self-separation of the Christian community from other non-Christian parts of the society and its self-definition as *holy* underlines once more the dichotomic character of *holy* in the sense of *separation-binding*.

All in all, the notion of *holy* in the New Testament is closely related to the notion of *holy* in the Old Testament since both focus on god. Thus, although the early Christian texts and the Bible in particular frequently adopt the same vocabulary that had been applied in pagan cults in antiquity, the meaning of the terminology of *holy* differs from the former pagan Greco-Roman use described above.

46 Among others, see 2 Cor 1:1; Eph 1:1; Col 1:2; Heb 13:24; Rev 14:12.

2.4.3 Zoroastrianism and the Persian Influence

That the Persian world had a significant impact on the Armenian politics, culture, and religion, particularly in the era before the seventh to eighth centuries CE (starting as early as with the rise of the Medes and Achaemenids), has already been exposed in the previous sections of this chapter. This considerable influence can be witnessed in many different facets of Armenian culture until today, among them the language (Hübschmann 1897). Another aspect are the religious controversies which led to continuing conflicts between Armenia and the Persian empire in late antiquity.⁴⁷ The tensions mainly arose between the Zoroastrian Sasanians who supported a strict religious policy and the newly emerging Christian “nations” in the Caucasian region, such as Armenia. These religious conflicts were of crucial importance for the Armenian elites during the fifth to seventh centuries CE and the overall identity formation processes in this region.⁴⁸ In this respect, Aleksidze states:

(...) for early Armenian writers, who served the needs of the religious elites, the opposition between Zoroastrianism and Christianity and, in a broader sense, the opposition against the Iranian expansion, is the central narrative framework within which the self-perception of the ‘martyr nations’ rests. (Aleksidze 2018, 143)

The outcome of these struggles and the corresponding identity formation processes of Armenia as a Christian nation against both paganism and “wrong” interpretations of Christian dogma have repercussions until today and therefore go far beyond the particularities of a single politically or religiously motivated historical conflict (Aleksidze 2018; Russell 1987). It is not the aim of this book to review the whole history of this often problematic relationship between the different Persian empires and the region of Armenia. However, the ties between both regions and the religious conflicts highlighted in the previous sections compel us to trace Persian terminologies related to *holy* in the emerging Christian religion in Armenia.

2.4.3.1 The Notion of Holy in Zoroastrianism

According to various studies, there was no distinct terminology related to *holy* in the (ancient) Zoroastrian religion and the Iranian/Persian realm. The notion of *holy* was part of the semantic field of purity. This notion was so important that scholars such as Luhrmann prompted comments on Zoroastrianism such as the following one:

47 See the previous section on the historical overview in this chapter.

48 For corresponding discussions in the secondary literature, see the works by Russell (1987) and de Jong (2015).

In some sense Zoroastrianism is no more than a ritualistic commentary upon purity and pollution. (Luhmann 1996, 101)

According to Luhmann and others, the question of purity in Zoroastrianism is linked to the wider Zoroastrian theology. The basis of this theology is the dualism between good and evil and the disorder in the world that the latter causes. Therefore, “the maintenance of purity is the symbolic and performative restitution of order in a disorderly world” (Williams 2015, 346). This assumption echoes with what Mary Douglas wrote in the introduction to her famous book “Purity and Danger” about the essential status of dirt (and impurity).

As we know it, dirt is essentially disorder. (Douglas 2001, 2)

The dichotomy between purity and dirt or purity and disorder can be interpreted as an example of the dichotomies related to *separation-binding* discussed in the previous sections of Chapter 2. Consequently, the purity/impurity distinction⁴⁹ might be regarded as a marker for a semantic field of *holy* that is expressed with a vocabulary related to a semantic field of purity.

To give another example, Alan Williams also argues that the topic of purity and pollution “runs through almost every aspect of Zoroastrianism” (Williams 2015, 347) and is particularly essential for the human stages of “birth, menstruation, and death” (Williams 2015, 346). As previously indicated, the concept of purity is closely related to the distinction between good and evil; thus, “[p]urity rules in the religious tradition are in essence an attempt to re-establish the primordial separation of good from evil” (Williams 2015, 350). Since humankind is currently in a mixed state, it is mainly this in-betweenness at the origin of the impure state that rituals must seek to reverse.

The most articulate Zoroastrian theological texts, written in Pahlavi, define the central challenge of human existence as the need to fight for the good in the battle currently being waged against the forces of evil in the present time of the ‘mixed state’ of existence (Pahl. *gumēzišn*). The opponents at war in the world are personifications of the processes of purification and pollution. The absolute purity of an entity such as the human body and soul, or the element water, is not possible in this *gumēzišn*, and belongs to the time before the great assault (Pahl. *ēbgat[īh]*) upon creation by the Evil Spirit and his forces. Original purity is thus the pristine state of being un-mixed. Since the assault, the creations may be returned to a state of purity and rescued from the conditions of the mixed state, but only temporarily, through ritual practice and other religious action (prayer, devotions). (Williams 2015, 349)

49 That also stands for the distinction between good and evil and, therefore, already goes beyond a pure/impure distinction.

2.4.3.2 aməša spənta

Even though the focus in Zoroastrianism seems to lie on word fields related to the notion of purity instead of *holy*, the hypothesis that the Zoroastrian terminology around purity might also be linked to a wider notion of *holy* is hinted at in contemporary English translations of Zoroastrian texts. These translations often interpret the corresponding Zoroastrian terminology of purity with English terms related to *holy*. For example, in translations by Mary Boyce, Persian words that are usually assumed to be related to purity are translated as “holy.”

The spiritual creation of the seven ‘holy immortals’ (Av. aməša spənta, Pahl. amahraspand) is matched in the physical creation of the seven material prototypes. (Williams 2015, 351)

One of these “holy immortals” who stand for the elements of creation is translated as “Holy Piety/Devotion” (*Spandarmad*). A central aspect in this context is the attempt to keep the “holy” elements pure from the defilement of the evil spirit.

Impurity (Av. irimant-, Pahl. rēmanih) is removed by the religious act of purification, by formulae of prayerful words (nirang) and by righteous intention. All three are brought together in the purification rites of Zoroastrianism, of which there are three principal varieties: 1) Pādyāb, 2) Sādenāhn and nāhn-e sī-šūy, 3) Barašnūm-ī nō šab for the removal of increasingly serious pollutions. (Williams 2015, 355)

The conclusion of the same article further underlines:

This example is just one of hundreds that could be brought to show that purity and pollution are not just signifiers of holiness and its opposite but, as Mary Douglas wrote, echoing Durkheim a century ago, “The dangerous powers imputed to the gods are, in actual fact, powers vested in the social structure for defending itself, as a structure, against the deviant behaviour of its members” (Douglas 1975: 54–55). (Williams 2015, 361)

This passage demonstrates how a focus on a particular research question or term narrows down the meaning of connected notions; in this case, the notion of *holy*. In the article by Williams, the emphasis lies wholly on the notion of purity – the notion of *holy*, on the contrary, is regarded as part of a very general religious vocabulary and not further reflected upon (for instance, in its relation to purity).

But what do the *Aməša Spəntas* stand for? Do they have any relation to a semantic field of *holy*? Or is the translation of *holy* only a misleading convention? The corresponding entry in the *Encyclopaedia Iranica* by Mary Boyce defines the *Aməša Spəntas* as follows:

AMƏŠĀ SPƏNTA, an Avestan term for beneficent divinity, meaning literally “Holy/Bounteous Immortal” (Pahl. Amešāspand, [A]mahraspand). Although the

expression does not occur in the Gāthās, it was probably coined by Zoroaster himself. *Spənta* is a characteristic word of his revelation, meaning “furthering, strengthening, bounteous, holy.” (Boyce 2011)

The six deities that belong to the *Aməša Spəntas* are *Vohu Manah/Vahman* “Good Purpose”; *Aša Vahištā Ardvahišt* “Best Righteousness”; *Xšaθra Vairya/Šahrēvar* “Desirable Dominion”; *Spənta Ārmaiti/Spendārmad* “Holy Devotion”; *Haurvatāt/Hordād* “Wholeness”; and *Amərətāt/Amurdād* “Immortality.” These yazatas are part of the divine essence but remain distinct beings. Later, they do also become personifications of different elements.

Ohrmazd is the especial guardian of the just man, Vahman of cattle, Ardvahišt of fire, Šahrēvar of metals, Spendārmad of earth, Hordād of water, and Amurdād of plants (...) The doctrine of the great Aməša Spəntas thus links spiritual, ethical, and material in a unique manner, and, together with radical dualism, gives its special character to Zoroastrianism. (Boyce 2011)

In sum, the term *spənta* seems to be often translated as “holy” or “immortal,” interestingly neither due to its connection to purity nor due to its relation to any form of *separation-binding*. Instead, it is translated as “holy” because of its connection to power, strength, and fostering character. As seen in the previous sections, these characteristics are also typical of the modern semantic field of *holy* in academia.

Besides the *Aməša Spəntas*, the spirit of Ahura Mazda is also commonly translated as “Holy Spirit” (*Spənta Mainyu*), similar to the widespread attribution of *holy* to the spirit of god in the Gospels and in other Christian texts. Lutzky (based on Benveniste) underlines that the term *spənta* refers to the notion of wholeness (Lutzky 1993, 291), a term that indicates another important subfield of the notion of *holy*, namely “perfection” and “unification.”

2.4.3.3 yaoždata

Another Zoroastrian term closely connected to the field of the *holy* and purity is the Avestan word *yaoždata* that, according to Benveniste and Lutzky (Lutzky 1993, 290), derives from the Indo-European stem **ieuos-* (as in Latin *ius*) meaning “binding force.” The term *yaoždata* denotes both a binding and an implicit separation of pure and impure, which makes it a perfect candidate for a potential *holy/profane* distinction, at least for Harriet Lutzky. The fact that this assessment stems from Lutzky is not surprising. As has been outlined in the previous sections of Chapter 2, Harriet Lutzky was particularly interested in word fields that express a form of *separation-binding*. Consequently, the word field of *yaoždata* with its ties to **ieuos-*

seemed to represent best what she saw as central for the idea of *holy*, namely a form of *separation-binding*.⁵⁰

2.4.3.4 Summary

In conclusion, the Zoroastrian emphasis on physical and theological purity, as described in secondary literature, makes it difficult to locate a distinct notion of *holy* in Zoroastrianism. However, the observation that many of the Zoroastrian terms are (primarily) related to a pure/impure or good/evil distinction, and thus indicate the presence of *separation-binding*, shows that an idea of *holy* might have been present in the Zoroastrian material as well, although not in the form of a distinct notion.

Indications of the underlying presence of a concept of *holy* in the Zoroastrian terminology on purity are:

- the observation that translators often tend to translate the terminology as “holy”⁵¹
- the already-mentioned observation that aspects of *separation-binding* or a notion of power seem to be at work in the Zoroastrian terminology as well

Even though a more profound analysis is beyond the scope of this book, not least due to a lack of language skills, the question arises whether a notion of *holy* existed as part of the semantic field of purity in the Zoroastrian world and in how far this overall focus on purity shaped the Armenian understanding and use of *surb*.

2.4.4 Excursus: The Greek terminology of purity

That the notions of purity and *holy* could already be applied separately in ancient languages by using distinct lexical fields for each notion can be demonstrated with the example of the ancient Greek terminology.

In the Greek context, the two terms most frequently used for (religious) purity were *kat' arós* and *hagnós*. As an opposite term (“stained”) of both words, the word *míarós* was applied (Robertson 2013). According to Robertson, both words were used in inscriptions in temples and sanctuaries from the sixth century BCE onwards. The

50 This idea, in turn, relied heavily on concepts developed by Durkheim and others, see the previous sections in Chapter 2.

51 It is important to note here that this observation does not mean that these translators discovered a form of latent holiness in an essentialist sense in the material. It means that according to their cultural background knowledge and estimations, the terminology in the Persian sources seemed to be equivalent or at least closely related to other word fields with a different semantic background in the English language, in this case, the terminology around “holy.”

term *kat'arós*, however, was “at all times a commoner word by far” (Robertson 2013, 197). It was also used in the broader sense of non-religious purity (“clean”) that could be attributed to clothes, floors, metals, grains, and other objects. The term *hagnós*, on the other hand, was commonly used for deities and “sacred” natural spaces such as sanctuaries and groves.

Therefore, it is hardly surprising to find this term to be prevailing in inscriptions in sanctuaries (Robertson 2013, 199). The term *hagnós* could thus be regarded as what in the article by Manfred Hutter on Anatolian notions of purity is denoted as a second-degree of purity that serves specific religious purposes (Hutter 2013). As the last example in this context, Noel Robertson writes:

In general, the purity denoted by *hagnós* is less urgent and elementary than the *kat'arós* kind. It consists in abstaining for a time from actions or associations most of which are not objectionable in themselves; it is demanded only for the purpose of entering a sanctuary or taking part in ritual. A person will therefore be *hagnós* on occasion but not as a rule. A person must always be *kat'arós*, however, since the purity so denoted is essential to the health and safety of people at large; as its opposite, one who is *miarós* “impure” poses a general threat. [my transliterations] (Robertson 2013, 199)

It will be an essential part of the examination part in this book to look at the Armenian terminology of *surb* and evaluate its standing within this complicated situation of notions of purity and *holy* as part of either distinct or shared word fields in ancient languages in (late) antiquity.

2.4.5 Summary

This short overview of important lexical fields related to *holy* in Greek, Latin, Hebrew, and Persian has revealed the following results.

Firstly, contemporary discussions on the notion of *holy* in the broader field of academia had and still have a strong impact on the interpretation of word fields in ancient languages. For instance, the influence of the concept of *separation-binding* promoted by Lutzky and others is closely related to ideas of the contemporary discussions on the *holy* and *profane* distinction.

Secondly, scholars state that there was a shift in the meaning of the ancient terminology due to the growing influence of Christian traditions. As an example taken from the Greek and Latin contexts, instead of using common words such as *sacer* or *hierós*, Christians tended to use less frequently applied terms such as *hágios* or *sanctus* to translate and signify what was commonly related to a field of *holy*. Furthermore, the Christian ideas were heavily based on the Old Testament and, thus, the Hebrew terminology as well. Opposed to the Greco-Roman traditions,

the concept of *holy* in Christianity was more closely related to a single god and his might and power that also granted the status of *holy* to places, objects, and persons.

Lastly, particularly the Zoroastrian terminology showed a very close relationship between the (potential) existence of a notion of *holy* and its coincidence with word fields related to purity. The question that arises in this context is whether the coincidence of the semantic fields of purity and *holy* is a historical exception or commonly found in ancient societies and languages. This is a topic that mainly arises in studies that deal with purity but only to a lesser degree in (contemporary) literature on *holy*. For instance, one of the central questions posed by Christian Frevel and Christophe Nihan in their introduction to the volume “Purity and the Forming of Religious Traditions in the Ancient Mediterranean World and Ancient Judaism” is:

How is the difference between “purity” and “holiness” (or, respectively, “pure” and “holy”) evaluated (differences, congruence, interdependence)? (Frevel and Nihan 2013, 3:21)

Opposed to the academic discourses related to purity, the close interrelation between *holy* and purity is only seldom mentioned or discussed in the overviews of the contemporary understanding of *holy* in academia (and the study of religion in particular), although overlappings are always briefly stated. Therefore, this book attempts to make this relationship more present, not only when discussing word fields related to purity, but when focusing on notions of *holy* as well.

2.5 Corpus Linguistics and Distributional Semantics

The major research questions in this book are concerned with the semantic problem of the contextual meaning of specific words. In order to analyze the meaning of the notion of *surb* in Ancient Armenian and its relationship with different concepts of *holy*, I apply and adopt ideas, methods, and tools based on distributional semantics and quantitative text analysis (corpus linguistics).

The fields of distributional semantics and (computer-driven) corpus analysis are very closely linked to one another (Gries 2015, 50). The growing importance of computer-driven methods and tools, as well as the growing popularity of *Digital Humanities*,⁵² have made applications of distributional semantics more present and fruitful than ever, not least due to the increased processing power of modern computers and the expanding amount of digitized textual data. Yet, the core

52 A disparate field with blurry boundaries that includes much more than the application of computer-driven methods and tools for the examination of large text corpora (Jannidis, Kohle, and Rehbein 2017; Sahle 2015).

ideas behind distributional semantics, distributional techniques in general, as well as corpus linguistics, are much older than the age of computers. They go back to scholars and philosophers such as Zellig S. Harris (Harris 2015), J. R. Firth (Firth 1957), Ludwig Wittgenstein (Wittgenstein and Schulte 2003), and the founding fathers of distant reading (Underwood 2017).

The core idea of distributional semantics, which does not depend on computers, is that “the different senses of an ambiguous word could be revealed by looking at the different contexts in which the word occurs” (Clark 2015, 496). This idea, in turn, is based on more general ideas about the distributional character of language. In this context, a distribution is defined as follows:

The distribution of an element will be understood as the sum of all its environments. An environment of an element A is an existing array of its co-occurents, i.e. the other elements, each in a particular position, with which A occurs to yield an utterance. A's co-occurents in a particular position are called its selection for that position. (Harris 2015, 146)

Studies based on distributional semantics do not necessarily need to be addressed by using vast amounts of data and text statistical approaches. In theory, it would also be possible to restrict the analysis to just a few sentences. However, this book focuses on a larger amount of data that can only very hardly be processed without the help of computers, at least in certain cases (for example, the Twitter data). The advantage of focusing on vast amounts of data lies primarily in the potential to get a more comprehensive and representative impression of the use of certain notions or words by examining a huge number of use contexts (sentences/texts in which the words or notions appear). The idea that not only the syntactical use of a lexical unit but its meaning as well could be derived from its context was prominently put forward by Hinrich Schütze (Schütze 1998) and Ludwig Wittgenstein (Pulman 2012), among others.

A huge advantage of the focus on the contextual use of lexical units (distribution) is the possibility to represent the distribution as a vector. Vectors, in turn, are readable and processable by computers (Clark 2015). The possibility to represent textual data as a vector is one of the reasons why the idea of distributional semantics is so popular in the field of computer-driven corpus linguistics and other computational fields, such as natural language processing or machine learning (McEnery and Hardie 2012). The representation of word distributions as vectors and matrices makes it possible to further examine them with (statistical) approaches that only work with numerical data. A statistic of a word's or notion's distribution that is based on a large amount of (representative) data finally allows to make predictions about its meaning and to compare it to the numerical representation of the distribution (and thus meaning) of other words and notions.

3. Methodology

As outlined in the previous chapter, the methodological approach in this book is primarily based on distributional semantics. To obtain an overview of the words that regularly appear in the context of *surb* and *holy*, I apply text statistical methods and tools from the field of corpus linguistics, among them:

- word frequency lists
- keyword analysis
- collocation analysis

The results of the corpus analysis are collected in a synoptic table for every text of each corpus. The entries in these synoptic tables are tagged with two annotation schemes. The first scheme is an abstract “domain” annotation system that I have developed in line with the texts in the corpora. The second scheme is represented by the already existing semantic field annotation system called UCREL *Semantic Analysis System* (USAS).¹ The principal idea behind the annotation of the synoptic tables is to provide an abstract overview of the most frequently appearing “domains” and “semantic fields” in the context of *surb* and *holy*. A comparison of the annotated data between the synoptic tables is applied in Chapter 6 in the form of semantic graphs.

In addition to this quantitative approach, selections of single sentences from the corpus data are also analyzed on a smaller scale (microanalysis).

The tools used in the single steps of the corpus examination (macroanalysis) and preparation of the data are *LancsBox*² (text statistics and visualization), *Python*³ (data cleaning and analysis), and *Gephi*⁴ (visualization of the semantic graphs). The overall methodological pipeline is visualized in the corresponding figure 3.

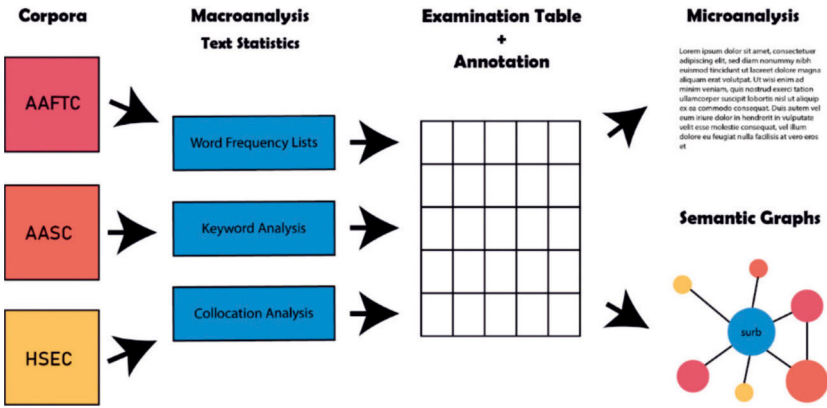
1 UCREL Semantic Analysis System (USAS): <http://ucrel.lancs.ac.uk/usas/> (last accessed: 11/02/2021).

2 Brezina, McEnery, and Wattam (2015).

3 A popular high level programming language.

4 A powerful graph visualization and exploration software. For more information, please see the official *Gephi website*: <https://gephi.org/> (last accessed: 11/02/2021).

Figure 3: Pipeline of methods used in this book.



3.2 Text Statistics

This section introduces the statistical methods applied in the macroanalysis. All of the following methods are well established and widely used in the field of corpus linguistics. Among others, they are further explained in Stefan Gries’ chapter “Quantitative designs and statistical techniques” (Gries 2015), Paul Rayson’s chapter “Computational tools and methods for corpus compilation and analysis” (Rayson 2015) in *The Cambridge Handbook of English Corpus Linguistics* (Biber et al. 2015), and in Vaclav Brezina’s book “Statistics in Corpus Linguistics” (Brezina 2018). All these articles and overviews illustrate how the applications of corpus linguistics and text statistics extend beyond the more limited scope of this book. The latter is mainly interested in semantic questions rather than in, for example, the analysis of morphological or syntactical features. Thus, only a few methods are relevant to its purpose (Trenter 2013).

3.2.1 Tokens, Types, Lemmas (Lexemes)

Before starting with the description of the applied text statistics, I will briefly introduce some fundamental terms and concepts that are necessary for the understanding of the subsequent section on the statistical approaches.

3.2.1.1 Tokens

Tokens best represent what is commonly understood as words in a sentence. Tokens are arrays of characters that are usually separated from each other by blank spaces or punctuation. Tokens are often called “running words.” As an example, ev-

ery single word in the following sentence counts as one token (or “running word”). Thus, the sentence below includes seven tokens.⁵

I eat oranges because you ate oranges.

3.2.1.2 Types

Types refer to distinct forms of a word. Thus, in the case of types, only the number of different words that appear in a text are counted. Consequently, the example sentence above includes six types and seven tokens because the word “oranges” (that counts as one single type) appears twice.

3.2.1.3 Lemmas

Lemmas are even more abstract than types by reducing each token in a text to its base version (usually the lemma found in a dictionary). For instance, the words “to go,” “went,” “goes” can all be reduced to their common lemma “to go.” The lemmatization of a text provides a good overview of the different word fields that appear therein. Consequently, a lemmatized text is particularly useful for a semantic analysis, which is usually more concerned with the word fields that appear in a distribution rather than with their morphological diversity. Thus, the exemplary sentence above includes only four (five) different lemmas:

- Personal pronouns (depending on the lemmatizer, the category of “personal pronouns” might also be split into different pronouns such as “I” and “you”)
- “to eat”
- “orange” (noun)
- “because”

3.2.1.4 Lexemes

Lexemes build yet another layer of abstraction by distinguishing between different meanings of a word (Brezina 2018, 40). Even though the different layers of meaning are central for semantic questions, the fact that there are still no reliable automated (computer-driven) approaches to conduct such a differentiation task across a whole corpus (not even for English texts), lexemes are not considered in this book.

Yet, tokens, types, and lemmas are all reflected. In particular, the lemmas and their frequencies play an essential role. With respect to the meaning of *surb* and *holy*, it is very effective to reduce the inflected words in their neighborhood to a standard lemma to obtain a better overview of the word fields that appear in the context of both notions. The analysis of the lemmatized versions is of particular

⁵ Note that the punctuation marks are sometimes also counted among the tokens. In this case, the example sentence would include eight tokens.

importance in the case of the highly inflected Armenian texts with their complex morphology. Since the morphological variation of the English texts is less problematic, the examination of the English corpora mostly considers types instead of lemmas.

The automated detection of lemmas is complicated to implement because a computerized lemmatizer would need to understand that in the example sentence above, “orange” is used as a noun and not as an adjective. Thus, the annotation of the data with parts-of-speech tagging (POS) is usually a necessary first step for a successful automated lemmatization of a text. There is an abundance of freely available and very useful English lemmatizers but none exist for Ancient Armenian. Unfortunately, a manual POS tagging and lemmatization of the entire Armenian corpus was not possible due to the large amount of text. Therefore, only a subset of the Armenian text corpus (see below) has been lemmatized, partially manually, and partly based on a self-built lemmatizer written in *Python*.

3.2.2 Word Frequency Lists

Word frequency lists are the first statistic that is applied in the macroanalysis of the examination. The idea of word frequency lists is probably the most straightforward statistical method to gain insights into the distribution of lexical units in a text corpus. However, there are still several ways to approach this question.

Firstly, there needs to be a decision regarding the type of word that should be counted (usually either “types” or “lemmas”). Secondly, either the raw word frequencies (called absolute frequencies, short AF) or the normalized frequencies (relative frequencies, short RF) of the distribution of words in a text are calculated. The relative frequency represents the relation between the raw frequency of a word (AF) and the overall number of words in a text. This book applies both absolute frequencies (AF) and relative frequencies (RF). The use of relative frequencies is particularly useful in this book as the texts in the corpora differ profoundly in size. Another common way to express frequency values is to calculate the logarithm of the frequency (usually the natural logarithm). This method makes sense when there are huge gaps between single frequency values. There are more advanced techniques to cope with the frequency values and the distribution of “types” and “lemmas” in word frequency lists, but they are not applied in this book.⁶

Thirdly, another common problem in the context of word frequency lists are expressions that consist of several words such as “of course” or “in spite of” (so-called n-grams) and contractions such as “don’t”. These examples pose problems regarding the calculation of word frequencies that are mostly automatically dealt with by the natural language toolkits used in this book (for instance, *NLTK*, *spaCy*, or the

⁶ Such as relative entropy, see Gries (2015, 53).

built-in parser of *LancsBox*), at least in the case of the English corpora. However, as there are no freely available Ancient Armenian language parsers that consider multi-word expressions or contractions, the word frequency lists of the Armenian texts only count the individual “types” or “lemmas.”

A fourth question in the context of word frequency lists is in how far their results are valuable for the overall examination of the meaning of *surb* and *holy*. In the case of the smaller text corpora such as the AASC,⁷ word frequency lists grant valuable insights into the data because they can be regarded as an extended collocation analysis with flexible L-R values.⁸ However, in the case of texts where the words related to *holy* and *surb* only seldom appear, the question remains in how far the overall word frequencies clarify the meaning and use of both notions. Yet, word frequency lists can provide a comprehensive first impression of the general topic of a text. This overall topic, in turn, sheds light on the text genres and the context in which the terminology appears.

3.2.3 Keyword Analysis

Keyword analyses are the second text statistic applied in this book. The study of keywords is based on the comparison of word frequency lists between a corpus of interest and a reference corpus in order to look for words that are exclusively used in each of the corpora (Brezina 2018, 80ff.; Gries 2015, 55). Words that are not exclusively applied in one of the corpora and thus appear to a similar extent in both corpora are called lockwords. The keyword analysis in this book, and most of the other statistical methods and approaches on the macro level, is done with *LancsBox*.

A corpus of interest (C), sometimes referred to as a ‘focus corpus’ (Kilgariff 2012) or ‘node corpus’ (Scott 1997), is compared with a baseline reference corpus (R) using a statistical measure to identify words that are used either more often or less often in C when compared to R. (Brezina 2018, 80)

In the survey in this book, the keyword analysis is only used in those cases where a reasonable reference corpus exists. For instance, in the case of the Armenian texts, the samples in the *Ancient Armenian Surb Corpus* (AASC) are always the corpus of interest that is compared to the *Armenian Full Text Corpus* (AAFTC) (reference corpus). For more information about the corpora used in this book, please see Chapter 4. The statistical measure used during the comparisons between the node and the reference corpus is “simple maths” (Brezina 2018, 85).

⁷ See Chapter 4.

⁸ For the details of collocation analyses, see the section below.

The results of the keyword analysis are of particular interest in the evaluation of the collocation analysis (“lemmas”) and the word frequency lists of the Armenian texts. The latter text statistics are only performed on the smaller *Ancient Armenian Surb Corpus* (AASC),⁹ which can be problematic in some cases. For instance, if the word “Z” is very prominent in the word frequency list of a particular text in the *Ancient Armenian Surb Corpus* (AASC), this does not necessarily imply that this word is central for the use and meaning of *surb*. Maybe “Z” only appears in the context of *surb* because it frequently occurs throughout the entire text from which the sentences in the AASC derive. Thus, looking at the negative keywords in the keyword analysis between the text in the AASC (node corpus) and the AAFTC (reference corpus) helps to identify words that only appear frequently in the context of *surb* due to their overall high frequency in the text.

In the case of collocation analyses, statistical measures such as log-likelihood (LL) usually take care of these issues by also considering the overall frequency of words. Yet, these statistical measures do not always work in the case of the Armenian texts because the collocation analyses are partially only conducted on the *Ancient Armenian Surb Corpus* (AASC) that does not include the complete texts. Therefore, a comparison of the AASC and the AAFTC via a keyword analysis is a necessary step to avoid potential misinterpretations of the other statistical results.

3.2.4 Collocation Analysis

Collocations are combinations of words that habitually co-occur in texts and corpora. (Brezina 2018, 67)

The collocation analysis is central to this book since it precisely tries to identify which other lexical units (and thus semantic fields) commonly appear in the context of *surb* and *holy*. An association measure is applied to evaluate whether certain words only appear in the context of *surb* or *holy* due to their high frequency in the full text.¹⁰ The association measure used in this book is log-likelihood (LL). For the math behind log-likelihood (LL), also see the explanations in Brezina (2018, 72). Log-Likelihood (LL) is an established association measure in corpus linguistics that is preferred in this book due to its non-exclusive character that also highlights the frequency of collocations (Brezina 2018, 74). The collocation window is usually L2-R2; however, there might be cases where L3-R3 is preferred.¹¹

9 Because only the smaller AASC has been lemmatized.

10 Note the problems in this context regarding the *Ancient Armenian Surb Corpus* (AASC) described in the previous section on the keyword analysis.

11 The L-R span indicates how many words left (L) of and how many words right (R) of the search term are considered in the examination.

This book adopts the “collocation parameters notation” (CPN) proposed by Brezina in his already cited work (Brezina 2018, 75). This notation consists of a) the statistics ID, b) statistic name, c) statistic cut-off value, d) L and R span, e) minimum collocate frequency (C), f) minimum collocation frequency (NC), and g) filters such as “no function words.” An example might look like this: (6-LL, 3, L3-R4, C5-NC5; *function words removed*).

The CPN above reads as follows: The association measure log-likelihood was applied (with the ID 6). The statistic cut-off value was 3.¹² The collocation analysis considered three words left of the search term and four words right of the search term. The collocate needed to appear at least five times in the overall text and five times in the defined collocation window (L3-R4) of the search term in order to be listed. Function words (stop words) such as “the” or “a” were ignored.

3.3 Synoptic Tables and Sentence Analysis

The top results¹³ of the word frequency lists, keyword analysis, and collocation analysis of each text in the corpora are grouped in a separate synoptic table. One row in the synoptic table represents one lemma that appeared among the top results of at least one of the three statistical analyses. It is important to note that the synoptic tables only include “lemmas.” Thus, different words deriving from a “type” analysis that stem from a common lemma such as “went,” “goes,” and “go” are represented together in one row in the synoptic tables (most likely as “go”).

The values in the columns display the lemma’s importance in each of the statistical investigations, a condensed “sum” value, and finally, the annotation of the lemma (see the section below). An ideal-type row could look like in the exemplary table 1 below. Note that the “domain” and “semantic_field” columns will be explained in the subsequent sections. In the columns below, “kw” stands for “keyword analysis,” “word_freq” is short for “word frequency list,” and “coll_freq” represents the collocate frequency value from the collocation analysis.

Table 1: Exemplary row in a synoptic table.

lemma	kw	word_freq	coll_freq	sum	domain	semantic_field
religion	1	1;200	3;43	98	concept	religion_S9

12 The cut-off value is usually not listed in the examinations below since *LancsBox*’s standard cut-off value has been kept.

13 Usually, the top fifteen results of each analysis are considered.

The idea behind the synoptic tables is to condense the results of the three statistical analyses in one single value that can subsequently be used to compare the synoptic tables in the form of semantic graphs. In the synoptic tables, this condensed value is represented in the “sum” column. The calculation of the “sum” values is straightforward and still has a lot of potential for future improvements (see below). In this example, the calculations of the “sum” values summarize the “kw” (keywords), “word_freq” (word frequency), and “coll_freq” (collocation frequency) values.¹⁴

The “kw” value represents the first part of the “sum” value. If the word (“lemma”) in the synoptic table appeared in the positive part of the corresponding keyword analysis (indicated by a “1”), 10 points are added to the total “sum” value. If the word appeared in the negative part of the keyword analysis, 10 points are subtracted from the total sum (indicated by a “-1”). If the word did not appear in the keyword analysis (o) or among the lockwords (2), nothing is added to the total “sum” value.

The values in the “word_freq” and “coll_freq” columns consist of two numbers separated via a semicolon. The first number indicates how often the lemma appeared in either the word frequency lists or the collocation analysis tables. This value is only important for the examinations of “types.” For instance, the words “religion” and “religions” might be among the top entries of a word frequency list. The fact that the lemma “religion” in the synoptic tables thus occurred twice in the statistical investigations is then indicated by a “2” before the semicolon. The second number represents the summarized value of a lemma in the results of a specific statistical investigation. For instance, if the type “religion” appeared 12 times and the type “religions” appeared 25 times among the top entries of a word frequency list, the value after the semicolon would be 37 (25 + 12). Both values are simply added up to the total “sum.”¹⁵

In the example of “religion” in the example row, the overall “sum” value is the result of the following addition:

14 Note that some synoptic tables might include more (or fewer) columns, depending on the corresponding analysis. For instance, not all examinations include a keyword analysis. In the case of the Armenian data, each of the statistical analysis was conducted for “types” and “lemmas.” Thus, the corresponding synoptic tables include columns such as “coll_freq_types” and “coll_freq_lemmas.”

15 This step needs future improvement to better represent the fact that individual “lemmas” might occur several times in the form of different “types.” By simply adding up the values, the multiple appearances of “lemmas” as different “types” is rather neglected in the total “sum” value. Another problem of the current calculation of the “sum” values is the overemphasis of the words in the “word_freq” columns, mainly when dealing with texts in which the notions of *holy* and *surb* seldom appear.

$$10 + 1 + 200 + 3 + 43 = 257$$

Yet, the overall “sum” value in the example row above is 98 and not 257, because the values in the “sum” column are further processed with a simple min-max normalization on the entire column.¹⁶ It is important to note that the values from the normalization are converted to integers, meaning that there are no fractional digits. Thus, two rows with different frequencies in the individual examination columns might have the same overall “sum” value because their different fractional values are cast to the corresponding integer value. It is even possible that some of the rows have a total “sum” value of 0, although they have values higher than 0 in the other columns.

This potential appearance of zero values in the “sum” column, as well as the simplistic way of adding up the frequencies from each column, is of minor importance for the later use of the synoptic tables. Still, future improvements in this calculation are desirable. In the context of this book, neither the synoptic tables nor their “sum” values are meant to provide a precise picture of the exact numbers and frequencies of each analysis. Instead, the “sum” value is supposed to serve as a general and condensed impression of the overall statistical results. Thus, simplification, normalization, and generalization are intended processes here. Furthermore, the idea that semantic fields or relations between words can accurately be summarized and represented in the form of numbers based on selected statistical examinations is misleading. Thus, the slight fuzziness of the overall “sum” values is less important since they should solely serve as approximate estimations for the creation and comparison of the semantic graphs.

The “sum” values are not only applied to create and compare the semantic fields (graphs) in Chapter 6 but they are also part of the selection of sentences for the microanalysis.

Based on the lemmas and their “sum” values in the synoptic tables, one sentence for each “domain” that is present in the synoptic table is selected from the corpus (for the “domain” and the “semantic field” annotation, see the section below). Usually, the lemma and respective sentence with the highest “sum” value is selected. There might, however, be exceptions in some cases. The sentence is then analyzed in a close-reading process by looking at the concrete use of *surb* or *holy* in these examples. The microanalysis of individual sentences complements the macroanalysis and allows to identify potential blind spots in the quantitative part of the examination.

16 $\frac{\text{value} - \text{column}_{\min}}{\text{column}_{\max} - \text{column}_{\min}} * 100$

To illustrate this selection process, I will briefly sketch the selection of sentences for the microanalysis based on the synoptic table of “The Life of Maštoc’” (AASC/AAFTC Koriwn) found in Chapter 5.

Several different “domains,” such as “concept” or “place,” appear in the synoptic table of the Armenian AASC/AAFTC Koriwn. As a first step, one lemma for each “domain” in the “domain” column is selected. Usually, this is the lemma with the highest “sum” value. Correspondingly, for the “concept” domain in the “domain” column, the Armenian lemma *սնունդ* (“name”) with a “sum” value of 81 is selected. As a second step, a sentence from the AASC/AAFTC Koriwn corpus is manually chosen in which the term *սնունդ* (“name”) appears in the context of *surb*. This sentence is finally discussed in the form of a microanalysis and compared to the results of the macroanalysis. This procedure is repeated for all the remaining “domains” that appear in the “domain” column, such as “place,” “verb,” “person,” etc.

3.4 Annotation Scheme

The two annotation schemes described in the subsequent sections are used in the annotation of lemmas in the synoptic tables. The annotations are applied to receive a more abstract overview of frequently appearing domains, discourse fields, and semantic fields in the context of *surb* and *holy*. The distribution of the domains and semantic fields between the different (sub-)corpora will be visualized, analyzed, and compared with one another in the form of semantic graphs in Chapter 6. Together with the “sum” values and their corresponding lemmas, the annotation scheme builds the fundamentals of the overall semantic fields of *holy* and *surb*.

The two annotation schemes applied in this book should be understood as complementary to one another. The first scheme has been developed by me based on the examination of the data. The second scheme derives from the already existing categorization system called *UCREL Semantic Analysis System* (short USAS; see previous sections) by Lancaster University.¹⁷

3.4.1 “Domain” Scheme

The “domain” scheme was developed in view of the textual data and by finding and extracting abstract classes while also considering existing ontologies. The advantage of the self-built “domain” scheme is that it was developed in adequation of the texts examined in this book. This means that it includes both abstract categories

17 UCREL Semantic Analysis System Website: <http://ucrel.lancs.ac.uk/usas/> (last accessed: 11/02/2021).

(such as “concept”) and more specific sub-categories such as “titles” without the necessity to integrate them into a complex ontology. The domains in this scheme are comparable to concepts (or classes) in description logics, which means that they are “groups of individuals that share similar characteristics” (Sakr et al. 2018, 20). The domain scheme used in this book is vaguely inspired by ontologies based on so-called *Web Ontology Languages* (OWL). One example in this context is the ontology built by *dbpedia.org*.

However, as the first scheme is primarily interested in the very abstract domains with as few specifications as possible, there was no need to manually build an entire ontology only for the texts examined in this book. The relational model applied in this book as part of the “domain” scheme only consists of Subject – Predicate – Object relations such as <lexical unit> <isOfType> <abstractClass> or short <lexical unit> “is a” <Class>. This structure and its formalization are closely connected to ideas of *Resource Description Formats* (short RDF).¹⁸ A concrete example is “An altar is a material object” (<Altar> <isOfType> <materialObject>).

The categories in the “domain” scheme have an exclusive character, meaning that they are disjunct. Consequently, a lemma from the synoptic table can only belong to one of the categories in the “domain” scheme. Even though this might be problematic in some cases,¹⁹ the need to choose one category also makes the overall analysis clearer. Furthermore, most of the ambiguous words can usually be ascribed to one of the “domains.” For instance, the term “church” (Եկեղեցի) in the Armenian texts is mostly used in the sense of “the Christian community” and not for a building. Thus, its attribution to the “concept” instead of the “place” or the “object” domain is a transparent and reasonable decision.

In the following part, the structure of the “domain” scheme will be described in more detail.

3.4.1.1 The Structure of the “Domain” Scheme

The “domain” scheme includes eight different classes. The assignment of the lexical units to these classes is restricted to one assignment only, meaning that the instances of the classes have no overlappings. They are disjunct. The “domains” can be organized into two major groups.

The first group includes words that can meaningfully be assigned to abstract classes such as “object” or “concept” (“world domain”). The second group comprises mostly verbs, fill words, adjectives, and other lexical units that cannot reasonably be subsumed under one of the classes from the “world domain.” Instead, they are grouped according to their syntactical function and consequently form the “syntax domain.” The classes in the first category (“world domain”) precede the classes from

18 For a good introduction to RDF schemes, see DuCharme (2013, 24ff.).

19 For instance, does the word “Bible” represent a concept or an object?

the “syntax domain,” meaning that whenever a lexical unit can be assigned to one of the classes from the “world domain,” it is not assigned to a class from the “syntax domain.” For instance, the verb “to purify” is assigned to the class “concept” from the “world domain” and not to the class “verb” from the “syntax domain,” because the verb “to purify” represents an abstract concept (purification) that is more dominant than its syntactical function as a verb.

3.4.1.2 World Domain

Time

This class includes all instances of lexical units that are related to “time” both in a concrete sense (“today”) and a more abstract one (“forever”).

Person

This category includes persons as living (historical/legendary) human beings. The question of whether a person should be regarded as a person or a concept is not always easy to answer in a Christian context. For instance, Christ could be considered as both a person and a concept (“god” or “son of god”). Usually, the term “Christ” is assumed to refer to the “concept” domain, whereas “Jesus” is mostly assigned to the “person” domain.

Place

Another important class is “places.” Similar to the “person” domain, it is sometimes difficult to decide whether a word counts as a “place,” a “concept,” or a material “object.” For instance, the term “church” can refer to a place, a building, and a more general concept in the sense of an institution. The same problem might arise in the context of a monastery.

Concept

A concept as a class of the “world domain” is an abstract and complex mental representation or a system without any direct referent in the form of a concrete material object. The attribute “complex” is important in this context because a chair might also be regarded as a concept; however, in this classification scheme at least, the chair is subsumed under the category “object” instead of “concept,” because it is not complex and has a material referent. Concepts also include actions (such as a kiss). However, very general concepts or actions such as “to take” are not subsumed under this category but are part of the “syntax domain” (“verb”).

Title

The class “title” includes institutional or traditional titles such as “bishop,” “king,” or “katholikos.” The class “title” derives mainly from the analysis of the Armenian data, where titles very frequently appear in the context of *surb*. In some cases, it might become difficult to distinguish between “concepts” and “titles” because “titles” are a subclass of “concepts.” A good example is the attribution of the Armenian title *vardapet* (“teacher”), a term that is often ascribed to *Maštoc’*. In this case, the *vardapet* (“teacher”) does not count as a “concept” but as a “title” because it is a traditional honorary title granted to important persons in the Armenian Church history. It is also important to note that titles were often used as synonyms or metonyms for (groups of) persons (“the bishops” or “the archbishop of Constantinople”). Due to pragmatic and coherency reasons, these appearances are also included in the “title” domain, although they stand for a person.

3.4.1.3 Syntax Domain

Verb

This category includes all the verbs that neither fall into one of the classes of the “world domain” nor into the category “attributes.”

Attribute

This category includes all lexical units that can be used as attributes for persons, objects, concepts, etc. This category has many overlappings with the “concept” domain from the “world domain,” and it is often challenging to determine whether a lexical unit belongs to the “attribute” or “concept” domain. For instance, the term “holiness” might count as an “attribute,” a “title” (“your holiness”), or a “concept.” In this book, “holiness” is considered to belong to the “concept” domain. Usually, only attributes that are very general are subsumed under the “attributes” category, such as “tall” or “small,” since the focus in the “concept” domain was on the complexity of its members.

Miscellaneous

The last class includes all lexical units that do not fit into any of the other categories. The “miscellaneous” domain contains mostly fill words or prepositions.

3.4.2 UCREL Semantic Analysis System (USAS)

The *UCREL Semantic Tagger* (USAS) by Lancaster University is applied in the tagging process of the “semantic field” annotation. The description of the full semantic tag set can be found on the corresponding website.²⁰

The USAS tag set and classification system is based on a classification theme that is in development since 1990. One among several other interesting parts of this classification theme is that it facilitates the attribution of multiple semantic tags to one lemma.²¹ Opposed to the “domain” scheme, a lemma can thus be part of numerous semantic fields. The multiple attribution enables an in-depth and pluralist annotation of the lemmas in the synoptic tables and thus constitutes an excellent complementation of the unambiguous categorization of the words to either a “world” or a “syntax” domain in the “domain” scheme. It also potentially enables a future integration of the data into other projects that use the *UCREL Semantic Tagger* (USAS) annotation scheme.

20 UCRELSemanticAnalysisSystemWebsite: <http://ucrel.lancs.ac.uk/usas/> (last accessed: 11/02/2021).

21 For a more in-depth introduction, see Archer, Wilson, and Rayson (2002).

4. Data

A corpus is a collection of texts (...) that can be analysed using a computer. (Brezina 2018, 15)

This chapter introduces the three text corpora that this book proposes to investigate. I have manually compiled two of these: the *Holy/Sacred English Corpus* (HSEC) and the *Ancient Armenian Surb Corpus* (AASC), which is a subcorpus of the *Ancient Armenian Full Text Corpus* (AAFTC). These will be at the center of my investigations while the third corpus, the *English Web corpus 2015* (*enTenTen15*) that is used with *Sketch Engine*,¹ is mainly employed to assess and complete the results of the HSEC and AASC/AAFTC analyses. This chapter starts with the shorter introduction of the HSEC data that concentrates on the structure of the English corpus and provides two short and optional *excursus* on the machine learning part and the localization of the Twitter data. The presentation of the Armenian data follows in the second part of Chapter 4, including content summaries of the Armenian texts.

4.1 Holy/Sacred English Corpus (HSEC)

The HSEC is a corpus of text samples that seek to be representative of the use of *holy* in the contemporary English-speaking world. Most of the texts stem from the twentieth or twenty-first centuries CE. The Catholic Canon Law appears to be an exception, but it is used in its current revised form from 1983 by pope John Paul II. It also stands for a contemporary understanding of the terminology of *holy* based on a Catholic Christian tradition in its institutionalized form. Despite these grey areas, the HSEC should not be seen as a diachronic sample but rather as a sample of the contemporary use of *holy*.

The sampling frame includes texts from the following contexts:

- The everyday use of *holy*
- The educated public notion of *holy*

1 *Sketch Engine Website*: <https://www.sketchengine.eu/> (last accessed: 11/02/2021).

- The academic notion of *holy*
- The religious use of *holy* in institutionalized religion

The **everyday use** is represented by a collection of tweets (HSEC Twitter Data) that contain both religious and non-religious use cases of *holy*. To some extent, the **educated public notion of holy** is also mirrored in the HSEC Twitter Data. Furthermore, it is part of the entry from the *Encyclopaedia Britannica* on “Sacred” and texts by Eliade and Otto.² The **academic use** of *holy* is represented by a selection of works by Eliade and Otto. The academic use is also present in the entry from the *Encyclopaedia Britannica*, albeit to a lesser extent. The **institutionalized religious use** of *holy* is observable in the Canon Law of the Catholic Church and parts of the HSEC Twitter Data.³

Even though the data in this corpus aims at being representative,⁴ it is biased in several respects. According to the list of potential biases in Brezina’s and McEnery’s & Hardie’s works (Brezina 2018, 16–18; McEnery and Hardie 2012, 8ff.), the overall selection of the subcorpora in this corpus is already biased. The collection of primarily digitized and publicly available texts is a good example of the “practical bias.” Furthermore, the selected texts differ profoundly in their length. To compensate for these biases, some of the results of the HSEC examinations are read against the larger and more representative *English Web corpus 2015 (enTenTen15)* that includes about 15 billion tokens. Although the HSEC should not be regarded as the definite corpus of contemporary uses of *holy*, it does strive to be comprehensive and is a referential corpus. Furthermore, its analysis already yields interesting results.

In the following parts, each of the subcorpora of the HSEC will be introduced in more detail. The introductions include a general overview of the relevant descriptive statistics, such as the number of tokens, and a brief report of how the data was acquired and cleaned.

All the following descriptive statistics of the texts in the HSEC were created with the help of *Lancaster Stats Tools Online* and the *LancsBox* software in particular. If necessary, an initial storing, cleaning, and processing of the texts was conducted with *Python*. Yet, since parts of the data derive from digitized versions of printed books that are publicly available on the Internet,⁵ the data was already noisy⁶ due

2 Eliade’s and Otto’s works had a significant impact on the broader non-academic notion of *holy*. This can be observed in the *Encyclopaedia Britannica* entry, among others.

3 Some of the tweets stem from religious actors such as members of different Christian Churches.

4 Evidently, it is not possible to be representative of a language that continually changes and is spoken by billions of people.

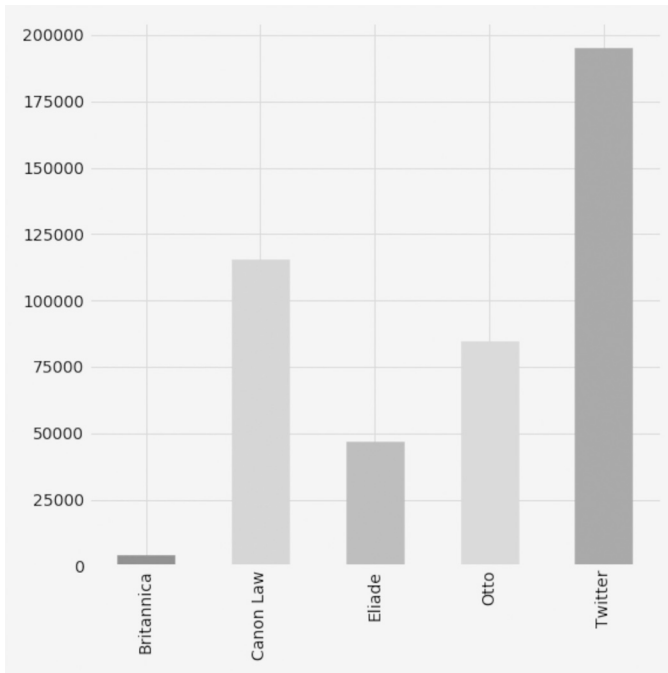
5 Particularly the *Internet Archive*: <https://archive.org/> (last accessed: 11/02/2021).

6 In this context, noisy means that the data contains meaningless information due to distorted and fragmentary texts.

to OCR (optical character recognition) issues. The prior data cleaning tried to polish the texts as thoroughly as possible. Yet, there is still much noise left that is very difficult to remove. For example, all foreign words from languages written in non-Latin alphabets in Rudolf Ott’s “The Idea of the Holy” were wrongly transcribed by the OCR. Consequently, the following statistics should always be treated with caution. Instead of taking them as absolute numbers, they should rather be considered as estimations.

The HSEC includes 446,318 tokens, 29,588 types, and 28,839 lemmas (including stop words;⁷ see the corresponding figure 4) and is constituted by the following texts.

Figure 4: Number of tokens (running words) in absolute numbers of the texts included in the HSEC.



⁷ “Stop words” are a synonym for “function words” and include words such as “and,” “or,” “but,” etc. without any specific meaning. Thus, they are mostly ignored in the statistical analyses.

4.1.1 Encyclopaedia Britannica

4.1.1.1 Data

The raw data of the article on “Sacred” by Frederik J. Streng⁸ was manually copied from the original website of the *Encyclopaedia Britannica* and cleaned with *Python*. The article is publicly and freely available on the website of the *Encyclopaedia Britannica*. It is important to bear in mind that the articles in the *Encyclopaedia Britannica* are updated on a regular basis. Therefore, at a later point, the length and the text of the article on the website might be different from the version used in this book. The text in the HSEC is based on the version of the article from the 27th of September 2018.

4.1.1.2 Rights

The *Encyclopaedia Britannica* and the author hold all the rights of this article. The article is used as part of the corpus data, but it will neither be part of the documentation nor be reprinted in this book. The article is publicly available and accessible online via the links provided in the footnotes and the bibliography.

4.1.1.3 Cleaning

There was only minimal cleaning involved. The text in the corpus has been set to lowercase. Furthermore, all numerical characters, as well as punctuation marks (except of full stops), brackets, etc. have been removed. The chapter titles have not been removed as they only appear once in the text.

4.1.2 Canon Law

4.1.2.1 Data

The raw text of the Canon Law in its digitized English version from 1983 has been taken from the Internet Archive as plain text.⁹ The official website of the Vatican conveniently grants access to the text in a hypertext format. Its repository also includes different related statistics and search engines.¹⁰ Due to coherency reasons, the search engines on the official website are not used in this book. Instead, the Canon Law text, together with the other parts of the HSEC data, are all examined with *LancsBox*.

8 Link to the article by Streng: <https://www.britannica.com/topic/sacred> (last accessed: 11/02/2021).

9 *The Canon Law on the Internet Archive*: https://archive.org/stream/pdfy-jhIKUqi6ojxepam/The%2BCode%2BOF%2BCanon%2BLaw_djvu.txt (last accessed: 11/02/2021).

10 *The Vatican Canon Law website*: http://www.vatican.va/archive/ENG1104/_INDEX.HTM (last accessed: 11/02/2021).

4.1.2.2 Rights

The Vatican holds the rights of this text. However, as this document is publicly and freely available, the text is also integrated into the HSEC. Similar to the rest of the corpus data, the Canon Law is neither reproduced in this book nor is it part of its documentation.¹¹

4.1.2.3 Cleaning

The table of contents, as well as the promulgation by Pope John Paul II from 1983, have been removed. Most of the titles, chapter titles, paragraphs (“\$abc,” “Canon XYZ”), etc. have also been deleted as well as Arabic and Latin numbers and many other non-word characters. The text has also been set to lowercase.

4.1.3 Mircea Eliade

4.1.3.1 Data

The data derives from a digitized version of the English translation of Eliade’s famous work “The Sacred and the Profane” by Willard R. Trask (Eliade and Willard R. Trask 1959) available on the Internet Archive.

4.1.3.2 Rights

As the translation was published in 1959, the rights are held by the publishing house (or its successors). However, the full text is not reprinted in this book but has only been used for the corpus analysis of specific terms.

4.1.3.3 Cleaning

Since the text available on the Internet Archive is based on a digitized version of the printed book, there was a lot of data cleaning involved. Chapters and books titles had to be removed in addition to special characters, numbers, etc. because they appear repeatedly on every other page and would skew the statistics. The text has also been set to lowercase. Despite this thorough cleaning, much noise remains due to the faulty OCR during the digitization of the printed book.

4.1.4 Rudolf Otto

4.1.4.1 Data

The data from Rudolf Otto’s “The Idea of the Holy” also derives from a digitized version of the English translation of the original German text (Otto 2004 [1917]) by John W. Harvey published in London in 1923 (Otto and Harvey 1923). The digitized text is also publicly available on the Internet Archive.

¹¹ To access the full text, please see the websites provided in the footnotes.

4.1.4.2 Rights

Since the translation was published in 1923, and the translation is publicly available on several platforms, there should be no legal issue in using this text. The full text is not reprinted in this book but has only been used for the corpus analysis of specific terms.

4.1.4.3 Cleaning

Since the full book has been digitized, there was a lot of data cleaning involved. For reasons similar to those discussed for the cleaning of the HSEC Mircea Eliade text, there is still much noise left. It is important to note that only the text until appendix XIII is considered.

4.1.5 Twitter Data

4.1.5.1 Data - General Information

The Twitter data consists of two different samples. The larger sample is exploited in the section on the differences between “holy” and “sacred” (see Chapter 5), among others, and includes 23,720 tweets. A smaller subset of this sample constitutes the second sample that only includes approximately 11,726 tweets. The reason for the use of two different samples is twofold.

The primary reason lies in the larger sample’s inclusion of tweets that mention “sacred,” whereas the smaller sample only consists of tweets that include the words “holy” and “holiness.” The second reason is related to a practical issue. The sample data in the smaller tweet collection was gathered between September 2018 and February 2019 (for more details, see below). As for the second sample, it includes the tweets from this period plus more tweets that were collected between February 2019 and July 2019. The smaller sample is primarily still part of this book as the analysis of the notion of *holy* in the English Twitter data (Chapter 5) was conducted in early 2019. At that time, the investigation of the Twitter data was based on a previous classification of the tweets in the smaller sample with the help of a machine learning model (see the excursus below). This machine learning model had been trained with a subset of the smaller Twitter data sample in which the term “sacred” seldom appeared.

Consequently, this model cannot reasonably be applied to classify the larger Twitter data sample. Expanding the analysis and the machine learning model to incorporate the large Twitter sample is a promising avenue for future research.

Despite these issues, the examination of the classified smaller Twitter data sample is a promising first foray into the contemporary application and meaning of “holy” in everyday language. It reveals many interesting insights into the data that will be elaborated upon in Chapters 5 and 6.

4.1.5.2 Data – Collection and Processing

The Twitter data has been collected with the help of the *Python* library *Tweepy*.¹² The smaller sample was collected between the 26th of September 2018 and 14th of February 2019 and consists of 11,726 tweets that include either “holy” or/and “holiness.” The larger sample includes the first sample and additional tweets collected between the 14th of February 2019 and the 23rd of July 2019. These tweets also contain the word “sacred.”

The tweets in the first sample are classified into religious and non-religious tweets. This classification was conducted with the help of a basic machine learning approach (for more details, see below). The script to obtain the tweets from the Twitter streamer with the help of *Tweepy* was set up in a *pythonanywhere.com* environment. It was automatically run twice a day during the period mentioned above. The captured data consisted of so-called “Tweet objects” that were stored in a JSON (JavaScript Object Notation) format. The “Tweet objects” are provided by the Twitter API and include a lot of information for every single captured tweet.¹³

The collected “Tweet objects” from the *Tweepy* streamer were then further processed in a *Python* script that only extracted the following information from the tweets that contain “holy” or “holiness”:

- text/full text,
- timestamp, and
- language

Furthermore, any tweets that had been marked as “retweets” were discarded to avoid duplicates of the same tweet.¹⁴ The new version of the data with only a few selected information was further processed with the help of another *Python* script and finally resulted in a table (more accurately: a *pandas* dataframe in *Python*) that builds the basis for the analysis of the tweets. Each row in the table represents an observation (a tweet) and each column a variable. The row below provides an impression of the stored information of each tweet. Several columns are not taken into consideration in this book as the information that they include is not processed by the methods used in this book. Yet, which methods would eventually be applied was not decided when the data was first collected. However, the remaining

12 *Tweepy website*: <http://www.tweepy.org/> (last accessed: 11/02/2021).

13 For more information about the data stored in a “Tweet objects,” please see the API documentation on the Twitter website under the following link to the *Twitter API documentation*: <https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/tweet-object> (last accessed: 11/02/2021).

14 The missing “retweeted_status” attribute in a “Tweet object” indicates the “no retweet” status of a tweet.

information has been preserved in the corresponding databases¹⁵ for future purposes. For instance, one may envisage a study of the potential correlation between the values of the sentiment analysis¹⁶ and the classification into religious/non-religious tweets. For the analysis with *LancsBox* in this book, however, only the “text” and “ron” columns are further processed.

Table 2: Exemplary row after the processing of the “Tweet objects” with Python.

id	date	text	ron	col5	col6	col7	col8	col9	col10
29	2018-09-26 22:08:40	Holyshit this is amazing	o	holy; shit; amaz- ing	Holy; shit; this; is; amaz- ing	holy; shit; this; be; amaz- ing	NNP; NN; DT; VBZ; JJ	o.2	o.85

The corresponding legend to the column titles is found in the following part:

ld

A unique identifier in the form of an increasing integer value starting with o.

date

The timestamp from the “Tweet object” converted to the DateTime format YYYY-MM-DD HH:MM:SS.SSS.

text

The full cleaned text from the “Tweet object.” With respect to the *Python* script, noise such as abbreviations, smileys, links, and other data that was not pure text has been removed.

ron

“ron” is short for “religious or not.” This column indicates whether a tweet is classified as “religious” (1) or “not religious” (o). The classification as “religious” or “not religious” is based on a machine learning model trained with a manually annotated sample of 3,250 tweets from the smaller sample of the Twitter data. The manual

15 The databases and .csv files are not publicly available since I am not sure about the legal aspects of publishing huge amounts of Twitter data that also contains personal information.

16 In this context, a sentiment analysis is a computational attempt to identify affective states and subjective information in a text. For instance, whether a statement is positive or negative. For more information, see the corresponding *Wikipedia entry*: https://en.wikipedia.org/wiki/Sentiment_analysis (last accessed: 11/02/2021).

categorization of the training data already includes many subjective decisions by the annotator and is thus biased. As a result, the overall annotation is also biased because the machine learning algorithm used in this book classified new data on the basis of the previously annotated one. The excursus below elaborates on the machine learning algorithm used in this book.

The manual annotation of the training set (3,250 tweets) was conducted by one person only (the author of this book). During the manual annotation process, every use of “holy” was classified as “religious” if the direct contextual usage of *holy* had a strong religious potential. Consequently, religioid use cases were often classified as “religious” as well. A good example is the employment of “holy” in “holy land.” This combination might be used in many cases as a synonym for Israel (the modern state of Israel). Yet, in the classification that this book adopts, the “holy land” is mostly regarded as a religious application of *holy* – even though the speaker/author might only refer to the geographical region. This decision has been made because “holy land” implies that the state of Israel, as well as the entire geographic region, have a significant religious meaning and the expression thus still has the potential to fall back into a more religious state.

col5: tokens_old -OLD-

“tokens_old” is a remaining column from an older version of the *Python* script. This column includes tokens created with the *NLTK* tokenizer. Manual removal of stop words, as well as some prior data cleaning, have been applied, too.¹⁷ However, in the current version of the table, the “spacy_text” column is the preferred column for the tokens.

col6: spacy_text -OLD-

This column includes the tokenization of the tweets with the “en” model provided by *spaCy*. This column should be seen as the reference column for the exact tokenization of the tweets in the table.

col7: spacy_lemma -OLD-

This column includes the lemmas of the tweet texts. The lemmas have been extracted with *spaCy* as well. However, the final lemmatization for the statistical anal-

17 The initial tokenization and lemmatization of the tweets were made with *NLTK* and in the case of lemmatization with the *WordNetLemmatizer* from the *NLTK* library. For more information, see the *NLTK website*: <https://www.nltk.org/api/nltk.tokenize.html> (last accessed: 11/03/2021). However, the *WordNetLemmatizer* had some huge disadvantages and was only able to either lemmatize verbs or nouns at the same time. This is the reason why using *spaCy* was preferred in spite of the disadvantages that come with this package, as is well described in this *Medium Article*: <https://medium.com/activewizards-machine-learning-company/comparison-of-top-6-python-nlp-libraries-c4ce160237eb> (last accessed: 11/03/2021).

ysis is done with *LancsBox*. Furthermore, the analysis of the lemmatized version of the English texts plays a minor role in this book.

col8: spacy_pos_tags -OLD-

This column includes the detailed “Parts of Speech” (POS) tagging for the tokens from the “spacy_text” column (also made with *spaCy*). The tags should have the same position in the “spacy_pos_tags” list as their corresponding tokens in the “spacy_text” column. As this book is mainly concerned with the larger “semantic field” and “domain” network of the notions of *holy* and *surb*, this column plays no decisive role either. However, a future investigation of the sentence structures from a macro perspective offers potential for future research.

col9: sentiment_pol_tweet -OLD-

This column includes the first value (polarity) returned by the *TextBlob* sentiment analysis method.¹⁸ The float value in this column can range from -1 (very negative) to 1 (very positive). The text handed to the *TextBlob* sentiment analysis method is the full tweet from the text column.

col10: sentiment_sub_tweet

This column includes the second value (subjectivity) returned by the *TextBlob* sentiment analysis method.¹⁹ The float value in this column can range from 0 (very objective) to 1 (very subjective). The text handed to the *TextBlob* sentiment analysis method is the full tweet from the text column.

4.1.5.3 Excursus: Machine Learning

For the classification of the Twitter data, an instance-based offline machine learning classifier has been trained and applied. The overall testing and training of the machine learning models are documented in a *Jupyter* notebook in my GitHub repository. This short excursus is technical and only serves as a documentation and a general presentation of what has been done to classify the Twitter data in this book. A concise introduction of the basic machine learning technics, their application, and the corresponding terminology goes beyond the scope of this book. Furthermore, it would also require a basic to advanced understanding of programming languages from the reader. In order to get a concise overview of the most important ideas behind machine learning classifiers, such as the one used in this book, I recommend reading Géron (2019).

18 For more information about *TextBlob* sentiment analysis, see the *official TextBlob documentation*: https://textblob.readthedocs.io/en/dev/api_reference.html#textblob.blob.TextBlob.sentiment (last accessed: 11/03/2021).

19 See the previous footnote.

The classifier trained and applied for the annotation of the Twitter data is called “KNeighborsClassifier” and was taken from *Python’s scikit-learn* package.²⁰ The “KNeighborsClassifier” was trained with the standard arguments.²¹ The accuracy of the “KNeighborsClassifier” was evaluated with *scikit-learn’s* cross-validation function²² and then with a separate evaluation of the performance of the trained model on a test set of the overall annotated data.²³

To compare and assess the results and the performance of the “KNeighborsClassifier,” several other machine learning classifiers were trained and tested with the above-mentioned data and methods. All models were trained with their standard arguments. There was no fine-tuning of thresholds or other parameters. The training and testing of different models²⁴ included the following classifiers:

- KNeighborsClassifier
- SGDClassifier
- LogisticRegression
- RandomForestClassifier

The evaluation metrics used to compare the performance of the individual machine learning models were “Precision,” “Recall,” and the “F₁ score.”²⁵ In the case of the Twitter data, the “SGDClassifier” yielded the best scores for all three performance measures²⁶ (see also the corresponding figure 5 for the F₁ scores).

Even though the “SGDClassifier” yielded the best results, the current annotation of the Twitter data is still based on the “KNeighborsClassifier.” The decision to keep the less precise annotation by the “KNeighborsClassifier” was made because the comparison of the different machine learning models in this excursus was only later added and after most parts of the examinations had already been conducted. Thus, changing the classifier would have also implied replacing significant parts

20 For more details about different machine learning models, among them simple instance-based models such as the “KNeighborsClassifier,” see the third chapter in Géron (2019).

21 For more information about the (standard) parameters of *scikit-learn’s* “KNeighborsClassifier,” see the official documentation on the *scikit-learn website*: <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html> (last accessed: 11/03/2021).

22 For more information about cross-validation, see the second and third chapters in Géron (2019) and the entry on the *cross_val_score* function in the *scikit-learn* documentation on the corresponding website (see the previous footnote).

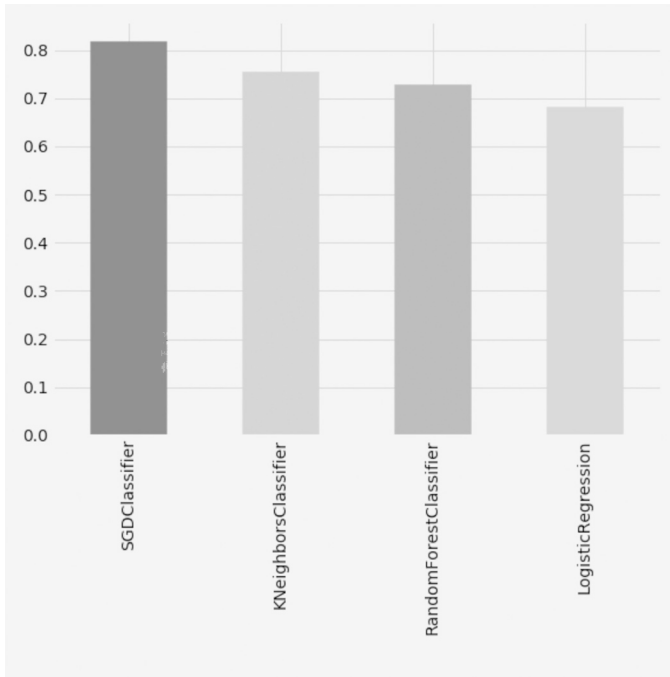
23 The test set includes 25% of the overall data of the 3,250 annotated tweets. Thus, the training set includes 2,438 tweets and the test set 812 tweets.

24 All models were implemented with *scikit-learn*.

25 For more information about the performance measures and the idea behind so-called confusion matrices, see the third chapter in Géron (2019).

26 Performance measures of testing the test set with a trained “SGDClassifier”: Precision: 84,66%, Recall: 79,26%, F₁ score: 81,87%.

Figure 5: F_1 scores of the trained classifiers.



of the entire database and to re-conduct the tests. Yet, the training and testing of the current sample of machine learning classifiers still does not include all the available classifiers.²⁷ This means that there might be a classifier that yields even better results than the “SGDClassifier.” As the applied “KNeighborsClassifier” was placed second in the training and testing in this book²⁸ and yielded good results as well, I have decided to stick with the trained “KNeighborsClassifier” as a machine learning model for the annotation of the Twitter data. For future extension of the analysis (for instance, with a larger Twitter set and more manually annotated data), however, the application of a different classifier seems to be a reasonable step.

4.1.5.4 Excursus: Localization of the Data

To contextualize the use of words related to *holy* in the Twitter data, it is necessary to take into consideration the geographic localization of the tweets used in the

27 For instance, support vector machines (SVM) are still missing.

28 Performance measures of testing the test set with a trained “KNeighborsClassifier.” Precision: 77,97%, Recall: 73,40%, F_1 score: 75,62%.

sample. Knowing about the geographic origin of the data is vital since there might be significant regional differences in the use of terms related to *holy*, for instance, between British English, American English, and non-native English speakers. It is, however, challenging to obtain a representative overview of the localization because Twitter users are not obliged to add geographic data to their tweets. Instead, the specification of this data is optional. To obtain an impression of the localization of the tweets from the larger Twitter sample, the following information from the “Tweet objects” are analyzed.²⁹

- The “location” information in the “user” dictionary of the “Tweet object.” Note that the “location” is user-defined, meaning that the user can fill in whatever he or she wishes.
- The “coordinates” from the “Tweet object.” According to the Twitter API, the coordinates represent “the geographic location of this Tweet as reported by the user or client application.”
- The value in “place/country_code” of the “Tweet object.” This value can indicate a place to which a tweet is associated. The associated place is not necessarily the place from where the tweet was sent.

To recover the geographic distribution of the Twitter data, the most frequently appearing entries in each of the three categories are displayed in the following overview.

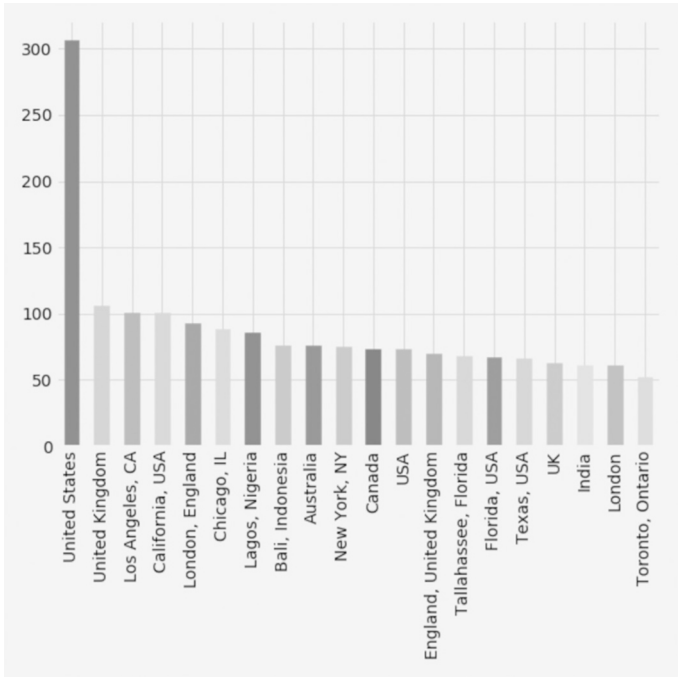
user_location data

25% of the “user/location” entries are *null* values, meaning that the user who sent the corresponding tweet did not add any information about the localization. Furthermore, many of the existing entries do not include the real geographic location of the user (for example, some entries state “hell” or “Hogwarts”). Thus, the corresponding figure 6 can only provide a vague estimation of the origin of the tweets.

A quick look at the entries in figure 6 reveals that the majority of the user locations are in the United States. The second most important group are users from the United Kingdom. Besides other English-speaking countries such as Canada or Australia, there are also significant amounts of tweets that were sent by users from Africa, India, and the Philippines.

29 For detailed information on the “Tweet objects” and the included information, see the Twitter API documentation.

Figure 6: Top 20 most frequent user locations in the Twitter data (large sample).



location_sent data

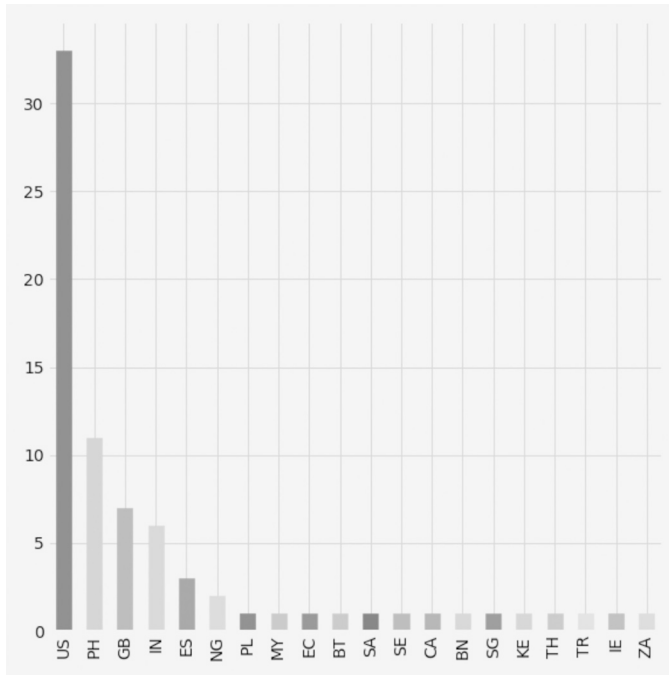
More than 99% of the “location_sent” column (consisting of coordinates from where the tweets were sent) are *null* values, meaning that the user added no value. There are only 77 entries with concrete geographic information. Thus, the corresponding figure 7 with the most frequently appearing countries is not representative of the overall geographic localization of the Twitter data.

In the case of the precise coordinates that were sent by the users, the impression from the previous analysis is underlined. The vast majority of locations are situated in the United States. Noticeably, the second largest user group is not located in Great Britain but in the Philippines.

location_attributed data

More than 96% of the “location_attributed” column (consisting of sites that were associated with the tweets) are *null* values, meaning that the user associated no place.

Figure 7: Top 20 most frequent locations from where the tweets in the Twitter data (large sample) were sent (countries).



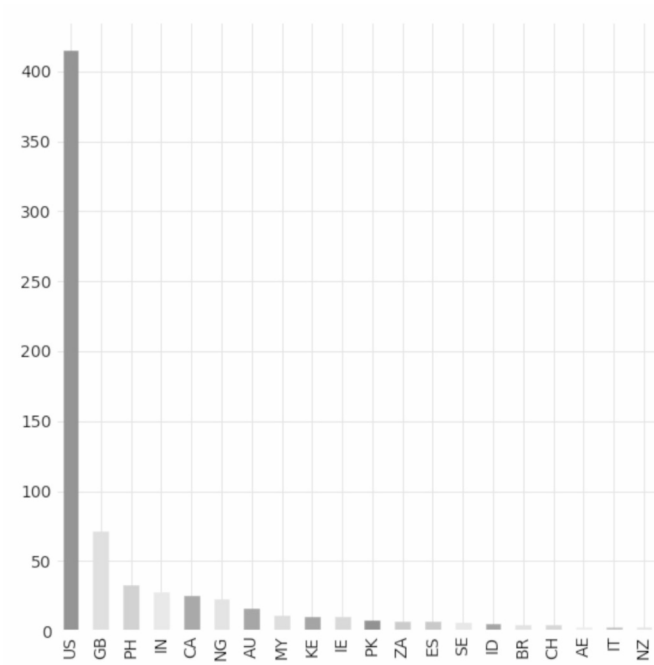
There are only 738 entries with concrete site information. Thus, the corresponding figure 8 with the most frequent associated sites is not representative either.

The previous rankings are also observable in the attributed locations analysis. Unlike in the case of the “coordinates” analysis, the second place goes to Great Britain, whereas the Philippines are in third place.

Summary

A cursory review of the geographic information included in the “Tweet objects” shows that most of the tweets stem from the United States or were sent by users located in the United States. However, as only very few of the tweets include geographic information, this result is not representative and needs to be treated with caution.

Figure 8: Top 20 most frequent associated locations (countries) in the Twitter data (large sample).



4.1.6 English Web corpus 2015 (enTenTen15)

4.1.6.1 Data

The *English Web corpus 2015* (enTenTen15) is an English corpus based on texts collected from the Internet. The corpus belongs to the *TenTen* corpus family, which is a set of web corpora with a target size of 10+ billion words.³⁰

4.1.6.2 Rights

A Campus license of the *Sketch Engine* software was used to access and examine the use of “holy,” “holiness,” and “sacred” in the *English Web corpus 2015* (enTenTen15).

30 For more information, see the TenTen description on the *Sketch Engine website*: <https://www.sketchengine.eu/ententen-english-corpus/> (last accessed: 11/03/2021).

4.1.6.3 Cleaning

There was no further cleaning involved as the received data did not have to be further processed with *LancsBox*. Instead, the examination of the *Sketch Engine* queries relies on the results of the text cleaning and statistical approaches (collocation, word frequencies, etc.) conducted by *Sketch Engine*.

4.2 Ancient Armenian Full Text Corpus (AAFTC) and Ancient Armenian Sub Corpus (AASC)

This subchapter introduces the two Armenian corpora *Ancient Armenian Full Text Corpus* (AAFTC) and *Ancient Armenian Sub Corpus* (AASC). In addition to the sections on data and copyright found in the previous part, this subchapter also includes brief content summaries and information about the provenance of the Armenian sources. The Gospels are an exception since their content will not be summarized.

4.2.1 Introduction of Armenian Texts

4.2.1.1 Agat'angelos – Transmission of the Text

Very little is known about Agat'angelos, the alleged author of the “History of the Armenians” (Agat'angelos and Thomson 1976, lxxv). The attribution of the “History of the Armenians” to a person named Agat'angelos appears to be a later development, possibly to underline the Greek origins of the account, however fictional these may be. The text is usually dated to the fifth century CE, although the surviving text witnesses are much later.

Thomson has challenged the accepted opinion that the “History of the Armenians” by Agat'angelos was an initially Greek text from the fourth century CE that was only later translated into Armenian (Agat'angelos and Thomson 1976, lxxv). Thus, the Armenian tradition of the text (short *Aa*) represents the original version. According to Thomson, all other versions of the A(x) branch, such as the Greek version (*Ag*) and the Arabic version (*Ar*), depend on the Armenian *Urtext* of the “History of the Armenians” (Agat'angelos and Thomson 1976, xxi ff.).

The earliest complete witness of the Armenian text (*Aa*) is a manuscript from the twelfth or thirteenth century CE (Agat'angelos and Thomson 1976, xxi). The Armenian version of the text was first published and printed in Constantinople in 1709. The first and up to now only critical edition of *Aa* was published in Tiflis in 1909. The following examination is based on both the critical edition of the Armenian text from Tiflis (1909) and the translation by Robert Thomson.

The text of the *Aa* edition is split into 900 paragraphs, of which only 395 paragraphs are part of the actual history. The greater part of the text from §225-730

is known as the "Teachings of Gregory," published separately (Agat'angelos and Thomson 1970). The "Teachings of Gregory" will not be discussed in this book.

4.2.1.2 Agat'angelos – Content

The "History of the Armenians" attributed to Agat'angelos deals with the Christianization of Armenia during the first half of the fourth century CE. Even though it was written much later, the account claims to describe these events from the perspective of an eye-witness.

The overall account is largely concerned with the acts and deeds of several persons, among them Gregory the Illuminator and king Trdat, and often has the character of a hagiography. The most important person in the "History of the Armenians" is Gregory the Illuminator, who will become the first bishop of Armenia subsequently to his consecration in Cappadocia presumably in 314 CE. The account by Agat'angelos narrates how Gregory the Illuminator converted the Armenian king of Parthian origin, Trdat, after a series of events.

At the beginning of the account, the Christian Gregory³¹ refuses to worship the pagan deities in Armenia. He is therefore summoned before Trdat and afterward tortured and imprisoned for several years. During Gregory's time in prison, a group of Christian nuns from Rome arrives in Armenia, among them Hřip'simê and Gayianê. According to the legend told in "The History of the Armenians," Hřip'simê had been persecuted by the Roman emperor Diocletian due to her refusal to marry him. To protect her chastity, Hřip'simê and her companions fled to the region of Armenia. The Armenian king Trdat is attracted by Hřip'simê as well and starts besetting her. Finally, the nuns are executed because of Hřip'simê's resistance towards the king's sexual advances.

The two narrative plots around Gregory and the nuns are connected. The Christian god punishes king Trdat for killing the nuns and transforms him into a wild boar. Gregory, who is still imprisoned, is the only person who can eventually cure the king. After releasing the king from his animal form, Gregory baptizes Trdat and travels to Caesarea to be consecrated as the first bishop of Armenia. The rest of the account (§731ff.) is dedicated to multiple stories related to the conversion of the land of Armenia by Gregory and Trdat. Among other narratives, these stories include conversion stories,³² a trip to Rome, and the participation of Gregory's son, Aristakes, in the First Council of Nicaea in 325.³³

31 Who, together with Trdat, had returned to Armenia from his exile in the Roman empire.

32 These conversions mostly consist of the destruction of former pagan sites and baptizing the Armenian peasants.

33 For a more detailed overview of the content and a concise comparison of the different versions of these events in the several different text traditions of "The History of the Armenians," see Agat'angelos and Thomson (1976, xxiv ff.).

4.2.1.3 Koriwn – Transmission of the Text

Only little is known about Koriwn, the author of the *Vita of Maštoc'* (Winkler and Koriun 1994, 245:80ff.). In the “Life of Maštoc’,” the author (Koriwn) describes himself as one of the disciples of Maštoc’, the famous inventor of the Armenian script who lived in the early fifth century CE. According to the “Life of Maštoc’,” Koriwn and another disciple named Łewondwas were sent to Constantinople as a part of an official delegation from Armenia. One of their primary tasks in Constantinople was to translate Christian texts from Greek and Latin with the help of the newly developed Armenian script (Winkler and Koriun 1994, 245:80ff.). After the death of his teacher Maštoc’, Koriwn is said to have written the present account of the life of his teacher Maštoc’.

Even though the dating and the historicity of early Armenian accounts are generally problematic, the account “Life of Maštoc’” is commonly dated to a period between 442–449 CE (Winkler and Koriun 1994, 245:81) and is thus among the oldest remaining Armenian sources. The assumption that the “Life of Maštoc’” can rightfully be called one of the earliest written testimonies of the newly developed Armenian script makes this work particularly interesting for the examination of the early Armenian use of *surb*. This importance mainly results from the fact that such early texts might still include older layers of the use of certain notions that were later overshadowed by other layers or changed due to continuing exchange processes.

There are two major versions of the text called *Koriwn I* and *Koriwn II* (Winkler and Koriun 1994, 245:44ff.). *Koriwn I* is not only longer than *Koriwn II* but is also likely to be the older version. Although it is much shorter, *Koriwn II* includes passages that are not found in *Koriwn I*. This leads to the question of whether the different versions of the “Life of Maštoc’” can reasonably be dated to the fifth century CE or should, instead, be seen as later reworkings of an older account from the fifth century CE that is now lost.³⁴

4.2.1.4 Koriwn – Content

In the center of the account of the “Life of Maštoc’” are Maštoc’ and Sahak, the last Armenian katholikos of the Gregorid dynasty (approximately 387–438 CE).³⁵ Due to its focus on the life and deeds of Maštoc’, the literary genre of the “Life of Maštoc’” is often associated with encomia or panegyrics (Winkler and Koriun 1994, 245:45).

Initially, Maštoc’ was an official at the court of the Armenian king Xosrov and held the position of a secretary or archivist with close connections to the military apparatus (Winkler and Koriun 1994, 245:208ff.). Towards the end of the fourth century CE, he resigned from his service at the royal court and became a monk.

34 Regarding this discussion, see Winkler and Koriun (1994, 245:45f.).

35 Regarding the problematic dating, see Winkler and Koriun (1994, 245:221ff.).

Maštoc' started preaching and doing missionary work in the region of Gołtn. During his evangelizing activities, Maštoc' grew aware of the inconvenience of not possessing a proper Armenian script. One consequence of this was that the church services had to be conducted in Greek and Syriac. According to Maštoc', this led, together with the lack of Christian-Armenian Church literature, to lingering pre-Christian traditions in the various regions of Armenia. Maštoc' eventually convinced the Armenian king Vřamřapuh and katholicos Sahak to establish a proper Armenian script.

Consequently, Maštoc' was told to invent an Armenian alphabet with the aid of a group of young disciples and bishops from Edessa in Samosata (Koriwn I §46ff.).

The invention and successive implementation of the Armenian alphabet resulted in important translation activities from Greek and Syriac to Armenian. These were often conducted by official Armenian delegations which were sent to the neighboring regions and cities of Armenia, such as Constantinople and Edessa (Koriwn I §133ff.).

The newly invented script was also instrumental in evangelizing several rural regions in Armenia and its neighboring countries, such as Georgia and Albania (Koriwn I §60ff. and §117ff.). The production of original Armenian literature emerged alongside the translations from other languages into Armenian (Koriwn I §145; Winkler and Koriun 1994, 245:382). The account of the "Life of Maštoc'" ends with the death of both Sahak and Maštoc' (Koriwn I §168ff.).

4.2.1.5 Epic Histories – Transmission of the Text

The "Epic Histories" had often been ascribed to a certain Faustus of Byzantium, a Greek scholar living in the fourth century CE (P'awstos and Garsoian 1989, 8:6). The account is usually divided into four books that start with book III and end with book VI.³⁶

However, that a historical person named Faustus of Byzantium (or P'awstos Buzandaran in Armenian) ever existed is highly doubted by scholars such as Nina Garsoian. According to Nina Garsoian, the word "Buzandaran" is related to the notion of "epic histories" (Perikhanian 1986; P'awstos and Garsoian 1989, 8:11, 14) and refers to an Iranian and Indo-European tradition of oral epics to which James Russel also compares it extensively (Russell 2012, IX; P'awstos and Garsoian 1989, 8:31ff.).

Whether one agrees with Nina Garsoian's interpretation of the author of the "Epic Histories" or not, its early dating, the closeness to oral traditions and the Persian world, as well as its provenance from the Southern territories of Armenia,

36 Yet, the overall structure of the "Epic Histories" is still debated (P'awstos and Garsoian 1989, 8:16ff.). Most of the information about the structure and text history of the "Epic Histories" are taken from Garsoian (1989, 8:2ff.).

make the “Epic Histories” particularly relevant for the analysis of *surb* in ancient Armenian texts from the fifth century CE.

The translation and overall commentary by Nina Garsoïan is based on the Venice edition of the text in 1832, although she has also considered other editions of the text (Petersburg 1883) (P’awstos and Garsoïan 1989, 8:57). The Armenian text used in this book derives from the digitized version of the “Epic Histories” on *digilib.am*.³⁷

4.2.1.6 Epic Histories – Content

From a historical perspective, the “Epic Histories” continue the “History of the Armenians” by Agat’angelos³⁸ and focus mainly on events taking place in the second half of the fourth century CE. Besides several hagiographical episodes, the “Epic Histories” deal with many political and military conflicts between the Armenian realm (with famous generals such as Mušel Mamikonean and Armenian kings such as Aršak II or Pap) and the Persian realm (Šāpūr II).

Furthermore, the “Epic Histories” also cover significant events of the Armenian Church history. In the center of the narrations and events around the Armenian Church stands *katholikos* Nersēs, the father of the already mentioned Sahak the Great (see the previous section on Koriwn), whose fate was entangled with the political and military conflicts between Armenia and the Persian empire. The involvement of ecclesiastical, religious, hagiographical, and military events in the “Epic Histories” together with their overtly more historical character³⁹ make this account a fascinating source for the examination of the use and meaning of *surb*.

4.2.1.7 Gospels – Transmission of the Text

The need for a translation of the Bible and particularly the Gospels was most likely the primary reason for the invention and establishment of the Armenian script in the early fifth century CE. Accordingly, the Proverbs of the Old Testament were among the earliest texts translated with the help of the newly invented Armenian script (Koriwn I §52). As was demonstrated in Chapter 2 and in the previous section on Koriwn, the invention of the Armenian script needs to be understood in the broader context of the Christianization of Armenia and its neighboring regions. The Gospels stand for an early example of translations into Armenian and constitute an essential source for the analysis of the use of *surb* in Armenian texts from the fifth century CE since they represent the earliest stages of the traceable use of *surb*

37 <https://digilib.aa.am/en/about/site> (last accessed: 11/03/2021).

38 Even though it is neither an official sequel nor directly connected to the account by Agat’angelos.

39 Even though their historical value needs to be treated with care, see Garsoïan (1989, 8:41ff.).

in the newly developed Armenian script. Furthermore, they illustrate how non-Armenian languages and literature have shaped the semantic field of *surb*. Some central issues in this context are:

- To which degree does the term *surb* rely on the terminology of other languages?
- Which terms from other languages such as Greek are translated with *surb*?
- What importance might these translations have had for the later use and meaning of *surb* in other Armenian texts?

Several versions of the Gospels in different languages were used for the translation of the Gospels into Armenian, most importantly the Greek and the Syriac ones. The importance of the Greek and Syriac influences is also perceivable in the “Life of Maš’oc’” by Koriwn. From Koriwn, we learn that two delegations were sent out to translate Christian texts into Armenian. The first delegation traveled to Edessa while the second was sent to Constantinople (Koriwn I §133ff.). In the “Life of Maš’oc’,” the Syriac influence is furthermore visible in the critical role of the Syriac bishop named Daniēl who, according to Koriwn, possessed a prototype of the later Armenian alphabet (Koriwn I §31ff.).

This book will focus on the comparison between the Armenian and Greek versions of the Gospels. However, the importance of the Syriac traditions and their potential influence on the shaping of the semantic field of *surb* will always be kept in mind during the examinations.

For the Armenian version of the Gospels, the so-called Zohrab Bible is used. Hovhannes Zohrabian of the Mechitarist Congregation originally published the Zohrab Bible in 1805 in Venice. The text is publicly available on the sites of the *Titus Project*⁴⁰ in Frankfurt and the website *arak29.am*.⁴¹ This book mainly uses the text from the *arak29.am* website. Regarding the Greek text, the Nestle-Aland edition (NA 28) is used. This version is also publicly available on the Internet (for instance, on *bibelwissenschaft.de*).

Even though the Zohrab Bible is still the standard version of the Armenian Bible and is most widely used in both scholarly and public contexts, it is “a primary desideratum in the field to supersede it with an edition more representative of the version’s broader manuscript tradition” (Cowe 2012, 255). There have been more recent attempts to create critical editions of single books from mostly the Old Testament, but except for some in-depth analysis of the Gospel of Luke, I have not found any recent and established version of the Gospels in Armenian that is also publicly available (Cowe 2012, 264).

40 <http://titus.uni-frankfurt.de/texte/etcs/arm/zohrab/armnt/armnt.htm%7C> (last accessed: 11/03/2021).

41 <https://www.arak29.org/bible/book/> (last accessed: 11/03/2021). See also Cowe (2012, 255).

The earliest codices that include the almost complete version of the Armenian Gospels stem from the ninth to thirteenth centuries CE, among them the Moscow and Ējmiacin Gospels (dating to 887 and 989 CE). Research on the *Urversion* of the Armenian Bible and the Gospels concluded that two different versions existed among the first Armenian translations. The first (*Arm 1*) is dated to the very beginning of the fifth century CE (around 406 CE) and the second (*Arm 2*) to the time after the Council of Ephesos in 431 CE (Cowe 2012, 265f.). Especially the research conducted by Lyonnet in the middle of the twentieth century CE demonstrated the Syriac origins of the first translations of the Gospels (Cowe 2012, 266). However, later revisions of the whole Bible and the Gospels then also considered the Greek text.

In conclusion, the impacts of other languages on the Armenian version of the Gospels, together with their significance and early dating, make them particularly relevant when looking for early strata of *surb* and external influences on its meaning. Of course, the Gospels are only one part of the entire Bible translation. This book focuses on the Gospels for chronological (there are early Armenian manuscripts of the Gospels) and practical reasons. However, for the verification of the findings in the context of the Gospels, it would be exciting to also examine the entire (Armenian) Bible text in a future study and compare it with its Greek (and hopefully also Syriac) counterparts.

4.2.2 Ancient Armenian Full Text Corpus (AAFTC)

4.2.2.1 Data

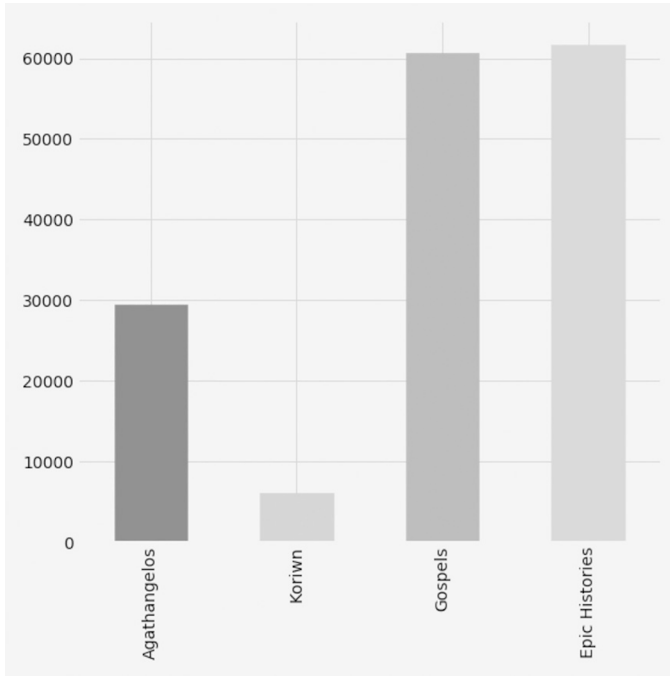
The *Ancient Armenian Surb Corpus* (AASC) is a subcorpus of the *Ancient Armenian Full Text Corpus* (AAFTC) that only includes sentences and passages from the AAFTC in which the notion *surb* occurs. The AAFTC is applied for the evaluation and contextualization of the *Ancient Armenian Surb Corpus* (AASC) and in the collocation analysis (“types”). The main part of the examination of the use of *surb*, however, is based on the *Ancient Armenian Surb Corpus* (AASC).

The AAFTC includes the following ancient Armenian texts. Most of the full texts derive from the corresponding versions on *digilib.am* or, in the case of the Gospels, from *arak29.am*.

- “The History of the Armenians” by Agat’angelos
- “The Life of Maštoc’” by Koriwn
- The “Epic Histories”
- The Armenian text of the Gospels following the so-called Zohrab Bible

The full text of the AAFTC (including stop words) includes 157,852 tokens and 22,030 types (see the corresponding figure 9).

Figure 9: Distribution of tokens (running words) in the Ancient Armenian Full Text Corpus (AAFTC).



“The History of the Armenians” by Agatangelos

The Armenian text follows the edition used by Robert Thomson (Agat’angelos and Thomson 1976). A digitized version of the Armenian text is available on *digilib.am*. The “History of the Armenians” by Agat’angelos includes 29,436 tokens and 8,339 types (including stop words). It covers 18.6% of the total AAFTC text volume.

“The Life of Maštoc” by Koriwn

The Armenian text follows the edition used by Gabriele Winkler for her translation of the “The Life of Maštoc’” (Winkler and Koriun 1994). A digitized version of the Armenian text is available on *digilib.am*. The “Life of Maštoc’” by Koriwn includes 6,141 tokens and 2,728 types (including stop words). It covers 3.9% of the total AAFTC text volume.

“Epic Histories”

The Armenian text follows a digitized version of the Armenian text on www.digilib.am. The translation is based on Nina Garsoïan’s translation “The Epic Histories” (P’awstos and Garsoïan 1989). The “Epic Histories” full text includes 61,631 tokens and 12,132 types (including stop words). It covers 39% of the total AAFTC text volume.

The Gospels

The Armenian text of the Gospels follows the so-called Zohrab Bible that has been digitized by the “Arak29” project on arak29.am. The Armenian text of the Gospels includes 60,644 tokens and 7,882 types (including stop words). It covers 38.4% of the total AAFTC text volume. The complete text of the Gospels is the only Armenian full text that has been wholly lemmatized because the “Arak29” project provides an already lemmatized version of the Gospels on their website. This lemmatization was also used as a basis for the lemmatization of the other texts in the AASC (see below).

4.2.2.2 Rights

All the Armenian text data is based on the digitized text versions publicly provided by digilib.am or arak29.am. The full texts will not be reprinted in this book.

4.2.2.3 Cleaning

The texts have been stripped of all the non-Armenian characters such as numbers, brackets, paragraphs, most of the punctuation marks, etc. As some of the texts (for instance, the “Epic Histories”) also included line breaks, these had to be removed as well. (Sub-)Chapter titles such as Գլուխ Ա. Have also been removed.

4.2.3 Ancient Armenian Surb Corpus (AASC)

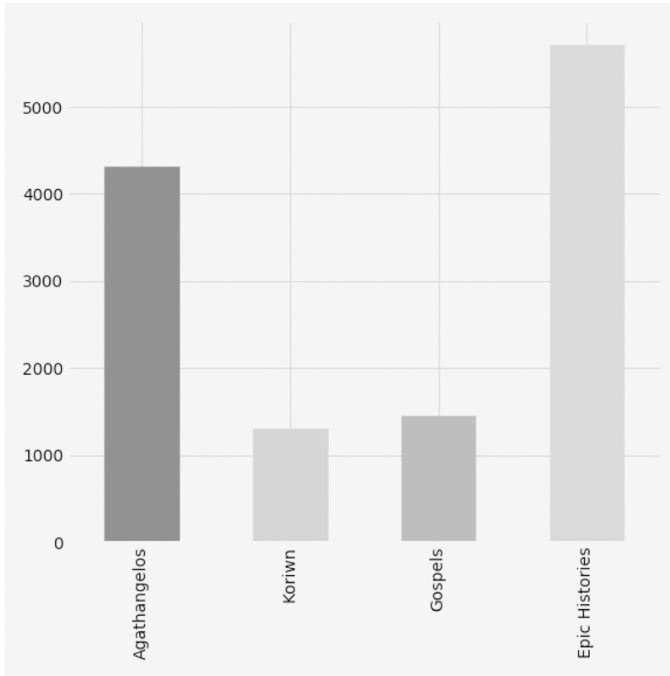
4.2.3.1 Data

The AASC includes 461 sentences from the texts in the AAFTC where words based on the root *surb* (*srb-*) appear. The full text of the 461 sentences (with stop words still included) comprises 12,788 tokens and 4,403 types and approximately 2,055 lemmas (see the corresponding figure 10).⁴²

The sentences stem from the following ancient Armenian texts in the AAFTC. For the details of the text versions and translations, please see the section on the AAFTC.

42 Note that the number of lemmas is too high. This is mainly due to the fact the automated lemmatizer might provide several lemmas for one type in certain cases. For instance, the lemmatizer suggests գթալ and գթել for the type գթացաւ.

Figure 10: Token (running words) distribution in the Ancient Armenian *Surb* Corpus (AASC).



“The History of the Armenians” by Agatangelos

The AASC includes 133 sentences from Agat’angelos’ “History of the Armenians” full text in which words based on the root *surb* (*srb-*) appear. The subcorpus of 133 sentences includes 4,318 tokens and 2,010 types (including stop words).

“The Life of Maštoc” by Koriwn

The AASC includes 49 sentences from Koriwn’s “The Life of Maštoc’” full text where words based on the root *surb* (*srb-*) appear. The subcorpus of 49 sentences (including stop words) includes 1,307 tokens and 746 types.

The “Epic Histories” by Faustos of Byzantium

The AASC includes 199 sentences from the “Epic Histories” full text in which words based on the root *surb* (*srb-*) appear. The subcorpus of 199 sentences (including stop words) includes 5,708 tokens and 2,327 types.

The Gospels

The AASC includes 80 sentences as full text in which words based on the root *surb* (*srb-*) appear. The subcorpus of 80 sentences (including stop words) includes 1,455 tokens and 603 types.

The “Rights” and the “Cleaning” parts of the AASC are identical to those in the description of the AAFTC above.

5. Examination

The following analyses of the two corpora start with the investigation of the Armenian texts in the AAFTC/AASC and continue with the examination of the English HSEC in the second part. The methods and tools used in Chapter 5 were described in Chapter 3. The results of these analyses build the basis for the comparative part in Chapter 6.

The introductory sentences to the single steps in the macroanalysis have been kept for each study. Even though this might be repetitive for those reading through the whole book, it helps readers that will only look at selected parts of the examination.

5.1 Armenian Corpora

Even though the major part of the analysis of the Armenian corpora is based on the *Ancient Armenian Surb Corpus* (AASC), the *Ancient Armenian Full Text Corpus* (AAFTC) is applied as well. The AAFTC is mainly exploited in the “types” part of the examination during the keyword analysis and during the collocation analysis. The reason for this lies in the fact that the AAFTC has not been lemmatized (see the discussion below). Consequently, it cannot be employed in the examination parts that focus on the lemmatized version of the Armenian texts.

The examination of the Armenian corpora is subdivided according to the subcorpora of the AAFTC and the AASC (Koriwn, Agat`angelos, Epic Histories, Gospels).

5.1.1 Overview of the Examination of the Armenian Corpora

At the beginning of each examination, a keyword analysis between the AAFTC “types” version (reference corpus) and the sample in the AASC “types” version (node corpus) is conducted (see Chapter 3).

The word frequency list analysis and the collocation analysis are conducted twice, once for the “types” version of the Armenian texts and then again for the lemmatized version.

The following two short sections will provide a comprehensive overview of the applied methods that have also been described in Chapter 3.

5.1.1.1 Methods Used in the “Types” Examination

Each examination of the “types” version of the Armenian texts includes the following methods:

- Keyword analysis between the AASC (node corpus) and the AAFTC (reference corpus).
- Word frequency lists of the AASC version of the subcorpus (top 15 entries).
- Frequency lists of the inflected forms of *surb* in the AASC version of the subcorpus.¹
- Collocation analysis of terms related to *surb* in the AAFTC.

5.1.1.2 Methods Used in the “Lemmas” Examination

The study of the “lemmas” will only be conducted for the texts in the AASC. Due to the lack of a freely available Ancient Armenian lemmatizer, the lemmatization of the AASC has been made with a provisional automated lemmatizer written in *Python*. The main part of the lemmatization was based on the available lemmatized version of the Gospels and the Bible on *arak29.am*. I have manually added the missing lemmas of the AASC. Although a full lemmatization of the larger texts in the AAFTC remains to be conducted, the lemmatized texts of the AASC provide sufficient material for our present purpose. Concerning the examination of the lemmas, the main methods applied in this section are listed in the following overview. Due to the missing lemmatization of the AAFTC, neither a keyword analysis nor a collocation analysis based on the larger AAFTC was possible in the case of the “lemmas” version of the Armenian texts.

- Word frequency lists of the AASC version of the texts (top 15 entries).
- Collocation analysis of terms related to *surb* in the AASC version of the texts.

The next step following the statistical examination of the “types” and “lemmas” in the Armenian corpora is to select a representative sample of sentences from the

¹ Note that this method has not been described in Chapter 3 since it is not part of the HSEC examination. The frequency list of different forms of words related to *surb* only serves as an overview of the morphological variance of the use of *surb* in the texts.

corpus data for the microanalysis. Among others, this microanalysis helps to identify potential blind spots in the macroanalysis. The sample of sentences for the microanalysis is taken from a synoptic table that is based on the results of the statistical macroanalyses described above (see Chapter 3). The words in the synoptic table are annotated according to the schemes presented in Chapter 3. They build the basis for the semantic graphs that will be used in the comparisons conducted in Chapter 6.

5.1.2 AAFTC/AASC Agatangelos

5.1.2.1 Keyword Analysis (“Types” Only)

The tables below present the ten most important keywords and lockwords. The keywords of the AASC node corpus are in the left column and those of the reference corpus AAFTC in the right column. The lockwords are in the middle column (words that appear to a similar extent in both corpora). The first table (see the corresponding figure 11) includes the original Armenian terms. The second table contains the English translations of the Armenian terms. All words related to *surb* have been removed because they obviously are the most widespread keywords in the node corpus AASC. The table with the English translations only lists the English lemmas of the inflected Armenian words. Consequently, it does not reflect the morphological variance of the Armenian words in the first keyword table. For instance, the Armenian word հոգւոյն (“of the spirit”) could either be in the genitive, dative, or ablative singular. Yet, the inflection is not mirrored in the English table that only lists “spirit.”

Figure 11: Keywords and lockwords in the AASC (+) and AAFTC (-).

1/1718	Keywords +	1/92	Lockwords	1/6554	Keywords -
1	տեղւոջն	1	զամենայն	1	աշխարհին
2	պատմեսցուք	2	որք	2	պարսից
3	զոսկի	3	իւրով	3	յանձն
4	հոգւոյն	4	ամենայն	4	տուեալ
5	ընկերաւքն	5	անցանէ	5	անհնարին
6	եղիցի	6	ասացեալ	6	իցեն
7	երրորդութեանն	7	աստուածութիւնն	7	պատասխանի
8	զառիս	8	աւար	8	իշխանն
9	զտաճար	9	աւելի	9	իցէ
10	ճանապարհ	10	բարեաց	10	մարդկանն

Table 3: English translation keywords and lockwords in the AASC (+) and AAFTC (-).

Keywords(+)	Lockwords	Keywords(-)
place	all	land/world
to tell	PRON	Persian
gold	PRON	person
spirit	all	undef
companion	to die/end	violent
to be	to say	to be
trinity	divinity	answer
lamb	disorder	prince
temple	more	to be
path	good	men

Since the keyword analysis is based on “types,” and since Ancient Armenian is a highly inflected language, the results of the study of the keywords need to be interpreted with caution. According to the keyword analysis, words related to “place” (տեղի) in the locative case (տեղւոյ) are much more present in the AASC than in the AAFTC.² There are other spatial terms such as “path” (ճանապարհ) and “temple” (տաճար) as keywords of the AASC as well.

Leaving aside particular forms of verbs that are commonly used in the AAFTC (such as պատմել, “to tell”), another significant category are religious concepts such as “spirit” (fourth place), “trinity” (seventh place), and attributes or objects such as “gold” (third place). Two other notable terms are “with the companions” (fifth place) and “lamb” (eighth place).

Regarding the lockwords, it is important to bear in mind that the term “all” (ամենայն) is equally present in the AAFTC and in the AASC. This is relevant for the analysis because “all” occurs regularly among the most frequent terms in both the word frequency lists and the top entries in the collocation analysis. Thus, its prominence seems to be accounted by its frequent occurrence in the whole text rather than by its importance for the notion of *surb*.

Interestingly, the keywords in the reference corpus are connected to what one might call the profane sphere, particularly when compared to the religious terms that appear among the keywords of the AASC. These are: “World” or “land” (աշխարհի), ethnic or social terms such as “Persians” (պարսից) or “men”

2 I would like to remind the reader that the AASC only consists of sentences that include *surb*. Thus, words that appear in the “keywords (+)” column can reasonably be interpreted as more or less exclusively appearing in the context of *surb*. Words that appear in the “lockwords” column are less exclusive and might be less relevant for the meaning of *surb*.

(մարդկանն), negative attributes such as “violent, excessive” (անհնարին), or titles such as “prince” (eighth place, իշխան).

5.1.2.2 Word Frequency Lists

Types

No *Surb*

The fifteen most common words in the AASC Agat'angelos subcorpus (no stop words, no words related to *surb*) are shown in following table and bar plot (see the corresponding figure 12).

Table 4: Most frequent words (top 15) in AASC Agat'angelos.

Type	Translation	Frequency
ամենայն	all	28
աստուծոյ	god	22
քրիստոսի	Christ	12
հոգւոյն	spirit	12
զամենայն	all	11
տէր	lord	11
գրիգորիոս	Gregory	11
աստուած	god	10
առաջի	before	10
գրիգորի	Gregory	9
արժանի	worthy	9
արարեալ	to make	9
գայիանէ	Gayianê	9
հռիփսիմէ	Hřip'simê	9
աստուածութեան	godness	8

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common words are (no *surb*; overall frequency: 180):

- Words related to religious “Concepts” (աստուծոյ, աստուած, աստուածութեան, տէր, քրիստոսի, հոգւոյն) (AF 74, RF 41%)
- Words related to “Persons” (գրիգորիոս, գրիգորի, հռիփսիմէ, գայիանէ) (AF 41, RF 23%)
- Words related to general “Concepts” (ամենայն, զամենայն) (AF 39, RF 22%)
- Attributes such as “worthy” (արժանի) (AF 9, RF 5%)
- Verbs such as “to make (արարեալ) (AF 9, RF 5%)

There is a strong focus on individual persons (Gregory, Hřip’simē, Gayianē, AF 41) and on mostly religious concepts (god, Christ, spirit, AF 74). The attribution of “worthy” is also present to a noticeable degree (AF 9).

Figure 12: Most frequent words (top 15) and their absolute frequencies in AASC Agat’ angelos.

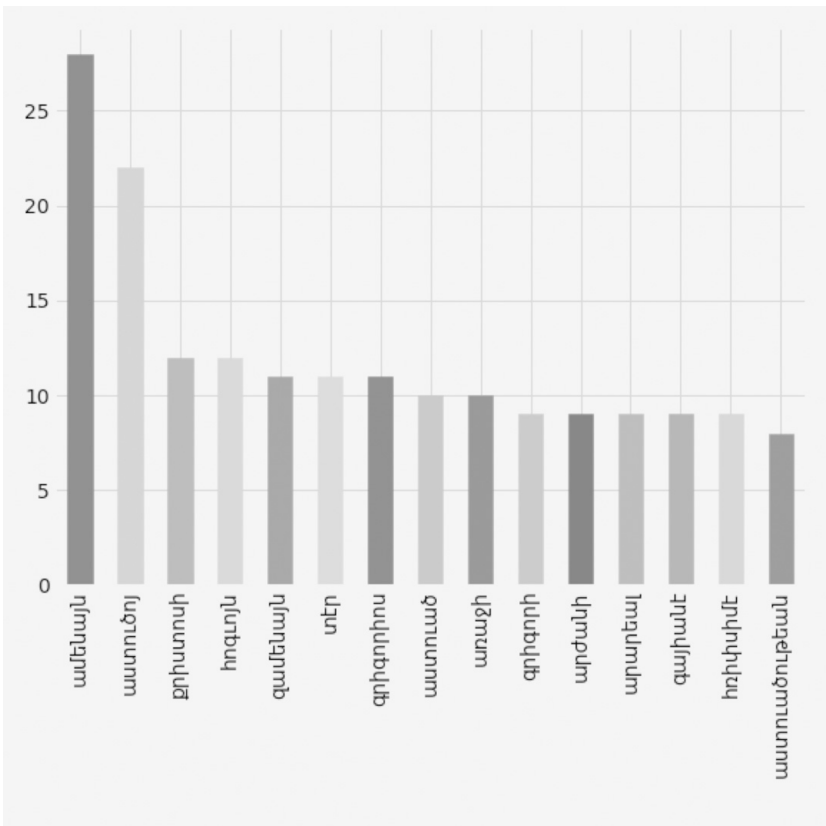


Table 5: All inflected forms of lexical units based on the root surb/srb- and their absolute frequencies in AASC Agat' angelos

Type	Frequency
սուրբ	39
սրբոյ	26
սրբոցն	19
սուրբն	17
զսուրբ	14
սրբոյն	11
սրբոց	8
զսուրբն	7
սրբութեան	7
սրբութեամբ	2
զսուրբսն	2
զսուրբս	1
սրբութեանդ	1
սրբեալք	1
սուրբք	1
սուրբս	1
սրբով	1
սրբութենէ	1
սուրբսն	1
սրբիչ	1
սրբեալ	1
սրբել	1
սրբութիւնն	1
սրբոցդ	1
զսրբութիւնս	1
սրբամատոյց	1
սրբասնելովն	1

Table 6: Most frequent lemmas (top 15) in AASC Agat'angelos.

Lemma	Translation	Frequency
ամենայն	all	42
առնել	make	39
աստուած	god	38
մօտ	near	31
հոգի	spirit	27
թագաւոր	king	24
որդի	son	23
աստուածութիւն	godness	21
անուն	name	21
տէր	lord	18
տեղի	place	18
միասին	only-begotten	17
ձեռն	hand	17
մեծ	great	17
տեսանել	to see	16

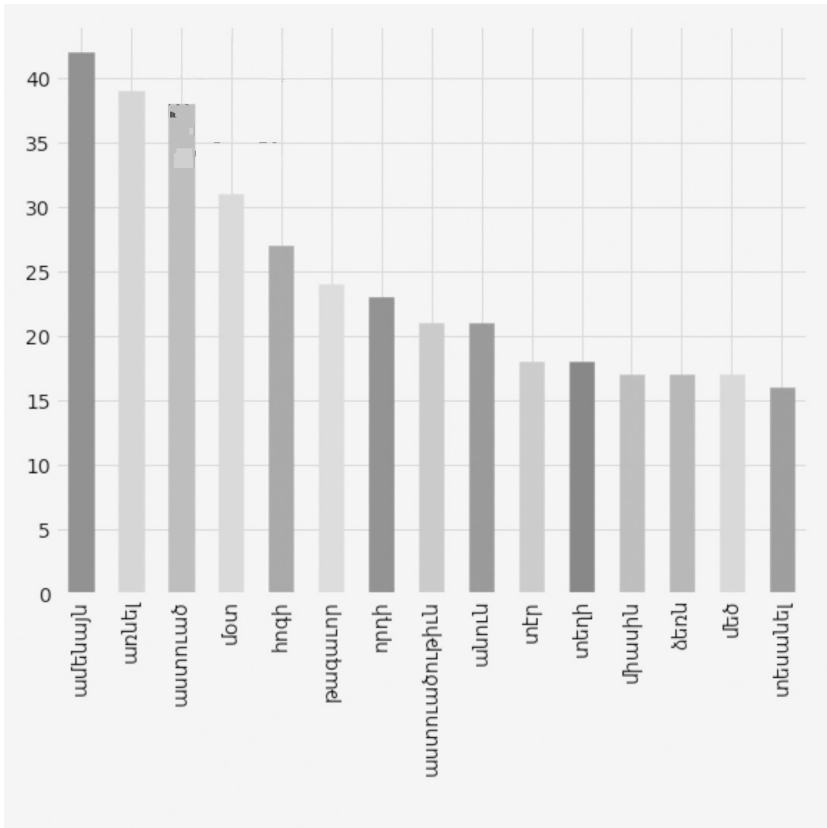
Lemmas

The fifteen most common lemmas in the AASC Agat'angelos subcorpus (no stop words, no *surb*) are shown in following table and bar plot (see the corresponding figure 14).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common lemmas are (no *surb*; overall AF: 369):

- Words related to religious “Concepts” (հոգի, միասին, աստուածութիւն, աստուած, տէր) (AF 121, RF 33%)
- Words related to general non-religious “Concepts” (անուն, ձեռն, որդի, ամենայն) (AF 102, RF 28%)
- Verbs such as “to do” (առնել, տեսանել) (AF 55, RF 15%)
- Words related to “Places” (տեղի, մօտ) (AF 49, RF 13%)
- Words related to “Titles” (թագաւոր, տէր) (AF 42, RF 11%)
- Attributes such as “great” (մեծ) (AF 17, RF 5%)

Figure 14: Most frequent lemmas (top 15) in AASC Agat’ angelos (bar plot).



Similarly to the word frequency list of “types,” religious concepts such as “spirit” or “god” dominate the word frequency list of “lemmas” as well. Yet, unlike what was highlighted in the case of “types,” the word frequency list of “lemmas” does not include individual persons. In addition, the category of non-religious “concepts” is much more present and includes words such as “name” and “all.” The “place” category was not part of the “types” version of the word frequency list either.

5.1.2.3 Collocation Analysis (AAFTC/AASC, Top 15)

Types

The fifteen most common words in the AAFTC Agat’ angelos subcorpus (no stop words) that appear in the context of words related to *surb/srb-* are shown in the

following table (6-LL, L2-R2, C3-NC3; no function words). The visualization is displayed in the corresponding figure 15.

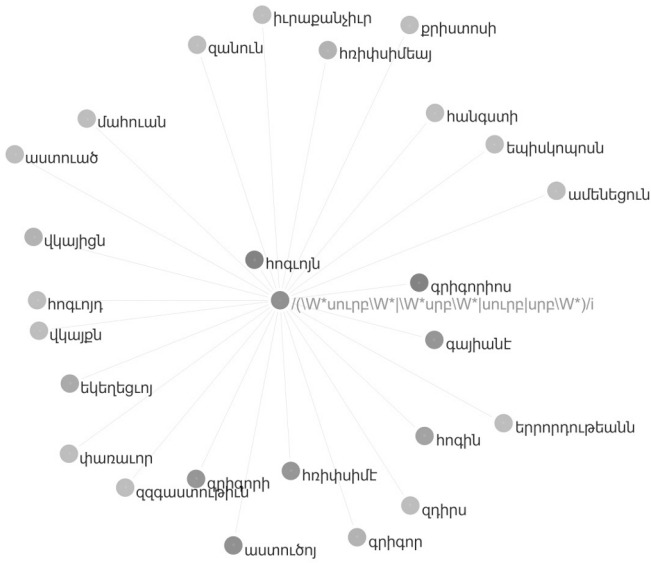
Table 7: Top 15 words collocation analysis (6-LL, L2-R2, C3-NC3; no function words) *surb/srb- in AAFTC Agat' angelos.*

id	pos	coll	translation	stat	freq_(coll)	freq_(corpus)
1	l	հոգւոյն	spirit	102.80	12	16
2	r	գրիգորիոս	Gregory (name)	68.37	12	46
3	r	գայիանէ	Gayianê (name)	61.29	8	14
5	r	գրիգորի	Gregory (name)	47.09	8	28
4	r	հռիփսիմէ	Hřip'simê (name)	56.94	8	17
6	l	հոգին	spirit	46.78	6	10
7	r	եկեղեցւոյ	church	36.33	5	10
8	l	զդիրս	place	30.05	3	3
9	l	զգգաստութիւն	chastity	30.05	3	3
10	l	հանգստի	repose	30.05	3	3
11	l	հոգւոյդ	spirit	30.05	3	3
12	r	վկայքն	martyrs	30.05	3	3
13	r	գրիգոր	Gregory (name)	27.78	4	9
14	r	աստուծոյ	god	27.64	9	122
15	r	հռիփսիմեայ	Hřip'simê (name)	25.75	4	11

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 91).

- Words related to “Persons” (գրիգորիոս, գայիանէ, հռիփսիմէ, գրիգորի, գրիգոր, հռիփսիմեայ) (AF 52, RF 48%)
- Words related to religious “Concepts” (եկեղեցւոյ, աստուծոյ, հոգւոյն, հոգին, հոգւոյդ) (AF 35, RF 38.5%)
- Words related to “Places” (զդիրս, հանգստի) (AF 6, RF 6.5%)
- Terms related to “Purity” (զգգաստութիւն) (AF 3, RF 3.5%)
- Words related to “Titles” (վկայքն) (AF 3, RF 3.5%)

Figure 15: Top 15 types collocation analysis (6-LL, L2-R2, C3-NC3; no function words) *surb/srb-* in AAFTC Agat’ angelos (graph) (c) LancsBox.



The collocation analysis of the “types” shows that terms related to *surb* are to a significant degree connected to individual persons such as Gregory or Hřip’ simē. This observation was already made during the analysis of the word frequency lists. Another important category consists of “concepts” that mostly have a religious connotation, such as “spirit” or “god.” Furthermore, the presence of concepts related to sexual purity (“chastity”) is noticeable as well. Yet, the latter domain is not part of the “domain” annotation system in Chapter 3. Nevertheless, since the discussion of the importance of purity and impurity in the context of the Armenian corpora will play a central role in the forthcoming discussions, it has been listed separately in this part of the analysis.

Lemmas

The fifteen most common words in the AASC Agat’ angelos subcorpus (no stop words) that appear in the context of the lemma *surb* are shown in following table (6-LL, L2-R2, C3-NC3; no function words). The visualization is displayed in the corresponding figure 16.

Table 8: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb in AASC Agat' angelos.

id	pos	coll_(lemma)	translation	stat	freq_(coll)	freq_(corpus)
1	l	հոգի	spirit	135.42	23	27
2	r	վկայ	martyr	88.31	13	13
3	r	գրիգորիոս	Gregory (name)	67.75	11	12
4	r	գրիգոր	Gregory (name)	47.60	9	12
5	r	հռիփսիմէ	Hr̄ip'simê (name)	42.98	9	14
6	r	գայիանէ	Gayianê (name)	41.38	8	11
7	m	անուն	name	39.39	10	21
8	r	Երրորդութիւն	trinity	28.36	5	6
9	l	գունդ	host	26.93	4	4
10	r	կաթողիկոս	katholikos	26.93	4	4
11	l	հանգիստ	repose	25.46	5	7
12	m	եկեղեցի	church	24.25	6	12
13	r	աստուած	god	21.24	9	38
14	l	մասն	_	20.18	3	3
15	l	վկայարան	martyr's tomb	16.12	4	8

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 123).

- Words related to primarily religious but also non-religious “Concepts” (հոգի, եկեղեցի, անուն, Երրորդութիւն, գունդ, եկեղեցի, աստուած) (AF 57, RF 46%)
- Words related to “Persons” (գրիգորիոս, գրիգոր, հռիփսիմէ, գայիանէ) (AF 37, RF 30%)
- Words related to “Titles” (կաթողիկոս, վկայ) (AF 17, RF 14%)
- Words related to “Places” (հանգիստ, վկայարան) (AF 9, RF 7%)

Figure 16: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3, no function words) *surb* in AASC *Agat’ angelos* (graph) (c) LancsBox.



Compared to the collocation analysis of the “types,” the collocation analysis of the “lemmas” reveals a greater importance of mostly religious concepts, such as “church” or “spirit.” That being said, the “persons” category, which mostly includes individual names, and its connection to *surb* is still very present. This is particularly the case when one considers that the “titles” domain often stands for persons as well.

5.1.2.4 Synoptic Table – Sentences Selection and Discussion (Microanalysis)

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses and their corresponding values and annotation has been created.

Table 9: Synoptic table of the AASC and AAFTC Agat'angelos.

id	lemma	kw	wf_t	cf_t	wf_l	cf_l	sum	do- main	semantic_field
1	տեղի	1	0;0	0;0	18	0	25	place	place_M2
2	պատմել	1	0;0	0;0	0	0	7	verb	narrate_Q2.1
3	ոսկի	1	0;0	0;0	0	0	7	at- tribute	golden_O4.3
4	հոգի	1	1;12	3;23	27	23	98	con- cept	spirit_S9
5	ընկեր	1	0;0	0;0	0	0	7	con- cept	companion_S3.1/ S2mf
6	երրորդութիւն	1	0;0	0;0	0	5	12	con- cept	trinity_Z3c
7	գառն	1	0;0	0;0	0	0	7	ob- ject	lamb_L2mfn
8	տաճար	1	0;0	0;0	0	0	7	place	temple_S9/H1
9	ճանա- պարհ	1	0;0	0;0	0	0	7	place	path_M3
10	ամենայն	-1	2;39	0;0	42	0	72	con- cept	all_N5.1+
11	աստուած	0	3;40	1;9	38	9	100	con- cept	god_S9
12	քրիստոս	0	1;12	0;0	0	0	10	con- cept	Christ_Z1m
13	տէր	0	1;11	0;0	18	0	27	title	Lord_S7.1+/S2.2m
14	գրիգոր	0	2;20	3;24	0	9	56	per- son	Gregory_Z1mf
15	արժանի	0	1;9	0;0	0	0	7	at- tribute	worthy_S1.1.4+
16	առնել	0	1;9	0;0	39	0	47	verb	make_A1.1.1
17	գայիանէ	0	1;9	1;8	0	8	24	per- son	Gayanê_Z1f
18	հռիփսիմէ	0	1;9	2;12	0	9	30	per- son	Rhipsime_Z1f
19	եկեղեցի	0	0;0	1;5	0	6	9	con- cept	church_S9/H1c
20	դիր	0	0;0	1;3	0	0	1	place	place_M2
21	զգաստութիւն	0	0;0	1;3	0	0	1	con- cept	chastity_S3.2/G2.2+
22	հանգիստ	0	0;0	1;3	0	5	6	place	repose_E3+

23	վկայ	0	0;0	1;24	0	13	36	title	martyr_S9/S2mf
24	մօտ	0	0;0	0;0	31	0	28	concept	near_N3.3-
25	թագաւոր	0	0;0	0;0	24	0	21	title	king_S7.1+/S2.2m
26	որդի	0	0;0	0;0	23	0	20	concept	son_S4m
27	աստուածութիւն	0	0;0	0;0	21	0	18	concept	divinity_S9
28	անուն	0	0;0	0;0	21	10	28	concept	name_A1
29	միասին	0	0;0	0;0	17	0	14	title	only-begotten_Z99
30	ձեռն	0	0;0	0;0	17	0	14	object	hand_B1
31	մեծ	0	0;0	0;0	17	0	14	attribute	great_A5.1+
32	տեսանել	0	0;0	0;0	16	0	13	verb	see_X3.4
33	գրիգորիոս	0	0;0	0;0	0	11	8	person	Gregory_Z1m
34	գունդ	0	0;0	0;0	0	4	1	concept	host_S1.1.1
35	կաթողիկոս	0	0;0	0;0	0	4	1	title	katholikos_S7
36	մասն	0	0;0	0;0	0	3	0	???	???
37	վկայարան	0	0;0	0;0	0	4	1	place	tomb_L1-/H1

As explained in Chapter 3, the selection of example sentences for the micro-analysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *surb* appears in the neighborhood of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

The domains shown in table 10 appear in the “History of the Armenians” by Agat’angelos (ordered by their summarized “sum” values).

Concept

With a “sum” value of 100, the lemma աստուած (“god”) is the most frequent word in the “concept” domain. The lemma աստուած (“god”) is not only the most

frequent word in the “concept” domain but the most important word in the overall examination of the “History of the Armenians” (with a “sum” value of 100).

Table 10: Domains in the AAFTC/AASC Agat' angelos in a descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	404
Person	118
Title	99
Verb	67
Place	47
Attribute	28
Object	21

Վասն կենսատու փրկութեանդ հասելոյ ձեզ յԱստուծոյ, որպէս մեզ պատմեցաւ, անպայման ուրախութեամբ գոհացաք զանզեղջական փառացն Քրիստոսի, եւ վասն սրբոց վկայիցդ Աստուծոյ որ ի ձերում միջի, յաշխարհիդ ձերում փառաւորեցան: (§821)

For the life-giving salvation bestowed on you by God, as has been narrated to us, with unbounded joy we have thanked the immutable glory of Christ, as also for the holy martyrs of God who were glorified among you, in your country.

Discussion

The use of “god” in the context of *surb* in the sentence above (§821) is a typical example of the general use of “god” in the Armenian texts. God is only seldom called *surb*. However, god often appears in a close context of words related to *surb*, frequently applied as another attribute for persons or things that are called *surb*, such as in this example (“(for) the holy martyrs of God” ([վասն] սրբոց վկայիցդ Աստուծոյ)). Consequently, it is not god who is in the center of *surb* in this sentence, but the martyrs who are granted several attributes, one of which is *surb*. The other attributes are “of God” (Աստուծոյ) and the verb փառաւորել (“to glorify”).

When looking beyond the syntactical boundaries of the direct context of *surb* in this sentence, it becomes clear that the entire sentence is full of different attributes. Most of them belong to semantic fields of “honor,” “glory,” or other very positively connotated expressions (among others, կենսատու փրկութեանդ [“live-saving salvation”], անպայման ուրախութեամբ [“with unbounded joy”], զանզեղջական փառացն Քրիստոսի [“immutable glory of Christ”], etc.). Thus, in §821, *surb* is only one among many other attributes used to honor the martyrs. The attribution “of god” is also one of these attributes.

The observation that the term “god” is often used together with *surb* is very much in line with what has been stated in Chapter 2. As indicated, in a Christian context, everything that is *holy* derives its status of holiness directly from god. This characteristic of *holy* in a broader Christian understanding is thus present in the use of *surb* in the “History of the Armenians” as well, thereby demonstrating the integration of the meaning and application of *surb* into the broader Christian cosmos of late antiquity.

Person

The second most prominent domain in the synoptic table is the “person” domain. With a “sum” value of 56, the lemma գրիգոր (Gregory) is the most conspicuous word in the “person” domain. Gregory the Illuminator is one of the most important persons in the “History of the Armenians.” According to the legend told in this account, Gregory was the first bishop of Armenia and the founder of the Armenian Church following his conversion of king Trdat (see Chapter 2 and the description of the “History of the Armenians” in Chapter 4).

Ապա աղաչեաց երանելի թագաւորն Տրդատ զսուրբն Գրիգորիոս, զի փոխանակ այնր, զի ոչ հաւանեցաւ նա կալ եւ շրջել ընդ նմա, եւ սիրեաց զմենաստորութեան կեանս՝ ձեռնադրեսցէ եւ տացէ նմա եպիսկոպոս զսուրբ որդին իւր, զոր ետ ածել ինքն՝ զԱրիստակէս: (§862)

Then the blessed king Trdat begged saint Gregory that since he had not agreed to remain and go around with him and because he loved the solitary life, he would in return ordain and give him as bishop his saintly son Aristakes, whom he had brought.

Discussion

On the one hand, this sentence is a typical example of the use of *surb* in the context of (individual) persons in the larger account of the “History of the Armenians.” On the other hand, it also includes a remarkable exception that will be discussed below.

What is typical about the use of *surb* in this sentence is its direct connection to a person. In contrast to the example of god in the “concept” domain, the word *surb* is often directly applied as an attribute for persons. Interestingly, all persons that appear in this sentence have a direct attribute. Two persons are called *surb* (Gregory and his son Aristakes) and one is called երանելի (“blessed”), namely the king Trdat whose conversion to Christianity started the official Christianization of Armenia. Besides other attributes in the form of titles, such as “king” in case of Trdat and “son” in case of Aristakes, Gregory and Aristakes are also characterized as “typical persons belonging to a religious institution” through their conduct of life (“solitary life”) and their titles (“bishop”). Unlike in most other Christian traditions,

the hereditary character of church titles is typical for the early Armenian Church, where the title of a bishop was commonly passed on from the father to one of his sons. Thus, judging by its use in this sentence, the attribution of *surb* seems to refer to persons with a religious conduct of life who also hold a higher position in the Church hierarchy.

Surprisingly, the adjective of “blessed” (Երանելի) is attributed to a worldly leader in this sentence, namely king Trdat. In the Armenian accounts used in this book, this attribute is usually exclusively granted to persons who hold higher positions in the church hierarchy or to biblical persons or martyrs.

A collocation analysis³ of Երանելի (“blessed”) in the AAFTC Agat ‘angelos demonstrates that “blessed” is almost exclusively used with Gregory. That Trdat is portrayed as “blessed” (Երանելի) is, therefore, an exception from this rule and might hint at the fact that the actual meaning of the word “blessed” (Երանելի) is not restricted to a religious sphere but describes a more general state (of happiness). The question, then, is whether the term *surb* has a more general meaning as well. Yet, in opposition to the use of “blessed” (Երանելի), the word *surb* is never used for non-religious persons in the Armenian accounts examined in this book.

However, the term *surb* is also employed in contexts that are not related to religion but which belong to the sphere of purity, even in a physical sense. Similarly to the attribution of “blessed” (Երանելի) to a worldly leader, these applications of *surb* are rather seldom. However, they might also point to an earlier layer of the semantic field of *surb* that is still present, even though the terminology has shifted its core semantic field towards religion. Some examples of this more general use of *surb* in the context of purity will be discussed in the course of this analysis.

Title

With a “sum” value of 36, the lemma վկայ (“martyr”) is the most conspicuous word in the “title” domain. Since the common occurrence of martyrs in the context of *surb* has already been discussed in the first example of this section, this part will focus on the third most frequently occurring word, namely “king” (թագաւոր).⁴

That the “king” is among the most frequent titles is unexpected since one would usually anticipate a religious title such as “bishop” (եպիսկոպոս). Even though titles such as “bishop” or “katholikos” are found in the synoptic tables as well, they are not among the most frequent words. King Trdat’s prominent role in the “History of the Armenians” may well account for the frequent occurrence of “king” in this context.

However, neither king Trdat nor any other king, for instance, Diocletian, is called

3 (6-LL, L2-R2, C2-NC2; no function words)

4 The second-placed word “lord,” a title that is almost exclusively used for god, has also been discussed already.

surb. Yet, they often appear in the context of persons, places, or objects that are called *surb*. The kings, as well as other worldly leaders, often have an ambiguous characterization in the “History of the Armenians.” On the contrary, persons that are called *surb* are always described with very positive attributes. The characterization of kings ranges from negative to neutral and positive. Since a positive representation of a king in the “History of the Armenians” has already been discussed in a previous part of this section, it is now time to present both an example of a negative characterization and a more neutral one.

Իսկ երանելին պարկեշտասերն Գայիանէ, սրբասնելովն Հռիփսիմեաւ հանդերձ եւ այլ ընկերաւքն իւրեանց, յիշեալ զուխտն սրբութեան, զարինաւոր կրանիցն զգաստութեան սրբութիւնն յոր մտեալ էին՝ ողբային յանձինս իւրեանց վասն նկարացոյց հրամանին, առ ի նկատել զպատկերս նոցա պիղծ եւ անաւրէն թագաւորին. (§143)

But the blessed and chaste Gaiane, with the saintly Rhipsime and their other companions, remembered the covenant of holiness, the religious rule of chastity into which they had entered, and lamented amongst themselves over the impure and impious emperor's command to have their portraits painted.

Եւ մատուցեալ թագաւորն եւ նախարարքն՝ բռնն հարեալ ունէին զտոյց սրբոյն Գրիգորի եւ ասէին. «Խնդրեմք ի քէն, թո՛ղ մեզ զյանցումն մեր, զոր մեք ընդ քեզ արարաք»: (§143)

The king and the princes approached, took hold of saint Gregory's feet and said: “We beg you, forgive us the crime that we committed against you.”

Discussion

The first example (§143) is crucial for sketching a semantic field of *surb* due to several reasons.

Firstly, it presents the use of *surb* in a long chain of attributes⁵ and characterizations applied to define two central persons in the “History of the Armenians” together with their companions. These two persons are the female martyrs Gayiané and Hřip' simé who were both killed by Trdat and who today remain major martyrs and saints of the Armenian Church. The episode reported in §143 describes them as antagonistic to the “impure” (պիղծ) and “impious” (անաւրէն) king Diocletian. The vocabulary used in this sentence (“impure” and “impious”) is often found in the context of *surb* and hints at two different aspects.

It demonstrates that the semantic field of impurity is regarded as an antonym of

5 The term “chain” has been deliberately chosen to underline that the attributes are intertwined. They are applied together to provide a general characterization of the respective persons.

surb. This, in turn, proves that *surb* is related to the semantic field of purity. This passage also shows that there existed a semantic field opposite to that of *surb*, namely that of impurity (պիղծ). Though trivial as it may seem, this observation is worthy of mention because it indicates that the semantic field of *surb* implies a sharp distinction (i.e. purity/impurity) similar to the *sacred/profane* distinction discussed in Chapter 2. This demarcation between two distinct spheres is here represented by the king on the one hand, and the nuns on the other, as well as by their respective characterization. This observation can rightfully be taken as indicative of the existence of a semantic field of *holy* in the Armenian terminology that corresponds to the standards articulated in Chapter 2.

Besides the opposition between the two spheres of purity and impurity inherent to the use of *surb*, it is also important to note that *surb* is once more employed as one attribute in addition to many others. The other attributes in this sentence (§143) are “blessed” (երանելի), “chaste” (պարկեշտասեր, զգաստութեան), and “religious rule” (զարհնաւոր կրանիցն). As previously mentioned, the attribute “blessed” (երանելի) is very common in the context of persons that are called *surb*. The religious dimension is also pervasive in this context, although the translation of զարհնաւոր կրանիցն as “religious rule” presents a problem of its own. Another dimension that has not been apparent in the microanalysis yet is that connected to sexuality and sexual purity in particular. The latter is an important motif in this sentence and was present in the macroanalysis (collocation analysis) as well.

Verb

With a “sum” value of 47, the verb առնել (“to do”) is the most important word in the “verb” domain.

«Քանզի չէաք իսկ հանդուրժող զամենայն արարեալս սրբոցն գտակաւ զիրաքանչիրսն նշանակել, այլ ի դիրագոյն եւ ի հեշտագոյն յառաքելականն անդր զանձինս պատսպարեցաք, որոյ անցեալ զբազմախոռն արգասեալ սրբոցն առ ի մանրակրկիտ առնելոյ՝ զկարեւորագոյնսն եւ զաւագտակարագոյնսն պատմեաց»:
(§898)

For we were not able to indicate precisely every detail of all that was done by the saints, but we have taken refuge in the easy and delightful and apostolic practice (of Luke), who passing over the many and various deeds of the saints omitted the details and narrated (only) the most important and most profitable points.

Discussion

It is challenging to abstract the typical use case of general verbs such as “to do” (առնել) in the context of *surb*. The lemma առնել (“to do”) often appears in ref-

erence to martyrs. The close relationship between “to make,” “martyrs,” and *surb* underlines the importance of deeds (as can be seen in this example) and of the general conduct of life (in the previous example of Gregory) for persons that are called *surb*. This is an important point because recent discussions on *holy* have largely overlooked the relevance of concrete actions or conduct of life for the signification of the contemporary notion of *holy*. As evoked in Chapter 2, contemporary debates on *holy* are more concerned with belief systems, religion, cognitive aspects, emotions, etc.

Conversely, the dimension of “actions” is pervasive in the “History of the Armenians.” This observation is in agreement with other early Christian traditions, where the focus on martyrs and their deeds shaped the early Christian notion of *holy* as well (Angenendt 2007; Brown 2015).

Place

With a “sum” value of 25, the lemma տեղի (“place”) is the most noticeable word in the “place” domain.

Իսկ երեք եւս, որ ունէին զտեղի վկայարանացն հանգստի սրբոցն՝
զի ի տեղուոջ հեղման արեանն իրեանց շինեսցին վկայարանք
հանգստի ոսկերաց իրեանց: (§748)

But the other three represented the places for the chapels of repose of the saints, because in the place where their blood was shed will be built chapels of repose for their bones.

Discussion

This sentence is an enlightening example as it also contains some of the other words in the “place” domain that frequently appear in the context of *surb*, namely the “chapels” or better “martyria” (վկայարան) and the “reposes” (հանգիստ). The saints are called սուրբք (*surbk*), which means that in the case at hand, *surb* is applied as a collective noun to represent a group of persons (mostly martyrs). The observation that saints and the semantic field of *holy* are closely related to places is rather common in the context of early Christianity (Angenendt 2007; Brown 2015). Consequently, the close connection between persons that are called *surb* and places where these saints were buried or where they had died is yet another indication that in Armenian texts, the term *surb* corresponds to other early Christian notions of *holy*.

Object

With a “sum” value of 14, the lemma ձեռն (“hand”) is the most important word in the “object” domain. However, as this word is frequently used in expressions that are detached from the meaning “hand,” the second most often mentioned lemma

գառն (“lamb”) will be discussed instead. Overall, the “object” domain plays a minor role in the “History of the Armenians” compared to the other domains, at least in the macroanalysis.

Եւ մի՛ տարցի զզգաստութիւն սուրբ հաւտի քոյ գազանն ապստամբ.
Եւ մի՛ յաղթահարեսցէ զգառնս հաւտի քոյ գայլն ապականիչ. (§145)
Let not the rebellious beast carry off the virtue of your holy flock. Let not the
corrupting wolf prevail over the lambs of your flock.

Discussion

As Thomson rightly pointed out, this passage is borrowing metaphors and allegories from the Old Testament and New Testament (Ez 34) where motifs related to shepherds and flocks are ubiquitous.

As it is often witnessed in a Christian context, the flock in this example refers to the entire Christian community; the latter is consequently characterized as *surb* as well. The idea that every Christian is regarded as *holy* (and not only a few chosen martyrs or exceptional persons) is also common in other early Christian traditions (see Chapter 2).

However, in the Armenian accounts studied in this book, the term *surb* mostly appears in reference to martyrs, high church officials, and other extraordinary persons. Therefore, this passage provides an interesting exemption from this rule that is still in line with broader early Christian conceptions of *holy*.

Attribute

In the attribute domain, the most important word is մեծ (“great”) with a “sum” value of 14. The second place is shared by two words that both have a “sum” value of 7. The first word is ոսկի (“golden”/“gold”) and the second word is արժանի (“worthy”). Since there are only three words in this category, example sentences for all three words are discussed in the following part.

Եւ մինչդեռ կազմէին զայն՝ սկսաւ ասել սուրբն Գայիանէ ընկերաւքն հանդերձ այսպէս. «Գոհանամք զքէն, Տէր, որ արժանի արարեր զմեզ մեռանել վասն մեծի անուանդ քոյ, եւ յարգեցեր զհողեղէն բնութիւնս մեր, զի լիցուք արժանի Աստուածութեանդ քում. եւ հաղորդեցեր զիս մահուան սրբոց քոց վկայիցն, Հռիփսիմեայ եւ ընկերացն: (§207)

And while they were setting these out, saint Gaiane with her companions began to speak as follows: “We thank you, Lord, for making us worthy to die on behalf of your great name, and for honoring our earthly nature so that we might become worthy of your divinity, and for making me share in the death of your holy martyrs, Rhipsime and her companions.

Զորս որդեգրութեան արժանիս արասցէ Աստուածութեանն, ջնջել

զմեղս ձեր եւ թողուլ զպարտիս ձեր, եւ ընդունել զմասն պսակի սրբոց որ ձեզս են: (§719)

He will make you worthy of divine adoption, wash away your sins and forgive your debts, and grant you a share of the crown of these saints who are among you;

Եւ բերէին զպարգետան տուեալս անդասին, զոսկին եւ զարծաթն եւ զպատուական կարասին, ի նուէր սպասու եկեղեցւոյն Աստուծոյ, եւ ի տունս նուիրաց սրբոյ վկայիցն դնէին. նաեւ զտուեալ զոսկի սպասս կայսերն ի նոյն հանգիստ սրբոցն դնէին: (§882)

And they brought the gifts they had been given as an offering for the service of the church of God, gold and silver and precious furniture; these they placed in the chapels of the holy martyrs. Likewise they placed in the same (chapels) of repose of the saints the gold vessels of the emperor that they had been given.

Discussion

The words մեծ (“great”) and արժանի (“worthy”)⁶ are attributes that commonly appear in the context of *surb* where they characterize religious and non-religious actors. Opposed to *surb*, the words արժանի (“worthy”) and մեծ (“great”) are also applied to non-religious persons or concepts, such as kings, nobles, or names.

The word ոսկի (“golden”/“gold”) is frequently found in the context of *surb* but it is never directly related to it. Ոսկի (“golden”/“gold”) is mostly connected to objects and is applied as an expression conveying gratitude and admiration through extraordinary gifts that are found in religious and non-religious contexts alike. Both words underline the extraordinary and royal atmosphere that is evoked in the context of *surb*. This atmosphere is created by including expensive material objects and by applying distinguishing notions such as “worthy” (արժանի) or “blessed” (երանելի).

5.1.2.5 Final Discussion

A microanalysis of the example sentences from each of the domains in the synoptic table has yielded the following results:

- The notion of *surb* frequently appears in the context of religious concepts, persons, or institutions. A telling example is the notion of god (աստուած) which is often used in the proximity of *surb* to characterize extraordinary religious persons (for instance, Gregory the Illuminator or the nuns Gayanê and Hřip ‘simê).
- The term *surb* is frequently used in direct connection to either individual persons or groups of persons. In this context, the terminology around *surb* either serves as an attribute or as an epithet for the persons. Most of the persons that

6 Both attributes could also be rightfully counted among the words in the “concept” domain.

are called *surb* can be described either as martyrs or as important persons in the Church hierarchy. Typical characteristics of these persons are their solitary life and their extraordinary deeds. Another illustrative example of this application of *surb* is found in §138 of the “History of the Armenians.” However, some exceptions hint at the fact that not only distinguished persons but also the entire community of Christians were regarded as *surb*.

- *Surb* is only rarely used in a direct combination with objects or places, at least from the perspective of the macroanalysis. That being said, as observed in the case of golden deposits in the repose of the saints, places, as well as objects, may play an important role in the wider context of *surb*.
- The term *surb* is often used in combination with other attributes. Particularly significant are the attributes արժանի (“worthy”) and “blessed” (երանելի) which frequently appear in the context of *surb*. Several other laudatory and distinguishing notions occur together with *surb* as well. For instance, the *surb* nuns Gayianê and Hřip’ simê in §143 are also called “virtuous” and “chaste.” It is interesting to note that some of the other attributes appearing in the context of *surb* can characterize religious and non-religious persons (for instance, արժանի, “worthy”). In contrast, others are only rarely or even never used for non-religious persons (երանելի, “blessed,” or *surb*).
- The passage from the “History of the Armenians” that is concerned with the martyrdom of Gayianê and Hřip’ simê reveals the close connection between *surb* and purity. In this context, the nuns Gayianê and Hřip’ simê, as well as their companions, are frequently described as *surb*, “virtuous,” and “chaste.” Due to its proximity to terms related to sexual purity or impurity (in the case of the emperor), the term *surb* also shifts to an almost similar meaning. The closeness between *surb* and purity is often reflected in the translations as well, for example, in the translation of §138 where սրբամատոյց is translated as “pure.”

Յայնմ ժամանակի եկեալ գտանէին ի քաղաքին Հռոմայեցոց արգելալանս մի կուսանաց՝ միանձնական լեռնական, ընդակերս, զգաստացեալս, պարկեշտականս, սրբամատոյց կանայս քրիստոսական հաւատոց, որ զցայգ եւ զցերեկ եւ յամենայն ժամանակի փառաւորութեամբ եւ արհնութեամբ զկատարեալ աղաթսն իրեանց առ Աստուած առաքել ի բարձունս արժանի լինէին: (§138)

Then they came and found in the city of Rome a convent of nuns, living solitary hermetic lives, eating vegetables, sober, modest, and pure women of the Christian faith, who day and night and the whole time by praising and blessing were worthy to raise to God in the heights their perfect prayers.

- The close connection between purity and *surb* is also observable in other Armenian sources in this book. At this point, one may ask in which way the notion of purity is congruent with the notion of *surb*. Is pure a subfield or an older stratum of a later more diverse semantic field of *surb* that also includes other connotations (and can thus not be restricted to the meaning of pure)? Or is the meaning of pure dominant in the semantic field of *surb*? In which case, should other passages where *surb* appears consequently also be translated as “pure” rather than “holy?”

5.1.3 AAFTC/AASC Epic Histories

5.1.3.1 Keyword Analysis (“Types” Only)

The tables below present the ten most important keywords and lockwords. The keywords of the AASC node corpus are in the left column and those of the reference corpus AAFTC in the right column. The lockwords are in the middle column (words that appear to a similar extent in both corpora). The first table (see the corresponding figure 17) includes the original Armenian terms. The second table contains the English translations of the Armenian terms. All words related to *surb* have been removed because they obviously are the most widespread keywords in the node corpus AASC. The table with the English translations only lists the English lemmas of the inflected Armenian words. Consequently, it does not reflect the morphological variance of the Armenian words in the first keyword table.

Figure 17: Keywords and lockwords in the AASC (+) and AAFTC (-).

1/1966	Keywords +	1/103	Lockwords	1/10038	Keywords -
1	կաթողիկոսն	1	ածել	1	պարսից
2	եպիփան	2	ամ	2	օրինակ
3	հոգւոյն	3	արարածոց	3	յօր
4	խաղ	4	բազմութեամբ	4	զաւրավարն
5	հոգւովն	5	գործեր	5	վասակ
6	անկցին	6	երթիցէ	6	շապիոյ
7	գունոց	7	զառաջին	7	ինձ
8	ներսէս	8	իրաց	8	սակայն
9	դանիելի	9	լինիցի	9	զգաւրսն
10	աշակերտք	10	լիցին	10	պատերազմ

Table 11: English translation keywords and lockwords in the AASC (+) and AAFTC (-).

Keywords(+)	Lockwords	Keywords(-)
katholikos	to carry	Persian
Epip'an (name)	year	example
spirit	creature	day
Xad (name)	quantity	general
spirit	to work	Vasak (name)
to fall	journey	Տօփուհ (name)
group	first	REFL-PRON
Nersēs (name)	PRON	CONJ
Daniel (name)	to be	army
disciples	to be	war

The keywords in the AASC node version of the “Epic Histories” are mostly related to persons, among them Epip’an, Xad, Nersēs, and Daniel. All of these persons belong to a religious sphere. Other keywords in the AASC node corpus also refer to persons by addressing groups of persons such as գունդ (“band” or “group”) and աշակերտք (“disciples”). The “spirit” (հոգևոյս) is very important as well. The top entry is held by a title, namely the “katholikos” (կաթողիկոսն).

Words that are very present in the AAFTC reference corpus and to a lesser extent in the AASC node corpus are the “Persians” and names of “worldly” leaders such as Տօփուհ and Vasak. Other word fields in the keywords list of the reference corpus are closely related to the military, such as զորավարն (“commander”), զգարսն (“armies”), and պատերազմ “war”).

Similarly to the previous analysis of the keywords in the section on “The History of the Armenians,” the keywords in the analysis of the “Epic Histories” also indicate a separation between two spheres that could be described with notions of *holy* and *profane*. In this case, the holy sphere is represented by religious persons, titles, and concepts in the node corpus, whereas the profane sphere is represented by “worldly” leaders, such as kings, military titles, armies, and war. That both spheres are still intertwined is observable in most of the Armenian sources examined in this book.

5.1.3.2 Word Frequency Lists

Types

No *Surb*

The fifteen most common words in the AASC Epic Histories subcorpus (no stop words, no words related to *surb*) are shown in following table and bar plot (see the corresponding figure 18).

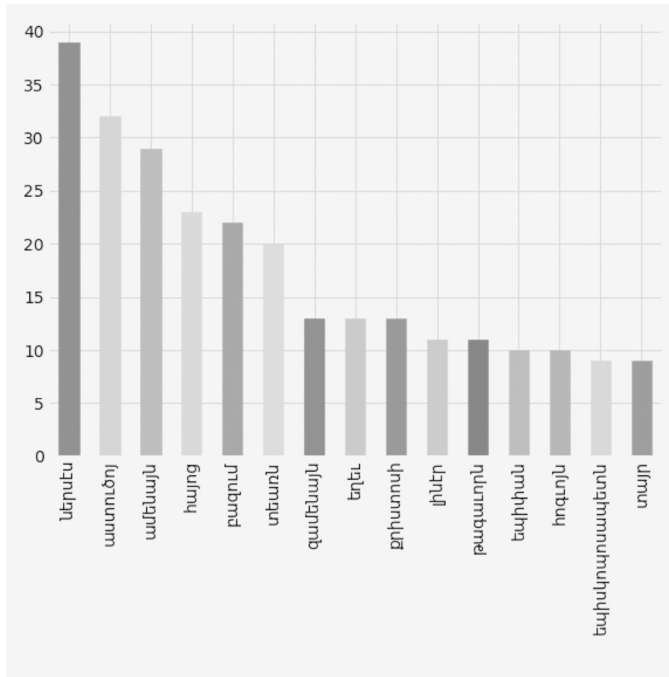
Table 12: Most frequent words (top 15) in AASC Epic Histories.

Type	Translation	Frequency
ներսէս	Nersēs	39
աստուծոյ	God	32
ամենայն	all	29
հայոց	Armenian	23
բազում	many	22
տեառն	lord	20
զամենայն	all	13
եղել	to be	13
քրիստոսի	Christ	13
լինէր	to be	11
թագաւորն	king	11
Եպիփան	Epip'an (name)	10
հոգոյն	spirit	10
Եպիսկոպոսապետն	archbishop	9
տայր	to give	9

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common words are (no *surb*; overall frequency: 264):

- Words related to non-religious “Concepts” (ամենայն, զամենայն, բազում) (AF 67, RF 25%)
- Words related to religious “Concepts” (աստուծոյ, տեառն, քրիստոսի) (AF 65, RF 25%)
- Words related to “Persons” (ներսէս, Եպիփան) (AF 49, RF 18.5%)

Figure 18: Most frequent words (top 15) and their absolute frequencies in AASC Epic Histories.



- Verbs such as “to be present, being” or “to give” (լինէր, եղև, տայր) (AF 33, RF 12.5%)
- Words related to “Places” (հայոց) (AF 23, RF 9%)
- Words related to “Titles” (թագաւորն, Եպիսկոպոսապետն) (AF 20, RF 7.5%)

The word frequency list of the “types” version reveals that individual persons (Nersēs, Epip’an, AF 49) are present but that they are less important than other domains such as religious “concepts” (“God,” “Christ,” “spirit,” AF 75). Furthermore, the region of Armenia (AF 23) is very present, as well as different religious and non-religious titles (“archbishop” and “king,” AF 20).

Even though the term “all” did not appear in the lockwords column of the keyword analysis in this section, it should be remembered that the terms “all” and “many” are frequently found among the keywords of the reference corpus (AAFTC)

of the other Armenian texts. Their impact on the meaning of *surb* thus needs to be interpreted with caution.

Surb

The root *surb/srb-* appears 232 times in the 199 sentences in the AASC Epic Histories. The visualization is displayed in the corresponding figure 19.

- *Surb* is most frequently used as an adjective or a noun (AF 207, RF 89%).
- The use of *surb* as a noun in the sense of *surb*-ness (սրբութիւն) is less common (AF 17, RF 7%).
- The root of *surb* is also used as a verb (սրբել + inflections) and in some other less common expressions (սրբասէր, ամենասուրբ) (AF 8, RF 4%).

Lemmas

The fifteen most common lemmas in the AASC Epic Histories subcorpus (no stop words, no *surb*) are shown in following table and bar plot (see the corresponding figure 20).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common lemmas are (no *surb*; overall frequency: 514):

- Verbs such as “to make” or “to say” (առնել, գալ, տալ, ասել) (AF 135, RF 26%)
- Words related to religious “Concepts” (աստուած, տէր, եկեղեցի) (AF 118, RF 23%)
- Words related to general “Concepts” (ամենայն, բազում, ազգ) (AF 109, RF 21%)
- Words related to “Persons” (ներսէս, մարդ) (AF 97, RF 19%)
- Words related to “Titles” (թագաւոր, տէր) (AF 68, RF 13%)
- Attributes such as “great” or “Armenian” (մեծ, հայ) (AF 53, RF 10%)⁷

7 Note, however, that the attribute “Armenian” is usually applied as a reference to the land of Armenia. Thus, it is listed as a “place” instead of an “attribute” in the synoptic table.

Figure 19: All inflected forms of lexical units based on the root *surb/srb-* in AASC Epic Histories (bar plot).

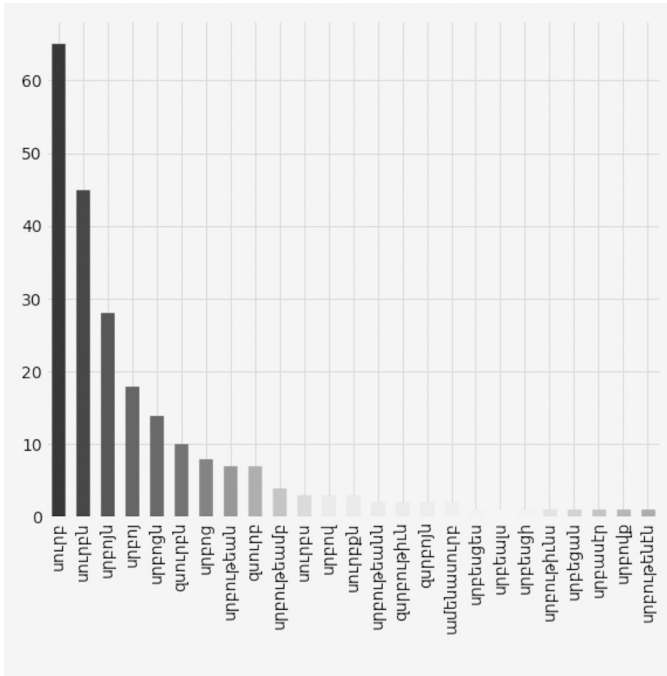


Table 13: All inflected forms of lexical units based on the root *surb/srb-* in AASC Epic Histories.

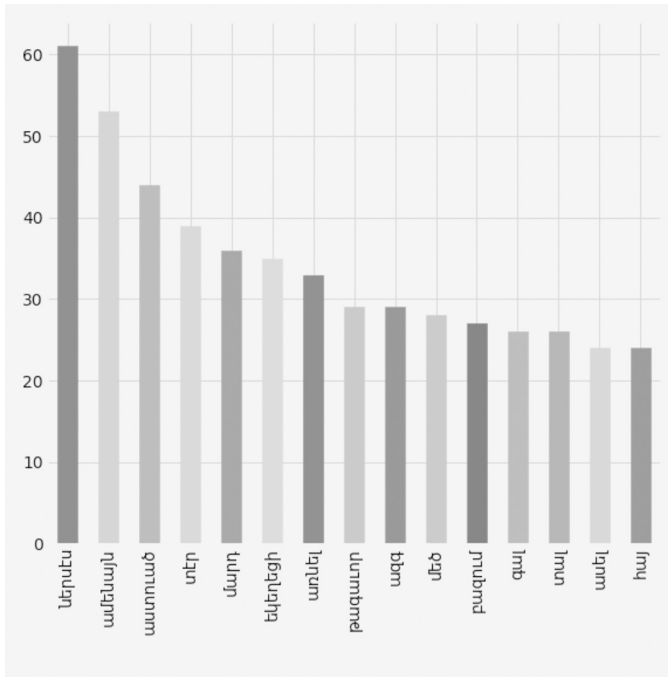
Type	Frequency
սուրբ	65
սուրբն	45
սրբոյն	28
սրբոյ	18
սրբոցն	14
զսուրբն	10
սրբոց	8
սրբութեան	7
զսուրբ	7
սրբութեամբ	4
սուրբս	3

սրբով	3
սուրբքն	3
սրբութեանն	2
զսրբութիւն	2
զսրբոյն	2
ամենասուրբ	2
սրբեսցես	1
սրբեալս	1
սրբեսցի	1
սրբութիւնս	1
սրբեցան	1
սրբասէր	1
սրբովք	1
սրբութենէն	1

Table 14: Most frequent lemmas (top 15) in AASC Epic Histories.

Lemma	Translation	Frequency
ներսէս	Nersēs (name)	61
ամենայն	all	53
աստուած	God	44
տէր	lord	39
մարդ	man	36
եկեղեցի	church	35
առնել	to make	33
թագաւոր	king	29
ազգ	nation	29
մեծ	great	28
բազում	many	27
գալ	to come	26
տալ	to give	26
ասել	to say	24
հայ	Armenian	24

Figure 20: Most frequent lemmas (top 15) in AASC Epic Histories (bar plot).



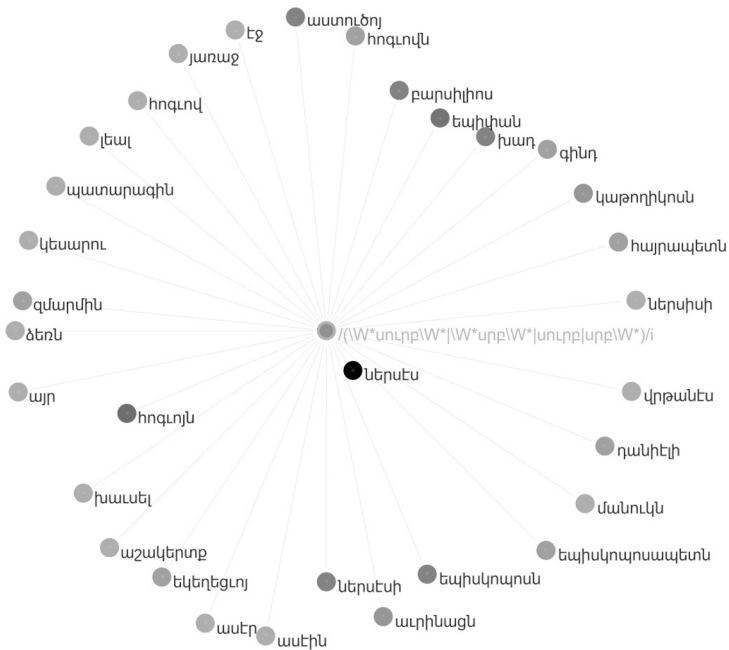
The distribution of the categories in the lemmas analysis of the AASC Epic Histories is nearly even. With the exception of the “attributes” and “titles” domains, each of the categories has a relative frequency between 19-26%. The notable prevalence of the “verb” domain in the “Epic Histories” might be accounted by the overtly descriptive and historical character of this account. Furthermore, the “persons” domain includes only a few entries compared to most of the other Armenian sources. This result reflects the general character of the “Epic Histories” that mentions persons but which has no main protagonists (except *katholikos Nersēs*). Most of the other words that appear in the word frequency lists are also commonly found in the word frequency lists of the other Armenian sources.

5.1.3.3 Collocation Analysis (AAFTC/AASC, Top 15)

Types

The fifteen most common words in the AAFTC Epic Histories subcorpus (no stop words) that appear in the context of words related to *surb/srb-* are shown in the following table (6-LL, L2-R2, C3-NC3; no function words). The visualization is displayed in the corresponding figure 21.

Figure 21: Top 15 types collocation analysis (6-LL, L2-R2, C3-NC3; no function words) *surb/srb-* in AAFTC Epic Histories (graph) (c) LancsBox.



The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 114).

- Words related to “Persons” (ներսէս, եպիփան, բարսիլիոս, շաղիտայի, խադ, գինդ, ներսէսի) (AF 68, RF 60%)
- Words related to religious “Concepts” (հոգւոյն, հոգւովն, եկեղեցւոյ) (AF 20, RF 18%)
- Words related to “Titles” (եպիսկոպոսապետն, հայրապետն, կաթողիկոսն) (AF 13, RF 11%)
- Words related to “Objects” (զմարմին) (AF 4, RF 4%)

Table 15: Top 15 words collocation analysis (6-LL, L2-R2, C3-NC3; no function words) *surb/srb-* in AAFTC Epic Histories.

id	pos	coll	translation	stat	freq_(coll)	freq_(corpus)
1	r	ներսէս	Nersēs (name)	240.69	31	77
2	l	հոգւոյն	spirit	106.68	11	14
3	r	եպիփան	Epip'an (name)	93.43	10	14
4	r	բարսիլիոս	Barsilios (name)	62.28	8	19
5	r	գինդ	Gind (name)	54.97	5	5
6	r	շաղիտայի	Šalitay (name)	43.96	4	4
7	r	խադ	Xad (name)	42.64	5	9
8	r	հոգւովն	spirit	38.97	4	5
9	l	եկեղեցւոյ	church	36.80	5	14
10	r	ներսէսի	Nersēs (name)	33.19	5	19
11	r	կաթողիկոսն	katholikos	32.91	4	8
12	l	յաղագս	for the sake of	32.13	9	150
13	l	զմարմին	body	28.75	4	12
14	r	հայրապետն	patriarch	28.75	4	12
15	r	եպիսկոպոսապետն	archbishop	28.14	5	30

Compared to the previous analysis of the word frequency lists, the collocation analysis of the “types” reveals an emphasize on persons who are all related to the Armenian Church (RF 60%). This result does not contradict the previous observation in the analysis of the word frequency lists where the “persons” domain was relatively irrelevant. It only illustrates that persons are important in the context of *surb*; this does not necessarily disprove that there are only few continuously present persons in the “Epic Histories.” The importance of exclusively church-related titles,

such as “katholikos” or “patriarch,” supports the overall importance of religious persons in the context of *surb* in the “Epic Histories.”

Lemmas

The fifteen most common words in the AASC Epic Histories subcorpus (no stop words) that appear in the context of the lemma *surb* are shown in following table (6-LL, L2-R2, C3-NC3; no function words). The visualization is displayed in the corresponding figure 22.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 141).

- Words related to “Persons” (ներսէս, խաղ, Եպիփան, բարսիլիոս, դանիէլ, գինը, բարսեղ) (AF 83, RF 59%)
- Words related to mostly religious “Concepts” (հոգի, եկեղեցի, յիշատակ, կոյս) (AF 36, RF 26%)
- Words related to “Titles” (կաթողիկոս, Եպիսկոպոս, վկայ, երիցապետ) (AF 22, RF 15%)

The importance of persons in the collocation analysis of the “types” can also be perceived in the collocation analysis of the “lemmas.” “Persons” and “titles” even have an almost similar percentage number (60% to 58% and 11% to 15%). Thus, the aspects mentioned in the previous collocation analysis of the “types” are also valid in the case of the collocation analysis of “lemmas.”

Table 16: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) *surb* in AASC Epic Histories.

id	pos	coll_(lemma)	translation	stat	freq_(coll)	freq_(corpus)
1	r	ներսէս	Nersēs (name)	220.43	43	61
2	l	հոգի	spirit	93.52	17	21
3	r	խաղ	Xad (name)	55.87	10	12
4	r	Եպիփան	Epip’an (name)	49.55	9	11
5	r	կաթողիկոս	katholikos	34.48	7	10
6	r	բարսիլիոս	Barsilios (name)	34.14	6	7
7	r	եկեղեցի	church	31.38	11	35
8	r	դանիէլ	Daniel (name)	28.57	6	9

Table 17a: Synoptic table of the AASC and AAFTC Epic Histories.

id	lemma	kw	wf_t	cf_t	wf_l	cf_l	sum	do- main	semantic_field
1	կաթողիկոս	1	0;0	1;4	0	7	10	title	katholikos_S7
2	Եպիփան	1	1;10	1;10	0	9	20	per- son	Epip'an_Z1m
3	հոգի	2	1;10	2;15	0	17	22	con- cept	spirit_S9
4	խաղ	1	0;0	1;5	0	10	12	per- son	Xad_Z1m
5	գունդ	1	0;0	0;0	0	0	3	con- cept	group_S5+c
6	ներսէս	1	1;39	2;36	61	43	100	per- son	Nersēs_Z1m
7	դանիէլ	1	0;0	0;0	0	6	6	per- son	Daniel_Z1m
8	աշակերտ	1	0;0	0;0	0	0	3	con- cept	disciples_P1/S2mf
9	աստուած	0	1;32	0;0	44	0	39	con- cept	god_S9
10	ամենայն	0	2;42	0;0	53	0	49	con- cept	all_N5.1+
11	հայ	0	1;23	0;0	24	0	23	place	Armenia_Z2
12	բազում	0	1;22	0;0	27	0	24	con- cept	many_N5+
13	տէր	0	1;20	0;0	39	0	30	title	Lord_S7.1+/S2.2m
14	քրիստոս	0	1;13	0;0	0	0	5	con- cept	Christ_Z1m
15	թագաւոր	0	1;11	0;0	29	0	20	title	king_S7.1+/S2.2m
16	Եպիսկոպոսապետ	0	1;9	1;5	0	0	6	title	archbishop_S9/ S2mf

Table 17b: Synoptic table of the AASC and AATC Epic Histories.

17	տալ	0	1;9	0;0	26	0	17	verb	give_A9-
18	բարսիլիոս	0	0;0	1;8	0	6	6	person	Barsilios_Z1m
19	եկեղեցի	0	0;0	1;5	35	11	25	concept	church_S9/H1c
20	գինդ	0	0;0	1;5	0	5	4	person	Gind_Z1m
21	շաղիտայ	0	0;0	1;4	0	0	1	person	šaitay_Z1m
22	մարմին	0	0;0	1;4	0	0	1	object	body_B1
23	հայրապետ	0	0;0	1;4	0	0	1	title	patriarch_S4m
24	մարդ	0	0;0	0;0	36	0	17	person	man_S2.2m
25	առնել	0	0;0	0;0	33	0	15	verb	make_A1.1.1
26	ազգ	0	0;0	0;0	29	0	13	concept	nation_G1.1c
27	մեծ	0	0;0	0;0	28	0	13	attribute	great_A5.1+
28	գալ	0	0;0	0;0	26	0	12	verb	come_M1
29	ասել	0	0;0	0;0	24	0	11	verb	say_Q2.1
30	եպիսկոպոս	0	0;0	0;0	0	8	2	title	bishop_S9/S7.1+/ S2mf
31	բարսեղ	0	0;0	0;0	0	4	0	person	Barset_Z1m
32	վկայ	0	0;0	0;0	0	4	0	title	martyr_S9/S2mf
33	յիշատակ	0	0;0	0;0	0	5	1	concept	remembrance_X2.2+
34	երիցապետ	0	0;0	0;0	0	3	0	title	archpriest_S9/ S7.1+/ S2.2m
35	կոյս	0	0;0	0;0	0	3	0	concept	virgin_S3.2

5.1.3.4 Synoptic Table – Sentences Selection and Discussion (Microanalysis)

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses as well as their corresponding values and annotation has been created.

As explained in Chapter 3, the selection of example sentences for the micro-analysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *surb* appears in the neighborhood of the selected lemma, the sentence that offers the most potential for the analysis is manually selected

The following domains appear in the “Epic Histories” (ordered by their summarized “sum” values):

Table 18: Domains in the AAFTC/AASC Epic Histories in a descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	184
Person	166
Title	69
Verb	55
Place	23
Attribute	13
Object	1

Concept

With a “sum” value of 45, the lemma ամենայն (“all”) is the most important word in the “concept” domain. However, as this lemma is very unspecific and regularly appears among the keywords of the reference corpus AAFTC, the second most important lemma աստուած (“god”) is selected instead.

(...) յիշել զսոսա առ սուրբ սեղանն Աստուծոյ ի ժամ պատարագին, յորժամ զանուանս սրբոցն կարգիցեն, եւ ապա զկնի նոցա զնոցայն. եւ մնացելոցն անկելոցն գութ եւ խնամ կալցին: (3,11,15)

(...) to commemorate them before God’s holy altar at that point of the liturgy when the names of the saints are enumerated, and after them. And [he ordered] that compassion and care be given to the survivors of the fallen.

Discussion

This sentence and its use of *աստուած* (“god”) in the context of *surb* is analogous to the example discussed in the microanalysis of the “History of the Armenians.” In both cases, the word “god” is used in the genitive case as an attribute to qualify persons or, as in this instance, objects.

The presence of an object in this example, namely the “holy altar,” is a significant addition to the general (macro-)analysis. According to the results of the macroanalysis, the “object” domain only plays a minor role in the “Epic Histories.” From the macro perspective, this assumption is justified if comparing the “object” domain to other domains, such as the “concept” or “person” domains. Nonetheless, the fact that *մարմին* (“body”) is the only word that occurs in the synoptic table as part of the “object” domain does not necessarily imply that it is the only object in the text that is related to *surb*. What it does indicate is that the “body” is the only object that has passed the various filtering and selection processes and thus seems to be more representative than, for instance, the altar (*սեղան*).

Another important topic in this sentence – one which was already encountered in previous parts of the macroanalysis – are the saints, the cult of the saints and their liturgical (*պատարագ*) embedding are introduced; this is expressed by their commemoration (*յիշել*) by enumerating their names (*զանուանսս սրբոցն կարգիցեն*). Many words in this sentence also arise in other parts of the macroanalysis of the “Epic Histories” and of other Armenian texts in this book.

Person

With a “sum” value of 100, the name *Ներսէս* (Nersēs) is the most conspicuous word in the “person” domain. The name *Ներսէս* (Nersēs) is not only the most frequent word in the “person” domain but also the most important word in the synoptic table.

Երկիւղած ի Տեառնէ Աստուծոյ, յոյժ ասանդապան պատուիրանաց նորա. մարդասէր, սուրբ զգաստ, սաստիկ իմաստուն եւ առանց ակնառութեան. իրաւանց իրաւարար, ցածուն քաղցր խնարհ աղքատասէր, ի սրբութիւնս ամուսնութեանցն արիճաւոր, եւ կատարեալ սիրովն Աստուծոյ: (4,3,10)

He [Nersēs] was fearful of the Lord God and a strict keeper of His Commandments, benevolent, holy, prudent, most wise and impartial, just in his judgments, humble, gentle, modest, a lover of the poor, observant of the sanctity of marriage, and perfect in the love of God.

Discussion

As stated in the examination of the word frequency lists in this section, Nersēs is the most prominent person in the “Epic Histories.” According to Nina Garsoïan’s

thorough prosopography included in the appendix of her translation of the “Epic Histories,” Nersēs was one of the most notable patriarchs in Armenia in the fourth century CE (approximately from 353–373 CE) (P’awstos and Garsoïan 1989, 8:395). The author even describes the entire account of the “Epic Histories” as “an uncritical panegyric [of its spiritual hero Nersēs], which provided the basis for the still more elaborately hagiographical vita” (P’awstos and Garsoïan 1989, 8:396). This, in turn, underlines again the close relationship between *surb* and outstanding religious persons in early Armenian Church history.

Furthermore, this sentence is another example of the series of attributes that also enhances the panegyric character of the account.⁸ In this example sentence, the (positive) attributes can be classified in the following way:

- general attributes that underline a positive personality (“benevolent”/ մարդասէր, “wise”/իմաստուն, “impartial”/ակնառութեան, “just”/ իրաւարար, “humble”/ցածուն, “gentle”/քաղցր, “modest”/խոնարհ, “lover of the poor”/աղքատասէր)
- attributes related to purity or religion (“modest”/զգաստ, “observant of the sanctity [սրբութիւնս] of marriage”/ի սրբութիւնս ամուսնութեանցն արիւնատր, “perfect in the love of God”/կատարեալ սիրովն Աստուծոյ)

The reason why *surb* is already counted among the religious attributes in this list is that – as previously observed (Chapter 5 “History of the Armenians”) – *surb* never refers to non-religious persons. Terms related to *surb* are used twice in this chain of attributes. Firstly, *surb* is used as an attribute (սուրբ, *surb*) preceding “chaste” (զգաստ). Secondly, *surb* is used as a noun (սրբութիւնս, *surb*-ness) in “sanctity of marriage.”⁹ The proximity of words related to sexual purity also confers a strong connotation of purity to the use of *surb* in this context. Therefore, this passage further demonstrates that *surb* often characterizes persons and sometimes acquires a meaning of sexual purity.

Title

With a “sum” value of 30, the lemma տէր (“lord”) is the most noticeable word in the “title” domain. Similarly to the previous section, this title is mostly used as a synonym for god and will thus not be discussed in this section. The second most frequently used lemma, “king” (թագաւոր), will not be addressed either. With a value of 20, the use of “king” in the “History of the Armenians” is identical to its

8 For another great example, see Epic Histories 4.7.6.

9 Note that the use of nouns in the genitive case before another noun has the function of an adjective and thus represents an idiosyncratic use of nouns in Armenian. Consequently, the meaning is not only “sanctity of marriage” but also “holy marriage.”

application in the “Epic Histories.” Instead, I have chosen to discuss the next word on the list, namely the “katholikos” (կաթողիկոս) with a “sum” value of 10.

զնոսա զամենեսին կազմեցին եւ արձակեցին բազում ընծայիւք եւ մեծամեծ պատարագաւք եւ հաւատարիմ հրովարտակաւք առ կաթողիկոսաց կաթողիկոսն Եւսեբիոս յերկիրն Գամրաց, եւ ի մայր քաղաքացն նոցա ի Կեսարիա, զի ձեռնադրեսցեն անդ զսուրբն Ներսէս ի կաթողիկոսութիւն Հայոց մեծաց: (4.4,5)

[All] of them got ready and were sent forth with many presents and the greatest gifts and authenticated letters-patent to the kat'olikos of kat'olikoi Ewsebios, to the land of Gamirk' and to its metropolis Caesarea, so that they might have St. Nersēs consecrated there as kat'olikos of Greater Armenia.

Discussion

The *katholikos* is a religious title that designates the highest ecclesiastical position in the Armenian sources examined in this book. The title is not only granted to the leaders of the Armenian Church but also to patriarchs from other regions, such as Constantinople. Besides the *katholikos*, there are many other religious titles, such as “bishop” (եպիսկոպոս), “patriarch” (հայրապետ), or “archpriest” (երիցապետ), which frequently appear in the context of *surb*. Many of these terms, such as “katholikos” or “bishop” (*yepiskopos*, եպիսկոպոս), are Greek loan words thereby pointing at institutional ties between the emerging Armenian Church and the Greek Church(es). Regarding the title “katholikos,” Nina Garsoïan argues that its use in the “Epic Histories” is anachronistic and that the church officials of the Armenian Church in the fourth century CE were more likely called patriarchs than *katholikai* (P’awstos and Garsoïan 1989, 8:537).

The application of the term *surb* to qualify official church titles underlines the religious and social standing of its bearer. Thus, the presence of titles can be regarded as another attribute highlighting the exceptional position of those who are also called *surb* in a religious and social hierarchy. Even though the title of a “king” (թագաւոր) also occurs in the macroanalysis, the kings are never directly called *surb*. Instead, kings often appear alongside persons who are called *surb*, thereby underlining the entanglement between the religious and non-religious spheres in the Armenian texts.

Verb

With a “sum” value of 17, the lemma տալ (“to give”) is the most pervasive word in the “verb” domain. This verb as well as other verbs, such as առնել (“to do”) and գալ (“to go”), are very generic and have already been partially discussed in the previous parts. Therefore, I will turn my attention to the rare use of *surb* as a verb. The verbal form is so scarce that it never appears in the synoptic tables.

Եւ կոչեցաւ յապարանս իւր ի Խախ ասանի յԵկեղեաց գաւառի. արար նմա ընթրիս, աղաչեաց զայրն Աստուծոյ բազմել նմա յարքունական գահոյսն, եւ իբրեւ այնու թէ սրբեսցի նա ի չարեաց անտի իւրոց, եւ ինքն այնուհետեւ յայնմ հետէ յապաշխարութիւն մտցէ: (5,24,5)

He invited him to his palace in the town of Xax in the district of Ekeleac, had a banquet prepared, and invited the man of God to recline on the royal throne, as though he might cleanse himself thereby from his evil deeds, and thereupon come repentance.

Discussion

In this passage, the verb սրբեսցի (third person aorist subjunctive from the infinitive սրբել, *srbel*) is closely connected to cleansing “evil deeds” (ի չարեաց անտի) and “repentance” (ապաշխարութիւն). Overall, the use of *surb* as a verb is a reliable marker of its relationship to the field of purity since it is often applied in the sense of purification. In this sentence, the syntactic and semantic embedding of *srbel* confirms the hypothesis that *surb*, in this context, should be associated with “cleansing, freeing, taking away.”

This tendency is also noticeable in many other instances of *surb* as a verb. However, as *surb* rarely appears as a verb, these examples seldom occur among the results of the macroanalysis. More examples of the application of *surb* as a verb can be found in the microanalysis of the “object” domain below and the examination of the Gospels.

Place

Compared to the “History of the Armenians,” the “place” domain is rather negligible in the “Epic Histories.” Հայ (“Armenia” or “Armenian”) is the only lemma that appears with a “sum” value of 23 in the synoptic table under the “place” domain.

Եւ նորա սիրով ընկալեալ զնոսա եւ բազում մեծարանաւք, եւ ըստ կանոնաց արինակի ժողովէր առ ինքն մեծ եպիսկոպոսապետն Եւսեբիոս զբազմութիւն եպիսկոպոսացն սրբոցն ըստ առաքելական արինացն, զի ձեռնադրեսցեն զսուրբն Ներսէս լինել եպիսկոպոսապետ Հայոց մեծաց: (4,4,8)

And he received them lovingly and with great splendor. [Then] following the canonical regulations, the great chief-bishop Eusebios assembled many holy bishops in accordance with the apostolic canons, so that they might ordain St. Nersēs as chief-bishop of Greater Armenia.

Discussion

Besides the importance of persons and titles in this sentence, one may also observe that the geographical information of “(Greater) Armenia” is used as an attribute for the title “chief-bishop,” namely by defining its sphere of influence. The term Հայոց մեծաց (“Greater Armenia”) should be interpreted as an indication of the separation process between the Armenian Church and other Churches. It further points to the function of accounts such as the “Epic Histories” in identity-building processes. This idea is also conveyed by terms such as ազգ (“nation”) in the synoptic table. The fact that these terms commonly occur in the vicinity of *surb* underlines the significance of religious events and actors for these political processes. They thus demonstrate once more the close relationship between religion and politics.

Attribute

With a “sum” value of 13, the term մեծ (“great”) is the most frequent (and only) word in the “attribute” domain. The term մեծ (“great”) is either used in the context of “Greater Armenia” (see the microanalysis of the “place” domain) or as another attribute in the series of attributes (see the previous examinations). For this reason, the term մեծ (“great”) will not be discussed with a microanalysis.

Object

The “object” domain is of lesser importance in the synoptic table of the “Epic Histories.”¹⁰ Մարմին (“body,” “flesh”) is the only lemma that appears in this category in the synoptic table with a “sum” value of only 1.

Եւ ոյք սիրեն զնա, սէրն իրեանց կերպարանէ զնա ի սիրտս իրեանց, եւ երկիր պագանեն նմա հոգովն ճշմարտութեամբ. սրբեալ զսիրտս իրեանց եւ զմարմինս իրեանց, առնեն տաճար Սուրբ Հոգւոյն. (4.5,42)

And those who love Him form Him in their hearts through their love, and they prostrate themselves before Him in spirit and truth - purifying their hearts and bodies they make a temple for the Holy Spirit.

Discussion

This passage reveals that the term *surb* is usually not used as a direct attribute for “body.” However, the application of *surb* in this sentence is a telling example of the close relationship between *surb* (particularly as a verb) and purity (see also the “verb” domain in this section). When looking at the syntax and the other words in this

10 Which does not mean that *surb* does not appear in the context of objects as has previously been demonstrated. It only proves that the “object” domain is, according to the selection criteria set up in the macroanalysis, less relevant from a specific point of view.

part, it becomes evident that the use of *surb* as a verb effectively “does something” by creating a “temple for the Holy Spirit” (առնեն տաճար Սուրբ Հոգւոյն)¹¹ through the “cleansing” (սրբեալ) of both the “body” (զմարմինս) and the “mind” (զսիրտս). The connection to physical and to spiritual purity in this example is implicit from the context. Acts of spatial cleansing or purification often accompany the creation of temples or churches both in Christianity and in other religions. Furthermore, the metaphorical idea of “cleansing/purifying the body and the mind” (in order to build a temple for god) is also present in the Gospels (1 Cor 6:18-19).

5.1.3.5. Final Discussion

The microanalysis of the “Epic Histories” has yielded the following results.

- The use of *surb* in the “Epic Histories” is closely connected to persons. The persons who are called *surb* are all related to religion or religious institutions. Furthermore, the persons called *surb* are extraordinary to some extent, either due to their (previous) deeds or because they belong to the highest ranks in church hierarchies.
- The use of *surb* in the context of outstanding religious persons is characterized by long attribute chains that often include general (positive) attributes and attributes related to religion or purity. These evoke panegyrics. This observation was also made in the context of the “History of the Armenians.”
- The term *surb* and particularly its use as a verb reveal connections to the field of purity. This relation has already been mentioned in the microanalysis of the “History of the Armenians,” where *surb* was often used in the sense of sexual purity. The results of the analysis of the “Epic Histories” show that *surb* and its connection to purity are not restricted to the field of sexual purity. Instead, *surb* can be applied in the broader sense of physical and spiritual purity.
- The microanalysis yielded several instances of the combination of *surb* and objects. This observation should be taken as an important reminder that there are frequent co-occurrences of *surb* in the proximity of objects in the material, although objects seldom appear in the synoptic tables due to the previously mentioned reasons.

11 The “Holy Spirit” is another ubiquitous example of the use of *surb* together with a religious concept (“spirit”) that regularly appears among the top entries in the synoptic tables of all the Armenian sources in this book. It will be discussed at a later point in this book.

5.1.4 AAFTC/AASC Koriwn

5.1.4.1 Keyword Analysis (“Types” Only)

The tables below present the ten most important keywords and lockwords. The keywords of the AASC node corpus are in the left column and those of the reference corpus AAFTC in the right column. The lockwords are in the middle column (words that appear to a similar extent in both corpora). The first table (see the corresponding figure 23) includes the original Armenian terms. The second table contains the English translations of the Armenian terms. All words related to *surb* have been removed because they obviously are the most widespread keywords in the node corpus AASC. The table with the English translations only lists the English lemmas of the inflected Armenian words. Consequently, it does not reflect the morphological variance of the Armenian words in the first keyword table.¹²

The node corpus’ (AASC) keywords table includes mainly religious terms, such as “pious” (բարեպաշտաւն), “church” (եկեղեցի), and “shrine” (խորան). The disciples (աշակերտ) also frequently appear in the node corpus as well as terms related to “death” (վախճան) and the verb “to intercede” (բարեխօսեմ). The words “death,” “shrine,” and “intercession” bespeak the importance of martyrs and thus once more highlight the relationship between *surb* and persons.

The word field around “day” (օր) in the node corpus is of no consequence for the interpretation of *surb*. The word “year” (ամ) has already been removed in advance (see the first footnote in this section including the *regular expression*). Both are mainly used for dating purposes. The keywords in the reference corpus are rather diverse and include many stop words which makes it difficult to trace any patterns.

12 In the case of “The Life of Maštoc,” other keywords have been removed as well. For instance, terms related to “year” (ամ) were very present in the original keyword table. As these terms were only used for dating and thus had nothing to do with the semantic field of *surb*, they have been removed from the keyword table. This is the *regular expressions* statement used for removing several words from the keyword table:
/(\W*սուրբ \W* / \W*սրբ \W* / ամք / ամաց / ւլնր / սուրբ / սրբ \W*)/i

Figure 23: Keywords and lockwords in the AASC (+) and AAFTC (-).

1/645	Keywords +	1/30	Lockwords	1/2034	Keywords -
1	բարեպաշտաւն	1	առաջնոյն	1	շնորհացն
2	խորանին	2	բայց	2	առն
3	վախճանի	3	բարեացն	3	արիիակ
4	անցելոց	4	դառնային	4	ժամանակս
5	աշակերտ	5	դասս	5	որով
6	աշակերտաւք	6	եպիսկոպոսի	6	աստուածեղէն
7	աւուրց	7	երթեալ	7	արկաներ
8	բարբառոյն	8	իջաներ	8	ժամանակի
9	բարեխաւսել	9	իրացն	9	ձեռն
10	եկեղեցոյն	10	իւրոց	10	մանաւանդ

Table 19: English translation keywords and lockwords in the AASC (+) and AAFTC (-).

Keywords(+)	Lockwords	Keywords(-)
pious	first	grace
shrine	but	to_make
death	good	example
to pass	to turn	time
disciple	order	PRON
disciples	bishop	divine
day	to march	to cast
language	to descend	time
to intercede	PRON	hand
church	PRON	rather

5.1.4.2 Word Frequency Lists

Types

No Surb

The fifteen most common words in the AASC Koriwn subcorpus (no stop words, no words related to *surb*) are shown in following table and bar plot (see the corresponding figure 24).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common words are (no *sub*; overall frequency: 103):

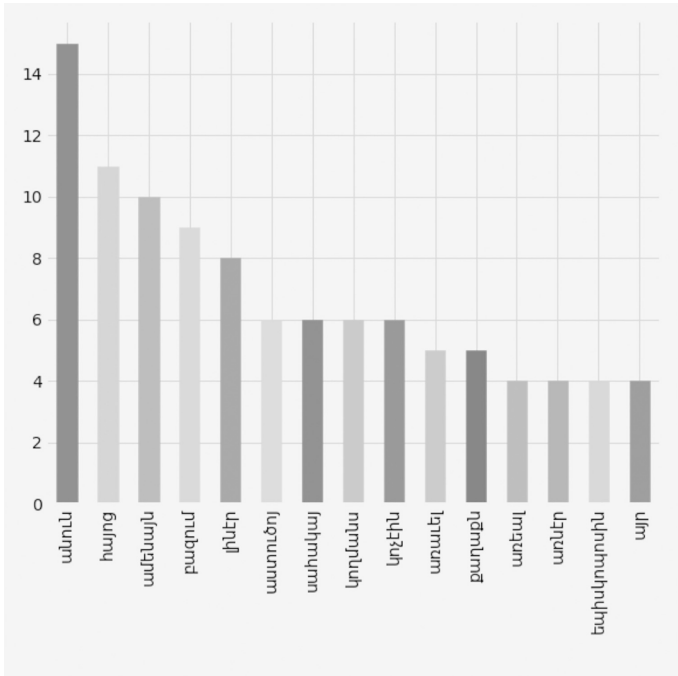
Table 20: Most frequent words (top 15) in AASC Koriwn.

Type	Translation	Frequency
անուն	name	15
հայոց	Armenia(n)	11
ամենայն	all	10
բազում	many	9
լինէր	to be	8
աստուծոյ	god	6
սահակայ	Sahak (name)	6
կողմանս	region	6
կոչէին	to call	6
առաւել	more	5
քաղաքն	city	5
առեալ	to take	4
առնէր	to make	4
եպիսկոպոսին	bishop	4
այր	man	4

- Words related to non-religious “Concepts” (ամենայն, բազում, առաւել, անուն) (AF 39, RF 39%)
- Words related to “Places” (հայոց, քաղաքն, կողմանս) (AF 22, RF 22%)
- Verbs such as “to call” or “to make” (լինէր, առեալ, առնէր, կոչէին) (AF 22, RF 22%)
- Words related to “Persons” (սահակայ, այր) (AF 10, RF 10%)
- Words related to religious “Concepts” (աստուծոյ) (AF 6, RF 6%)
- Words related to “Titles” (եպիսկոպոսին) (AF 4, RF 4%)

Compared to the analyses conducted so far, words related to places are more frequently found in the word frequency list of the AASC Koriwn, whereas categories such as “persons” and religious “concepts” are less important. Even though words

Figure 24: Most frequent words (top 15) and their absolute frequencies in AASC Koriwn.



related to “all” are very present as well, their influence on the overall meaning of *surb* has already been questioned.

Another interesting observation is that “god” (աստուծոյ) is the only religious concept that occurs in the word frequency list. This might suggest that the account by Koriwn is less directly concerned with religious topics than most of the other Armenian accounts so far.

Surb

The root *surb/srb-* appears 56 times in the 49 sentences in the AASC Koriwn. For the corresponding visualization, see figure 25.

- Most common is the use of *surb* as an adjective or noun (AF 53, RF 95%).
- The root of *surb* is also used as a verb (սրբել + inflections) and in some other less common expressions (սրբասեր, սրբամատոյց) (AF 2, RF 3.5%).

- The use of *surb* as a noun such as in *surb-ness* (սրբութիւն) is less common (AF 1, RF 1.5%).

Figure 25: All inflected forms of lexical units based on the root *surb/srb-* in AASC Koriwn (bar plot).

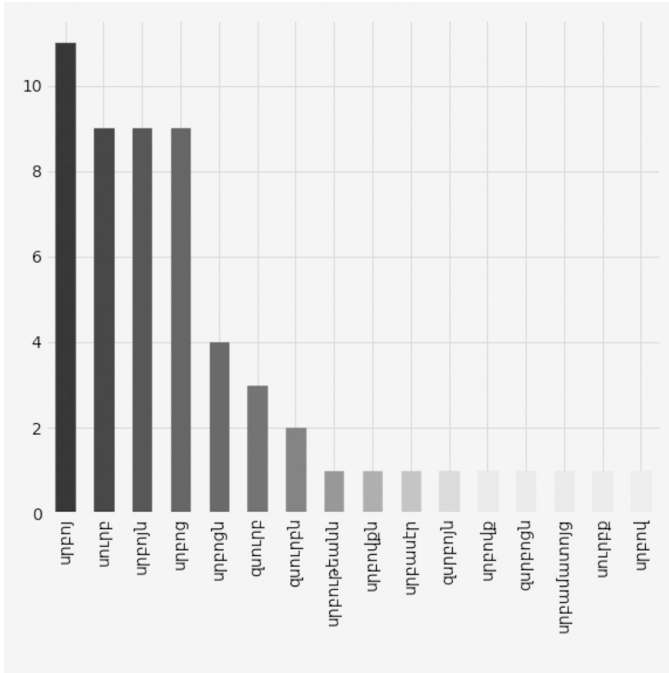


Table 21: All inflected forms of lexical units based on the root *surb/srb-* in AASC Koriwn.

Type	Frequency
սրբոյ	11
սրբ	9
սրբոյն	9
սրբոց	9
սրբոցն	4
զսրբ	3
զսրբն	2

սրբութանն	1
սրբովքն	1
սրբասէր	1
զարբոյն	1
սրբովք	1
զարբոցն	1
սրբամատոյց	1
սուրբք	1
սրբով	1

Table 22: Most frequent lemmas (top 15) in AASC Koriwn.

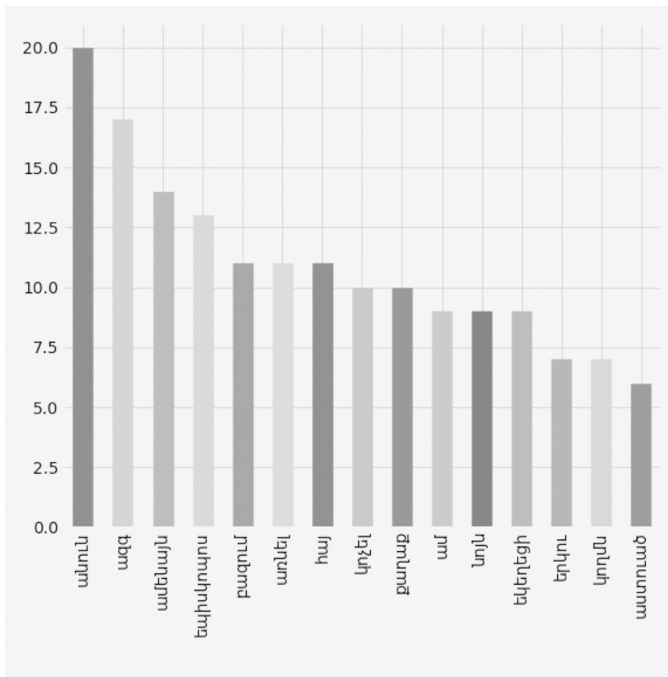
Lemma	Translation	Frequency
անուն	name	20
ազգ	nation	17
ամենայն	all	14
եպիսկոպոս	bishop	13
բազում	many	11
առնել	to make	11
հայ	Armenia(n)	11
կոչել	to call	10
քաղաք	city	10
ամ	year	9
նոյն	PRON	9
եկեղեցի	church	9
երկու	two	7
կողմն	region	7
աստուած	god	6

Lemmas

The fifteen most common lemmas in the AASC Koriwn subcorpus (no stop words, no *surb*) are shown in following table and bar plot (see corresponding figure 26).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common lemmas are (no *surb*; overall frequency: 164):

Figure 26: Most frequent lemmas (top 15) in AASC Koriwn (bar plot).



- Words related to non-religious “Concepts” (ամենայն, բազում, անուն, երկու, նոյն) (AF 61, RF 37%)
- Words related to “Places” (հայ, ազգ, քաղաք, կողմն) (AF 45, RF 27%)
- Verbs such as “to make” and “to call” (առնել, կոչել) (AF 21, RF 13%)
- Words related to religious “Concepts” (աստուած, եկեղեցի) (AF 15, RF 9%)
- Words related to “Titles” (եպիսկոպոս) (AF 13, RF 8%)
- Words related to “Time” (ամ) (AF 9, RF 6%)

Even though non-religious concepts are much more present in the “lemmas” frequency list compared to the “types” frequency list, most of them, such as “two” (երկու) or “himself” (նոյն), are of no relevance for the discussion of *surb*.

The “place” category is more widespread in the “lemmas” frequency list, whereas otherwise important categories such as “persons” or religious “concepts” are of lesser importance in the AASC Koriwn. These tendencies compare to the “types” word frequency list above.

5.1.4.3 Collocation Analysis (AAFTC/AASC, Top 15)

Types

The fifteen most common words in the AAFTC Koriwn subcorpus (no stop words) that appear in the context of words related to *surb/srb-* are shown in the following table (6-LL, L2-R2, C2-NC2; no function words). The visualization can be found in the corresponding figure 27.

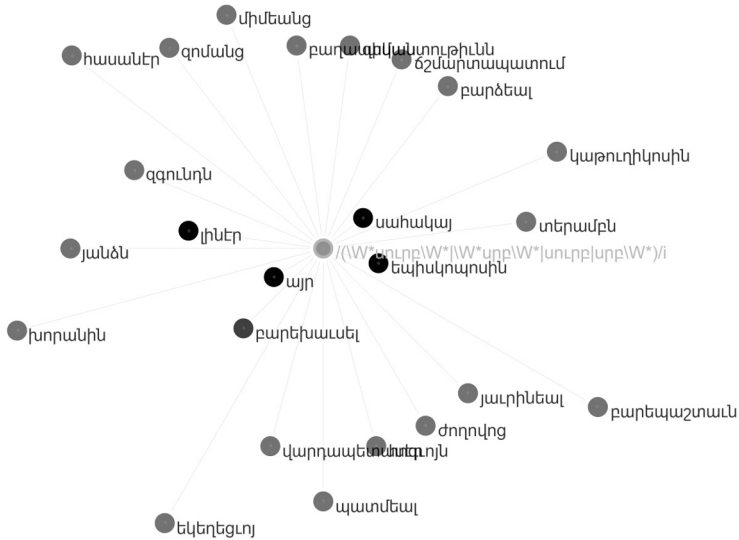
The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 32).

Table 23: Top 15 words collocation analysis (6-LL, L2-R2, C2-NC2, no function words) *surb/srb-* in AAFTC Koriwn.

id	pos	coll	translation	stat	freq_(coll)	freq_(corpus)
1	r	Եպիսկոպոսին	bishop	32.66	4	5
2	l	այր	man	23.35	4	11
3	m	ժողովոց	assemblage	21.52	2	1
4	m	բարեխաւսել	to intercede	18.75	2	2
5	r	սահակայ	Sahak (name)	16.83	3	9
6	r	բարեպաշտաւն	pious	14.95	2	3
7	l	խորանին	region	14.95	2	3
8	r	կաթողիկոսին	katholikos	14.95	2	3
9	m	բարձեալ	to remove	13.24	2	4
10	m	Եկեղեցոյ	church	12.08	2	5
11	l	լինէր	to be	11.00	3	22
12	l	առնէր	to make	9.34	2	9
13	l	հասանէր	to arrive	9.34	2	9

- Verbs such as “to intercede” and “to make” (բարեխաւսել, բարձեալ, լինէր, առնէր, հասանէր) (AF 11, RF 34%)
- Words related to “Persons” (այր, ժողովոց, սահակայ) (AF 9, RF 28%)
- Word related to “Titles” (Եպիսկոպոսին, կաթողիկոսին) (AF 6, RF 19%)
- Words related to religious “Concepts” (բարեպաշտաւն, Եկեղեցոյ) (AF 4, RF 13%)
- Words related to “Places” (խորանին) (AF 2, RF 6%)

Figure 27: Top 15 types collocation analysis (6-LL, L2-R2, C2-NC2; no function words) *surb/srb-* in AAFTC Koriwn (graph) (c) LancsBox.



Interestingly, the “places” domain which was very present in the word frequency lists is less relevant in the collocation analysis. Instead, the “persons” and “titles” domains are much more present compared to the previous examinations of the AASC Koriwn. This indicates that places, such as cities or regions, seem to play an important role in the wider context of *surb*; however, the closer vicinity of *surb* is dominated by “persons,” “titles,” and verbs related to personal actions, such as “to make” or “to intercede.” Therefore, it indicates once more the importance of saints and revises the results of the word frequency lists, at least to a certain extent.

Lemmas

The fifteen most common words in the AASC Koriwn subcorpus (no stop words) that appear in the context of the lemma *surb* are shown in following table and bar plot (6-LL, L2-R2, C2-NC2; no function words). For the visualization see the corresponding figure 28.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic

tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 46).

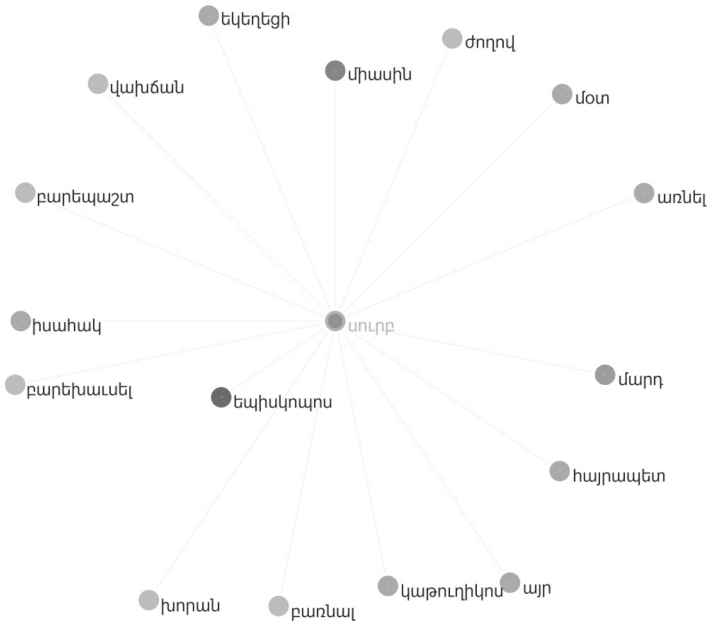
- Words related to “Titles” (եպիսկոպոս, կաթողիկոս, հայրապետ) (AF 15, RF 33%)
- Words related to “Persons” (մարդ, ժողով, այր, սահակ) (AF 12, RF 26%)
- Verbs such as “to arrive” or “to intercede” (բառնալ, հասանել, բարեխաւսել, առնել) (AF 10, RF 22%)
- Word related to (mainly) religious “Concepts” (եկեղեցի, բարեպաշտ, վախճան) (AF 7, RF 15%)
- Words related to “Places” (խորան) (AF 2, RF 4%)

Table 24: Top 15 lemmas collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb in AASC Koriwn.

id	pos	coll	translation	stat	freq_(coll)	freq_(corpus)
1	r	եպիսկոպոս	bishop	43.11	9	13
2	l	մարդ	man	20.85	4	5
3	r	կաթողիկոս	katholikos	19.27	3	3
4	l	հայրապետ	archbishop	19.27	3	3
5	m	բառնալ	to remove	15.59	2	1
6	m	ժողով	assemblage	12.81	2	2
7	l	այր	man	12.70	3	5
8	r	սահակ	Sahak (name)	11.20	3	6
9	l	հասանել	to arrive	11.19	3	6
10	m	բարեխաւսել	to intercede	9.07	2	3
11	l	եկեղեցի	church	8.3	3	9
12	r	բարեպաշտ	pious	7.43	2	4
13	l	խորան	dwelling	7.43	2	4
14	l	վախճան	death	7.43	2	4
15	l	առնել	make	7.02	3	11

The observations from the “types” collocation analysis are also valid in the case of the “lemmas” collocation analysis. On the one hand, they underline the fact that the “person” and “titles” domains are much more present in the collocation analysis than in the word frequency lists. On the other hand, the “places” domain is

Figure 28: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) *surb* in AASC Koriwn (graph) (c) LancsBox.



less significant. Yet, there is only one name in the “person” domain, namely Sahak, whereas the other terms mostly refer to either groups (ժողով) or abstract signifiers such as “man” (մարդ, այր).

Contrary to the initial presumptions resulting from the word frequency lists, religious “concepts” are also much more present in the collocation analysis, thereby indicating that the application of *surb* in “The Life of Maštoc” also has a strong religious connotation. However, the hypothesis that the overall account is less religiously imbued than the other Armenian sources is still valid.

5.1.4.4 Synoptic Table – Sentences Selection and Discussion (Microanalysis)

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses as well as their corresponding values and annotation has been created.

As explained in Chapter 3, the selection of example sentences for the microanalysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *surb* appears in the neighborhood of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

As explained in Chapter 3, the selection of example sentences for the microanalysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *surb* appears in the neighborhood of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

Table 25: Synoptic table of the AASC and AAFTC Koriwn.

id	lemma	kw	wf_t	cf_t	wf_l	cf_l	sum	do- main	semantic_field
1	բարե- պաշտ	1	0;0	1;2	0	2	46	con- cept	pious_S9
2	խորան	1	0;0	0;2	0	2	43	place	shrine_S9
3	վախճան	1	0;0	0;0	0	2	37	con- cept	death_L1-
4	անցել	1	0;0	1;5	0	0	50	verb	pass_M1
5	աշակերտ	2	0;0	0;0	0	0	0	con- cept	disciples_P1/S2mf
6	օր	1	0;0	0;0	0	0	31	time	day_T1.3
7	բարբառ	1	0;0	0;0	0	0	31	con- cept	speech_Q2.1
8	եկեղեցի	1	0;0	1;2	9	3	78	con- cept	church_S9/H1c

9	բարեխօսել	1	0;0	1;2	0	2	71	verb	intercede_S1.1.3+
10	ամենայն	0	1;10	0;0	14	0	78	concept	all_N5.1+
11	հայ	0	1;11	0;0	11	0	71	place	Armenia_Z2
12	բազում	0	1;9	0;0	11	0	65	concept	many_N5+
13	կողմն	0	1;6	0;0	7	0	43	place	region_M7
14	քաղաք	0	1;5	0;0	10	0	50	place	city_M7c
15	առնել	0	2;8	1;2	11	2	81	verb	make_A1.1.1
16	Եպիսկոպոս	0	1;4	1;4	13	9	100	title	bishop_S9/S7.1+/S2mf
17	այր	0	1;4	1;4	0	3	40	person	man_S2.2m
18	կոչել	0	1;6	0;0	10	0	53	verb	call_Q2.2
19	անուն	0	1;5	0;0	20	0	81	concept	name_A1
20	սահակ	0	1;6	1;3	0	3	43	person	Sahak_Z1m
21	աստուած	0	1;6	0;0	6	0	40	concept	god_S9
22	ժողով	0	0;0	1;2	0	2	15	person	assemblage_S5+
23	կաթողիկոս	0	0;0	1;2	0	3	18	title	katholikos_S7
24	բառնալ	0	0;0	1;2	0	2	15	verb	remove_M2
25	հասանել	0	0;0	1;2	0	3	18	verb	arrive_M1
26	ազգ	0	0;0	0;0	17	0	53	concept	free_S6-
27	ամ	0	0;0	0;0	9	0	28	time	year_T1.3
28	նոյն	0	0;0	0;0	9	0	28	concept	pron_??
29	երկու	0	0;0	0;0	7	0	21	concept	two_N1
30	մարդ	0	0;0	0;0	0	4	12	person	man_S2.2m
31	հայրապետ	0	0;0	0;0	0	3	9	title	archbishop_S9/S2mf

The following domains appear in the “The Life of Maštoc” (ordered by their summarized “sum” values):

Table 26: Domains in the AAFTC/AASC Koriwn in a descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	558
Person	268
Title	207
Verb	127
Place	110
Object	59

Concept

With a “sum” value of 81, the lemma անուն (“name”) is the most important word in the “concept” domain.

Եւ զնորին աշակերտ Թադիկ անուն, զայր զգաստ եւ բարեպաշտան, հանդերձ եղբարքք, երանութեան հասելոյ սպասաւոր սրբոյն ի փառս Աստուծոյ կարգէին: (§182)

Und als Diener des Heiligen, der die Seligkeit erreicht hatte, bestellte man zur Verherrlichung Gottes eines Schüler desselben namens T'adik, einen aufgeweckten und frommen Mann, zusammen mit [anderen] Brüdern.

Ընդդէմ լինէր սուրբ Եպիսկոպոսացն, որոց առաջնոյն Բաբիլաս անուն եւ երկրորդին Ակակիոս. (§44)

Er stellte sich den heiligen Bischöfen vor, deren erster Babilas hieß und der zweite Akakius.

Discussion

The concept “name” (անուն) is also present in most of the other synoptic tables. The reason why “name” (անուն) regularly appears among the top entries in the “concept” domain is due to the general importance of persons in the context of *surb*. As the previous microanalyses have already revealed, the use of *surb* in the context of persons can either be related to groups with an undefined number of members (for instance, “the saints” or “the holy bishops” as a collective term) or to individual persons.

The term “name” (անուն) often occurs in the context of the latter. These are frequently characterized by attribute chains and are called by their individual name. In place of the long attribute chains applied in the “Epic Histories” and the “His-

tory of the Armenians,” the “Life of Maštoc” often uses only two attributes to describe individual persons called *surb* in the form of a short subclause. The sentence below provides another example where the terms զգաստ (“modest”) and բարեպաշտան (“pious”) appear in the context of *surb*.

Verb

In order not to repeat my remarks concerning generic verbs such as առնել (“to do”) or կոչել (“to call”) (see the previous microanalyses), I will directly review the usage of the verb “to intercede” (բարեխօսել) that the microanalysis made apparent.

Եւ բարեխաւսել նորա վասն սրբոց եւ բարեխաւսել Հոգւոյն Սրբոյ՝ առ ի վարդապետելոյ մեզ, զի ընդ միմեանց բարեխաւսելն գիտելի է. քանզի միապատիւ է աստուածականն եւ ոչ բազմաբար: (§155)
Und was die Fürsprache der Heiligen anbetrifft und die Fürsprache des Heiligen Geistes, sie sind für unsere Belehrung gedacht, daß wir füreinander eintreten, denn das Göttliche ist gleich und nicht unterschiedlich verehrenswürdig.

Discussion

The verbs “to intercede” (բարեխօսել) and “intercession” (բարեխօսութիւն) are usually found with reference to martyrs and saints in the Armenian accounts. The potential to “intercede” (before god) is a typical characteristic of saints in early Christian traditions, as has been demonstrated by Angenendt and Peter Brown (Angenendt 2007; Brown 2015).¹³ In consequence, the prominence of persons who intercede before god on behalf of others in the “Life of Maštoc” further indicates that Armenian Christianity in late antiquity must be considered within the framework of the larger early Christian world.

Place

With a “sum” value of 71, the lemma հայ (“Armenia(n)”) is the most conspicuous word in the “place” domain.

Վասն որոյ առեալ երանելոյն Մաշթոցի դաս մի մանկտոյ հրամանաւ արքայի եւ միաբանութեամբ սրբոյն Սահակայ, եւ հրաժարեալք ի միմեանց համբուրի սրբութեանն՝ խաղայր գնայր ի հինգերորդ ամի Վռամշապհոյ արքային Հայոց, եւ երթեալ հասանէր ի կողմանս Արամի՝ ի քաղաքս Երկուս Ասորոց. որոց առաջինն Եղեսիա կոչի, եւ երկրորդին Ամիդ անուն: (§43)

Deshalb nahm der selige Maštoc՝ auf Befehl des Königs und mit dem Einverständnis des heiligen Sahaks eine Gruppe Jugendlicher mit. Und nachdem sie sich mit einem heiligen Kuß voneinander verabschiedet hatten, machte er sich im fünften

13 See also Chapter 2 “Holy and Persons.”

Jahr des armenischen Königs Vřamřapuh auf den Weg und gelangte dabei in die Gegenden von Aramāa, in zwei Städte Syriens, deren erste Edessa hieß, und die zweite den Namen Amida [hatte].

Discussion

The political implications of the word հայ (“Armenia(n)”) in the Armenian accounts and its indirect connection to the notion of *surb* has already been discussed in the previous parts of the analysis.

The synoptic table in the AAFTC/AASC Koriwn demonstrates that other spatial terms, such as քաղաք (“city”), իտրաւն (“land”), and կողմն (“region”), point into a similar direction. They are only rarely used in direct connection with *surb*. Yet, they indicate that “travelling to different regions or cities” (to do missionary work) was an important activity in the life of those persons called *surb*. In addition to the use of հայ (“Armenia(n)”) in this and other Armenian sources, these terms once more underline the importance of geographical and political demarcation processes at play in the Armenian accounts as well as the role of persons called *surb* therein.

Besides, the use of *surb* in this example sentence also illustrates another application of *surb* that evokes associations to purity, namely the “holy kiss” (literally: “a kiss of holiness/purity,” համբուրիւ սրբութեանն).

Title

With a “sum” value of 100, the lemma եպիսկոպոս (“bishop”) is the most important word in the “title” domain. The term bishop is also the lemma with the highest “sum” value in the synoptic table of the “Life of Mařtoc’.”

Ընդդէմ ինէր սուրբ եպիսկոպոսացն, որոց առաջնոյն Բաբիլաս անուն եւ երկրորդին Ակակիոս. (§44)

Er stellte sich den heiligen Bischöfen vor, deren erster Babilas hieß und der zweite Akakius.

Եւ յանձն արարեալ զնոսա սրբոյ եպիսկոպոսին Սամուէլի, այնմ՝ զոր ի վերոյ նշանակեցաք, ինքն դառնայր ի կողմանս Հայոց Մեծաց. (§131)

Und nachdem er sie dem von uns oben [bereits] erwähnten, heiligen Bischof Samuël anvertraut hatte, kehrte er selbst in das Gebiet von Groß-Armenien zurück.

Discussion

The incidence of titles such as bishop or *katholikos* on the meaning of *surb* has already been reviewed.¹⁴ My comments in this respect also apply to the use of titles in the “Life of Maštoc’.”

However, it is remarkable that the use of titles in the context of *surb* in the “Life of Maštoc’” seems to be more important than that of persons; this is at least what the “sum” values suggest. In truth, titles are often used as synonyms for persons in the “Life of Maštoc’” (“the bishop” instead of calling a person by his name). The category of “titles,” therefore, needs to be evaluated together with that of “persons.”

Person

With a “sum” value of 43, the name սահակ (Sahak) is the most significant word in the “person” domain.

Եւ իբրեւ ատուրս բազումս անդէն ի նմին դեգերեր, յարուցեալ այնուհետեւ հասանէր առ սուրբ կաթողիկոսն Հայոց Մեծաց, որոյ անունն ճանաչէր Սահակ, զոր պատրաստական գտանէր նմին փութոյ հաւանեալ: (§29)

Und nachdem er viele Tage damit zugebracht hatte, erhob er sich [und] begab sich zu dem heiligen Katholikos von Groß-Armenien, dessen Namens, Sahak, er sich entsann, [und] den er bereit fand, da er dessen Sorge teilte.

Discussion

Together with Maštoc’, Sahak is the most important person in the “Life of Maštoc’.” Sahak, the son of Nersēs, was a *katholikos* (or patriarch) in Greater Armenia from 386 CE onwards and the last descendant of the Gregorid dynasty.¹⁵ The frequent mention of a highly ranked ecclesiast in the Armenian Church in the context of *surb* (who, in this case, is also one of the main protagonists in the “Life of Maštoc’”) further underlines how *surb* may distinguish extraordinary religious persons.

Time

With a “sum” value of 31, the lemma օր (“day”) is the central word in the “time” domain, closely followed by ամ (“year”) with a “sum” value of 28. Yet, the collected data in the synoptic table reveals that the “time” domain is scarce. Furthermore, the lemmas in the “time” domain appear with a relatively low frequency and are mostly indications of dates. They are thus unrelated to the meaning of *surb* and will consequently not be discussed.

14 See the microanalysis of the “Epic Histories.”

15 Regarding the use of both titles in “The Life of Maštoc’,” see Winkler and Koriun (1994, 245:219f.) Most of the information about the life of Sahak stem from Winkler and Koriun (1994, 245:221ff.).

5.1.4.5 Final Discussion

The microanalysis of the “Life of Maštoc” mainly confirms aspects that have already been mentioned.

For instance, as in other sources, the use of *surb* in the “Life of Maštoc” is frequently found in reference to religious persons and titles. There are also indications of a link between *surb* and the field of purity. In this context, one may recall the example of the “holy kiss.” The link between *surb* and purity or healing is also evoked in the “Life of Maštoc” (§123 and §48).

Compared to the other examinations, the macroanalysis of the “Life of Maštoc” underscores the role of spatial notions. However, since these notions mostly appear in the word frequency lists and not in the collocation analysis, they are not directly linked to the terminology related to *surb*. Yet, they remain pertinent to the broader context of the account by Koriwn.

5.1.5 AAFTC/AASC Gospels

5.1.5.1 Keyword Analysis (“Types” and “Lemmas”)

Unlike the other texts examined so far, the analysis of the keywords in the Gospels takes into account the “types” and “lemmas.” This has been possible thanks to the publicly available lemmatization of the Gospels as well as of the entire Armenian version of the Bible on *Arak 29 Armenian Bible*.

Types

The tables below present the ten most important keywords and lockwords. The keywords of the AASC node corpus are in the left column and those of the reference corpus AAFTC in the right column. The lockwords are in the middle column (words that appear to a similar extent in both corpora). The first table (see the corresponding figure 29) includes the original Armenian terms. The second table contains the English translations of the Armenian terms. All words related to *surb* have been removed because they obviously are the most widespread keywords in the node corpus AASC. The table with the English translations only lists the English lemmas of the inflected Armenian words. Consequently, it does not reflect the morphological variance of the Armenian words in the first keyword table.

Figure 29: Keywords and lockwords (“types”) in the AASC (+) and AAFTC (-).

1/512	Keywords +	1/33	Lockwords	1/7687	Keywords -
1	բորոտք	1	ասել	1	ասեն
2	մկրտեսցէ	2	ելեալ	2	ասեմ
3	հոգւովն	3	զինչ	3	ասեին
4	հոգւով	4	քան	4	ամեն
5	ջրով	5	է	5	այս
6	իրով	6	զյիսուս	6	եկն
7	հոգին	7	յիս	7	աշակերտքն
8	հոգւոյն	8	սկսաւ	8	մեզ
9	աղքատք	9	աստուծոյ	9	ել
10	ամօթ	10	էին	10	այն

Table 27: English translation keywords and lockwords in the AASC (+) and AAFTC (-).

Keywords(+)	Lockwords	Keywords(-)
leper	to say	to say
to baptize	to go (out)	to say
spirit	what	to say
spirit	more	Amen
water	to be	this
fire	Jesus	external (?)
spirit	PRON	disciple
spirit	to begin	PRON
beggar	god	to go (out)
shame	to be	PRON

The examination of the keywords (“types”) adds new insights into the meaning of *surb*. Leaving aside the keywords in the reference corpus (AAFTC) and the lockwords which are mostly related to “to say” (ասեմ in several inflected forms) and hence less relevant, it is interesting to note that the first-placed term in the “types” AASC column are the “lepers” (բորոտք). Needless to say, the importance of lepers in the keyword analysis adds to the discussion on the relevance of the semantic fields of purity and healing for *surb*. This relevance is further strengthened

by other keywords in the node corpus, such as “to baptize” (մկրտեմ), “fire” (հուր), and “water” (ջուր). Other interesting notions are “beggar” (աղքատ) and “shame” (ամօթ). All of these words are at least potentially related to the semantic field of purity as well.

Lemmas

The tables below present the ten most important lemmatized keywords and lockwords. Similarly to the previous keyword analysis, the first table (see the corresponding figure 30) includes the Armenian terms, whereas the second table displays the English translations of the items in the first table. As usual, all words related to *surb* have been removed because they obviously build the most present keywords in the node corpus AASC.

Here again, the examination of the keywords (“lemmas”) brings crucial points to light. The keywords in the reference corpus (AAFTC) and the lockwords show that verbs such as “to listen” (լսել), names such as Herodes, and concepts such as “power” (զորութիւն) are found to a similar extent in both corpora. In the AAFTC corpus, however, words such as “people/multitude” (ժողովուրդ), “to believe” (հաւատալ), and “Petros” are much more frequent.

When compared to the keywords in the AASC (“types”) corpus, the keywords in the AASC show similarities and differences. For instance, “spirit” (հոգի), “lepers” (բրորոտք), “shame” (ամօթ), and “leprosy” (բրորութիւն) are among the top lemmas in the AASC Gospels “types” and “lemmas” corpora. However, other keywords from the AASC “type” keyword study are missing in the AASC (“lemmas”) analysis of keywords.

Table 28: English translation keywords and lockwords in the AASC (+) and AAFTC (-).

Keywords(+)	Lockwords	Keywords(-)
leprosy	to_listen	multitude
leper	power	house
spirit	to sew	to believe
shame	more	to thank
joist	all	PRON
gratuitous	Herodes (name)	to eat
structure	excessive	Amen
ravaging	sins	brother
inextinguishable	to approach	Petrus
to reach	man	priest

Figure 30: Keywords and lockwords (“lemmas”) in the AASC (+) and AAFTC (-).

1/236	Keywords +	1/17	Lockwords	1/2030	Keywords -
1	բորոտութիւն	1	լսել	1	ժողովուրդ
2	բորոտ	2	զօրութիւն	2	տուն
3	հոգի	3	կարել	3	հաւատալ
4	ամօթ	4	քան	4	խնդրել
5	հեծան	5	ամենայն	5	դա
6	ծրի	6	հերովդէս	6	ուտել
7	յարդ	7	առաւել	7	ամէն
8	յափշտակութիւն	8	մեղք	8	եղբայր
9	անշէջ	9	մերձենալ	9	պետրոս
10	հասուցանել	10	այր	10	քահանայապետ

For instance, neither “baptize, wash with water” (մկրտեմ), nor “fire” (հուր) or “water” (ջուր) are among the top keywords in the AASC Gospels (“lemmas”). Instead, lemmas such as “gratuitous” (ձրի), “ravaging” (յափշտակութիւն), and “inextinguishable” (անշէջ) are found. Therefore, even though the semantic fields of purity and healing are still perceptible, they lack some of the notions that appeared in the AASC Gospels (“types”) corpus keyword analysis.

5.1.5.2 Word Frequency Lists

Types

No *Surb*

The fifteen most common words in the AASC Gospels subcorpus (no stop words, no words related to *surb*) are shown in following table and bar plot (see the corresponding figure 31).

Table 29: Most frequent words (top 15) in AASC Gospels.

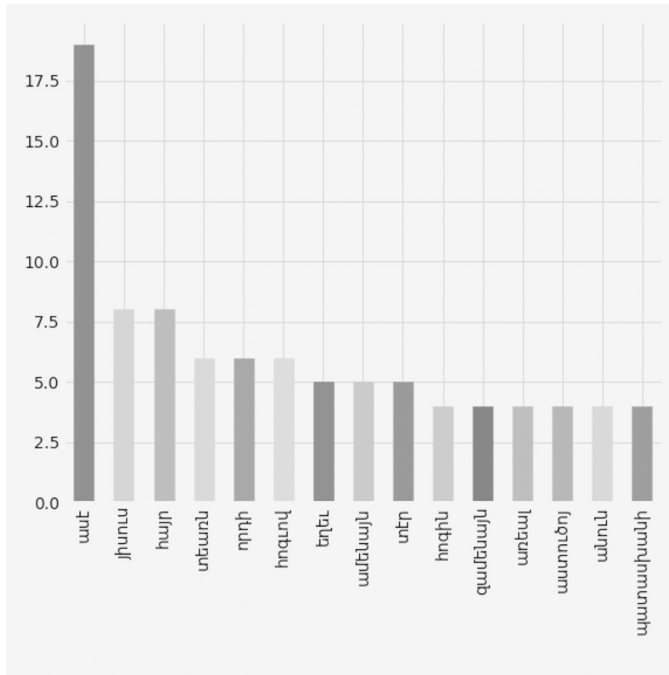
Type	Translation	Frequency
ասէ	to say	19
յիսուս	Jesus	8
հայր	father	8
տեառն	lord	6
որդի	son	6
հոգով	spirit	6
եղէւ	to be	5
ամենայն	all	5
տէր	lord	5
հոգին	spirit	4
զամենայն	all	4
առեալ	to take	4
աստուծոյ	god	4
անուն	name	4
պատասխանի	answer	4

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common words are (no *surb*; overall frequency: 92):

- Words related to (mostly) religious “Concepts” (հայր, տեառն, աստուծոյ, տէր, յիսուս, հոգով, հոգին, որդի) (AF 43, RF 47%)
- Verbs such as “(to) speak, answer” or “to receive” (ասէ, պատասխանի, եղէւ, առեալ) (AF 32, RF 34%)
- Words related to non-religious “Concepts” (հայր, անուն, որդի, ամենայն, զամենայն) (AF 21, RF 23%)

Overall, there is a strong focus on religious “concepts” such as “god” and “spirit” (AF 43, RF 47%) as well as on verbs. Non-religious concepts already encountered during the previous examinations, such as “all” and “name,” are observable as well. It is interesting to note that the “person” and “title” domains are missing in the case of the Gospels. Yet, as both the establishment of a church hierarchy and the phenomenon of Christian saints are later developments that are subsequent to the Gospels, it is

Figure 31: Most frequent words (top 15) and their absolute frequencies in AASC Gospels.



not surprising that the two domains do not appear in the word frequency list of the Gospels.

Surb

The root *surb/srb-* appears 84 times in the 79 sentences in the AASC Gospels. For the visualization, see the corresponding figure 32.

- Most common is the use of *surb* as an adjective or noun (AF 50, RF 59.5%).
- At least compared to the other Armenian sources in the Armenian corpus, the root *surb/srb-* in the Gospels is relatively frequently used as a verb (սրբել) (AF 26, RF 31%).
- The use of *surb* as a noun, such as in *surb-ness* (սրբութիւն), is less common (AF 8, RF 9.5%).

Table 30: All inflected forms of lexical units based on the root *surb/srb-* in AASC Gospels.

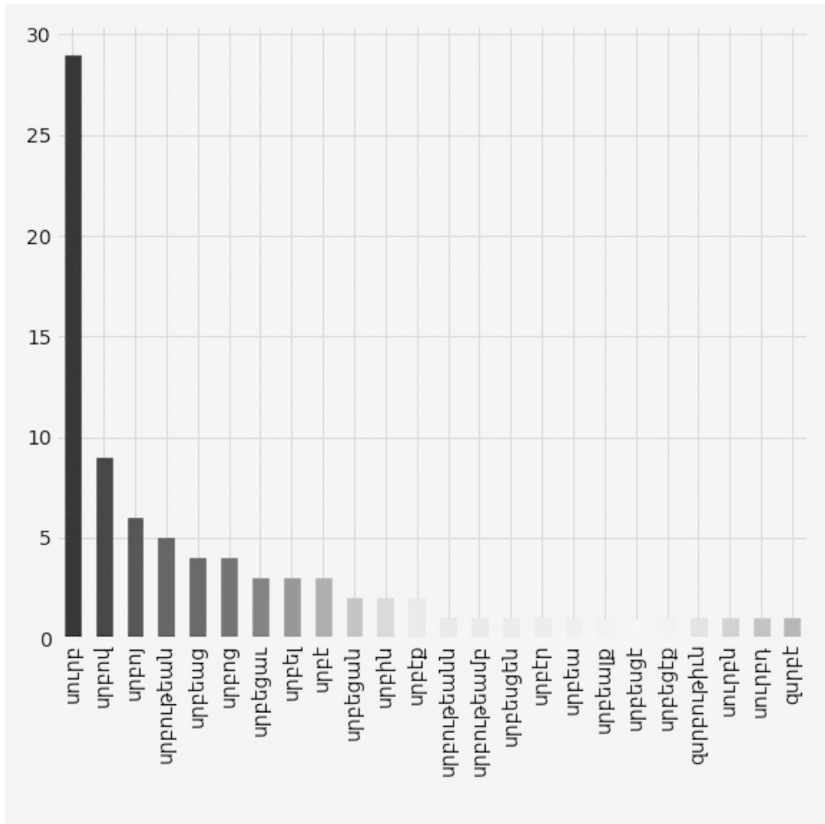
Type	Frequency
սուրբ	29
սրբով	9
սրբոյ	6
սրբութեան	5
սրբեաց	4
սրբոց	4
սրբեցաւ	3
սրբել	3
սրբէ	3
սրբեցան	2
սրբին	2
սրբեք	2
սրբութեանն	1
սրբութեամբ	1
սրբեսցեն	1
սրբեր	1
սրբեա	1
սրբեալք	1
սրբեսցէ	1
սրբեցէք	1
զսրբութիւն	1
սուրբն	1
սուրբդ	1
զսրբէ	1

Lemmas

The fifteen most common lemmas in the AASC Gospels subcorpus (no stop words, no *surb*) are shown in following table and bar plot (see the corresponding figure 33).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common words are (no *surb*; overall frequency: 187):

Figure 32: All inflected forms of lexical units based on the root *surb/srb-* in AASC Gospels (bar plot).



- Verbs such as “to make” or “to speak” (առնել, խօսել, առնուլ, մկրտել, գալ, տալ, ասել, տեսանել) (AF 100, RF 54%)
- Words related to religious “Concepts” (տէր, հայր, որդի, յիսուս) (AF 44, RF 24%)
- Words related to non-religious “Concepts” (ամենայն, որդի, հայր, անուն) (AF 38, RF 20%)
- Words related to “Titles” (տէր) (AF 13, RF 6%)

Table 31: Most frequent lemmas (top 15) in AASC Gospels.

Lemma	Translation	Frequency
ասել	to say	36
հոգի	spirit	26
տէր	lord	13
հայր	father	12
գալ	to come	12
ամենայն	all	10
յիսուս	Jesus	10
տեսանել	to see	10
տալ	to give	10
մկրտել	to baptize	9
որդի	son	9
առնուլ	to take	8
առնել	to make	8
անուն	name	7
խօսել	to speak	7

The results in the word frequency list of “lemmas” are more or less equal to the results of the “types” analysis of the Gospels, although the importance of verbs is even more prominent in the case of the “lemmas.” Unlike what was observed during the “types” analysis, the “title” domain is present in the word frequency table of the “lemmas” as well. However, the only title that appears is “lord,” a common Christian synonym for god.

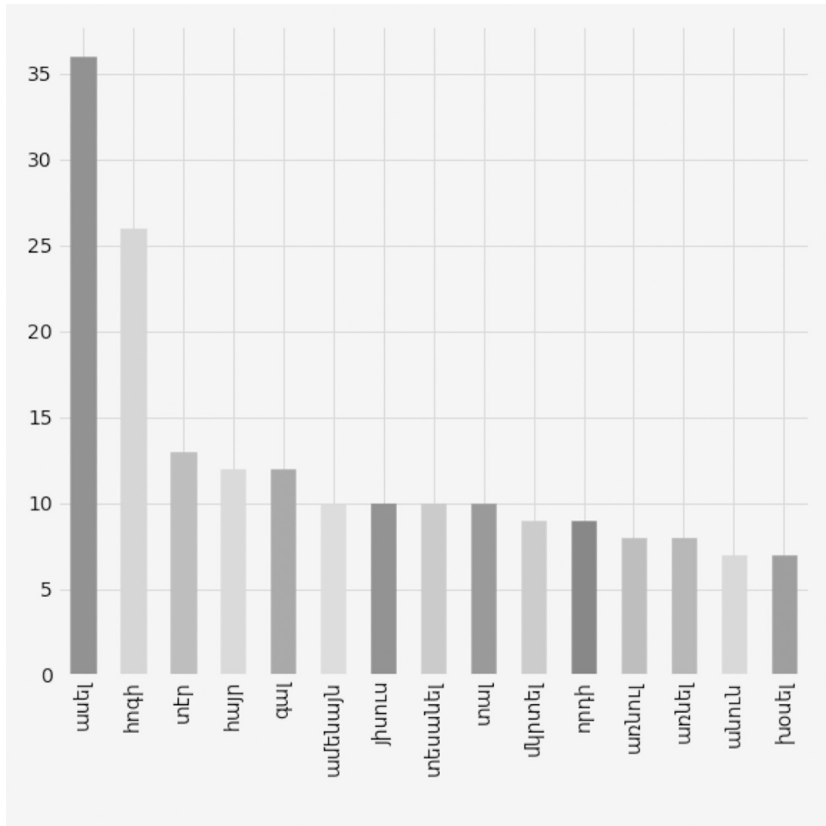
5.1.5.3 Collocation Analysis (AAFTC, Top 15)

Types

The fifteen most common words in the AAFTC Gospels subcorpus (no stop words) that appear in the context of words related to *surb/srb-* are shown in the following table (6-LL, L2-R2, C2-NC2; no function words). The visualization can be found in the corresponding figure 34.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 42).

Figure 33: Most frequent lemmas (top 15) in AASC Gospels (bar plot).



- Words related to religious “Concepts” (հոգուվ, հոգին, հոգւոյն, հոգւովն, հոգի, զհոգին, հրեշտակաց) (AF 24, RF 57%)
- Words related to non-religious “Concepts” (խոսք, բորոտք, անուն) (AF 7, RF 16%)
- Verbs such as “to want” (զնան, կամիմ) (AF 5, RF 12%)
- Words related to “Objects” (զկալ, զձեռն) (AF 4, RF 10%)
- Words related to “Persons” (յովսէփ) (AF 2, RF 5%)

The results of the collocation analysis (“types”) are discussed together with those of the collocation analysis (“lemmas”) below.

Table 32: Top 15 words collocation analysis (6-LL, L2-R2, C2-NC2; no function words) *surb/srb-* in AAFTC Gospels.

id	pos	coll	translation	stat	freq_(coll)	freq_(corpus)
1	l	հոգւով	spirit	61.43	6	13
2	l	հոգիւն	spirit	41.73	4	8
3	l	հոգւոյն	spirit	41.73	4	8
4	l	հոգւովն	spirit	31.26	3	6
5	l	հոգի	spirit	28.13	3	9
6	r	զկալ	floor	26.36	2	2
7	r	խուլք	deaf	26.36	2	2
8	l	կամիմ	to intend	24.16	3	16
9	l	բորոտք	leper	22.54	2	3
10	l	գնան	to come	20.82	2	4
11	r	անուն	name	17.56	3	46
12	l	զհոգիւն	spirit	16.84	2	9
13	r	զձեռն	hand	15.95	2	11
14	l	հրեշտակաց	angel	15.95	2	11
15	m	յովսէփ	Joseph	15.95	2	11

Lemmas

The fifteen most common lemmas in the AAFTC Gospels subcorpus (no stop words) that appear in the context of the lemma *surb* are shown in following table (6-LL, L2-R2, C2-NC2; no function words). The visualization can be found in the corresponding figure 35.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 60).

Figure 34: Top 15 types collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb/srb- in AAFTC Gospels (graph) (c) LancsBox.

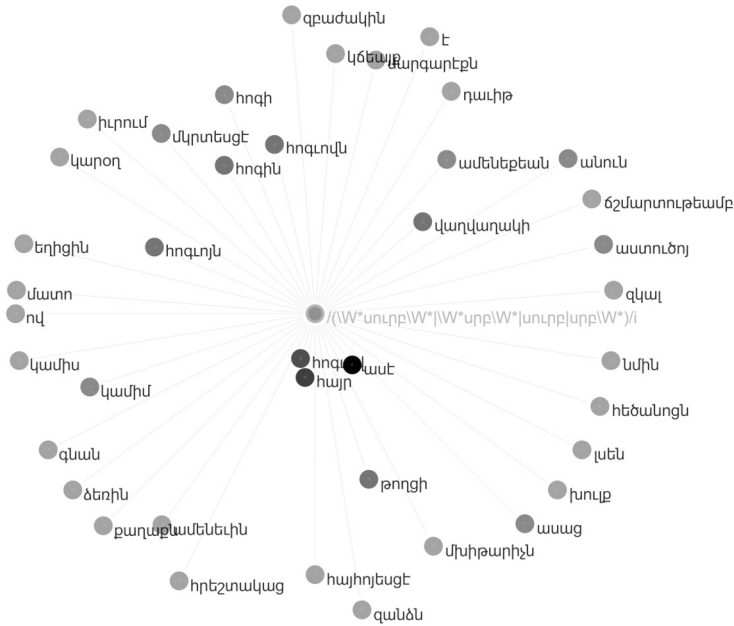
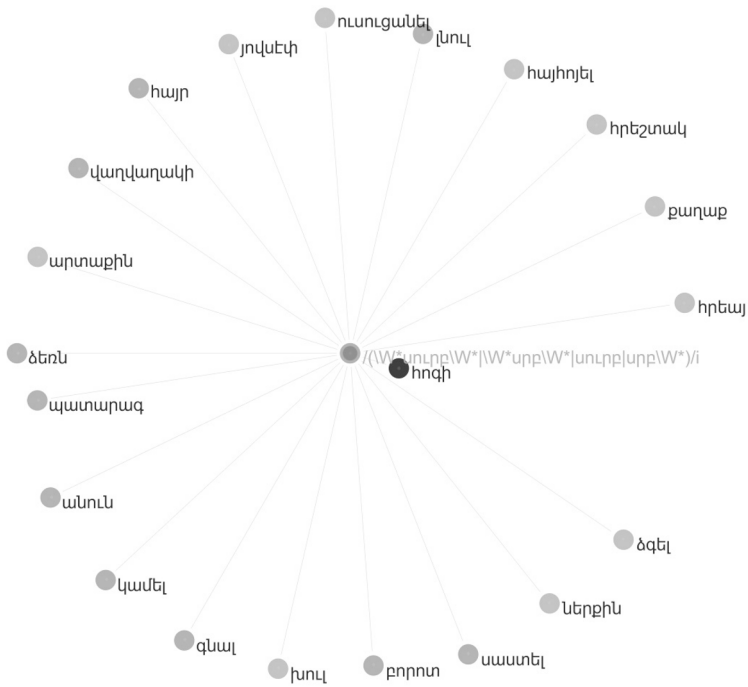


Table 33: Top 15 lemmas collocation analysis (6-LL, L2-R2, C2-NC2; no function words) surb in AAFTC Gospels.

id	pos	coll	translation	stat	freq_(coll)	freq_(corpus)
1	l	հոգի	spirit	241.12	24	61
2	l	բորոտ	leper	28.08	3	9
3	r	սպասարագ	sacrifice	25.50	3	13
4	r	ներքին	inner	22.51	2	3
5	r	սաստել	to rebuke	19.69	3	32
6	r	խուլ	deaf	19.60	2	5
7	m	արտաքին	external	16.81	2	9
8	l	լնուլ	to fill	16.27	3	56
9	m	հայհոյել	to blaspheme	15.92	2	11
10	r	վաղվաղակի	suddenly	15.37	3	65
11	r	ձգել	to stretch	13.58	2	19

12	l	գնալ	to go	13.07	3	96
13	m	յովսէփ	Joseph	12.62	2	24
14	r	անուն	name	12.33	3	109
15	r	ձեռն	hand	11.88	3	118

Figure 35: Top 15 lemmas collocation analysis (6-LL, L2-R2, C3-NC3; no function words) surb in AAFTC Gospels (graph) (c) LancsBox.



- Words related to religious “Concepts”(պատարագ, հոգի) (AF 27, RF 45%)
- Verbs such as “to rebuke” (սաստել, լնուլ, հայիոյել, ձգել, գնալ) (AF 13, RF 22%)
- Words related to non-religious “Concepts” (բորոտ, խուլ, անուն) (AF 8, RF 13%)
- Misc. (ձեռն, վաղվաղակի, արտաքին, ներքին) (AF 6, RF 10%)

- Words related to “Persons” (յովսէփ) (AF 2, RF 3%)

Similarly to the previous collocation analysis (“types”), the term “spirit” frequently appears in the context of *surb* in the Gospels. Other religious concepts, however, are only rarely mentioned. Instead, there are several words related to sickness and impurity, such as “deaf” (խուլ) and “leper” (բորոտ). Some of these terms also appeared in the keyword analysis of the Gospels and once more underline the potential relevance of purity and healing (sickness) for the use of *surb* in the Gospels. Similarly to the previous word frequency lists, persons seldom occur in the collocation analysis (the only exception is Joseph).

5.1.5.4 Synoptic Table – Sentences Selection and Discussion (Microanalysis)

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses as well as their corresponding values and annotation has been created.

As explained in Chapter 3, the selection of example sentences for the microanalysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be exceptions in some cases. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *surb* appears in the neighborhood of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

In contrast to the other microanalyses, the parallel Greek version will be provided in addition to the Armenian text. The comparison of the two versions of the Gospels allows assessing the Greek influence on the notion of *surb*. The limitations of a study that cannot take into consideration the Syriac version have already been announced in Chapters 2 and 3. The results of the forthcoming analysis need to be interpreted accordingly.

The following domains appear in the Gospels (ordered by their summarized “sum” values):

Table 34a: Synoptic table of the AASC and AAFTC Gospels.

id	lemma	kw_t	kw_l	wf_t	cf_t	wf_l	cf_l	sum	do- main	semantic_field
1	բորոս	1	1	0;0	1;2	0	3	29	con- cept	leper_B2-/S2mf
2	մկրտել	1	0	0;0	0;0	9	0	22	verb	baptize_S9
3	հոգի	5	1	2;10	6;22	26	24	100	con- cept	spirit_S9
4	ջուր	1	0	0;0	0;0	0	0	14	ob- ject	water_O1.2
5	աղքատ	1	0	0;0	0;0	0	0	14	con- cept	beggar_I1.1-/ S2mf
6	հուր	1	0	0;0	0;0	0	0	14	ob- ject	fire_O4.6+
7	ամօթ	1	1	0;0	0;0	0	0	23	con- cept	shame_G2.2-
8	ասել	-1	0	1;19	0;0	36	0	48	verb	say_Q2.1
9	յիսուս	-1	0	1;8	0;0	10	0	13	per- son	christ_Z1m
10	հայր	0	0	1;10	0;0	12	0	26	con- cept	father_S4m
11	տէր	0	0	2;11	0;0	13	0	29	title	lord_S7.1+/ S2.2m
12	որդի	0	0	1;6	0;0	9	0	20	con- cept	son_S4m
13	ամենայն	0	0	2;9	0;0	10	0	24	con- cept	all_N5.1+
14	առնուլ	0	0	1;4	0;0	8	0	17	verb	take_A9+
15	աստուած	-1	0	1;4	0;0	0	0	0	con- cept	god_S9
16	անուն	0	0	1;4	1;3	7	3	22	con- cept	name_A1
17	պա- տասխանի	0	0	1;4	0;0	0	0	9	con- cept	answer_Q2.2
18	զկալ	0	0	0;0	1;2	0	0	7	ob- ject	floor_H2
19	խուլ	0	0	0;0	1;2	0	2	9	con- cept	deaf_X3.2-
20	կամիլ	0	0	0;0	1;3	0	0	8	verb	intend_X7+
21	գնալ	0	0	0;0	1;2	0	3	10	verb	go_M1

Table 34b: Synoptic table of the AASC and AAFTC Gospels.

22	ձեռն	0	0	0;0	1;2	0	3	10	object	hand_B1
23	հրեշտակ	0	0	0;0	1;2	0	0	7	concept	angel_S9mf
24	յովսէփ	0	0	0;0	1;2	0	3	10	person	joseph_Z1m
25	բորոտիւն	0	1	0;0	0;0	0	0	14	concept	leprosy_B2-
26	ձրի	0	1	0;0	0;0	0	0	14	concept	gratuitous_S6-
27	հեծան	0	1	0;0	0;0	0	0	14	object	joist_O2
28	յարդ	0	1	0;0	0;0	0	0	14	object	straw_L3
29	յափշտակութիւն	0	1	0;0	0;0	0	0	14	concept	ravaging_A1.1.2
30	անշեջ	0	1	0;0	0;0	0	0	14	attribute	inextinguishable_Z99
31	հասուցանել	0	1	0;0	0;0	0	0	14	verb	reach_M1
32	գալ	0	0	0;0	0;0	12	0	16	verb	come_M1
33	տեսանել	0	0	0;0	0;0	10	0	14	verb	see_X3.4
34	տալ	0	0	0;0	0;0	10	0	14	verb	give_A9-
36	առնել	0	0	0;0	0;0	8	0	12	verb	make_A1.1.1
37	խօսել	0	0	0;0	0;0	7	0	11	verb	speak_Q2.1
38	պատարագ	0	0	0;0	0;0	0	3	7	concept	sacrifice_S9
39	ներքին	0	0	0;0	0;0	0	2	6	attribute	inner_M6
40	սաստել	0	0	0;0	0;0	0	3	7	verb	rebuke_Q2.2
41	արտաքին	0	0	0;0	0;0	0	2	6	attribute	external_M6
42	լնուլ	0	0	0;0	0;0	0	2	6	verb	fill_N5.1+
43	հայհոյել	0	0	0;0	0;0	0	2	6	verb	blaspheme_Q2.2/ S7.2-/S9
44	վաղվաղակի	0	0	0;0	0;0	0	3	7	misc	suddenly_N3.8+
45	ձգել	0	0	0;0	0;0	0	2	6	verb	stretch_N3.6

Table 35: Domains in the AAFTC/AASC Gospels in a descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	332
Verb	211
Object	73
Title	29
Attribute	26
Person	23
Misc	7

Concept

With a “sum” value of 100, the lemma հոգի (“spirit”) is the most important word in the “concept” domain. It is also the most significant word in the synoptic table according to the “sum” value.

Ես մկրտեմ զձեզ ջրով յապաշխարհութիւն. բայց որ զկնին իմ գայ՝ հօրագոյն է քան զիս, եւ ես չեմ բաւական բառնալ զկօշիկս նորա. նա մկրտեսցէ զձեզ ի հոգին սուրբ եւ ի հօր: (Matthew 3,11)

Ἐγὼ μὲν ὑμᾶς βαπτίζω ἐν ὕδατι εἰς μετάνοιαν, ὁ δὲ ὀπίσω μου ἐρχόμενος ἰσχυρότερός μου ἐστίν, οὗ οὐκ εἰμι ἰκανὸς τὰ ὑποδήματα βαστάσαι· αὐτὸς ὑμᾶς βαπτίσει ἐν πνεύματι ἁγίῳ καὶ πυρί·

I baptize you with water for repentance, but he who is coming after me is mightier than I, whose sandals I am not worthy to carry. He will baptize you with the Holy Spirit and fire.¹⁶

Discussion

The combination of *surb* and հոգի (*hogi*, “spirit”) is ubiquitous in most of the other Armenian sources as well. The “spirit” usually appears among the top entries in every “concept” domain.

The “spirit” is central in the Old Testament and the New Testament alike. Its role in Christianity and in other religions such as Judaism is a debated topic. Reviewing the debate is beyond the scope of this book (Schmidt, Schäfer, and Berger 1980). However, the combination of *surb* (or “holy” in the English and *hágios* in the Greek version of the Gospels) and *hogi* (English “spirit” and Greek *pneúma*) is crucial for our purpose since it frequently appears in the Armenian and the English sources. The common use of the “holy spirit” as a fixed expression in the form of the *surb*

16 The English translation of the Gospels follows the English Standard Version (ESV).

hogi in all the Armenian sources testifies to the influence of non-Armenian sources such as the Gospels and their translations on the Armenian notion of *surb*.

The selected sentence illustrates another important dimension of the use of *surb* in the Armenian Gospels, namely that of baptism (see also the microanalysis of the “verb” domain).¹⁷ The idea of baptism is closely connected to purification in the New Testament. It is also closely linked to water and fire. The occurrence of these two notions in the synoptic table further indicates their relevance for *surb*.¹⁸ In the passages related to baptism, the “holy spirit” stands for the spiritual element of the purification that is accompanied by a physical element (water). The use of an attribute such as *surb* underlines the importance of qualifying the “spirit” as related to purification. Yet, the fact that the “holy spirit” is almost always called “holy spirit,” and not exclusively in situations linked to baptism or other types of purification, suggests that the use and meaning of *surb* in this context goes beyond the application of a specific attribute that stands for a purificatory function of the “spirit.”¹⁹

In sum, this passage demonstrates that the close relationship between *surb* and the semantic field of purity is not only present in the original Armenian sources but is also traceable in non-Armenian sources such as the Gospels. The importance of purity within the Armenian translation of the Gospels and its impact on the meaning of *surb* will be further dealt with in the final part of this section.

Verb

With a “sum” value of 22, the lemma մկրտել (“to baptize”) is the second most conspicuous word in the “verb” domain. The most important word is ասել (“to say”) with a value of 48. However, as ասել (“to say”) is also among the lockwords in the analysis of keywords, it will not be discussed.

Գնացէք այսուհետեւ՝ աշակերտեցէք զամենայն հեթանոսս.
մկրտեցէք զնոսա յանուն Հոր՝ եւ Որդւոյ՝ եւ Հոգւոյ սրբոյ: (Matthew
28:19)

πορευθέντες οὖν μαθητεύσατε πάντα τὰ ἔθνη, βαπτίζοντες αὐτοὺς
εἰς τὸ ὄνομα τοῦ πατρὸς καὶ τοῦ υἱοῦ καὶ τοῦ ἁγίου πνεύματος,
Go therefore and make disciples of all nations, baptizing them in the name of the
Father and of the Son and of the Holy Spirit,

17 Regarding the importance of baptism and John the Baptist in particular for the notion of “spirit” in the New Testament, see Schmidt, Schäfer, and Berger (1980, 178ff.).

18 See the microanalysis of the “object” domain below.

19 Again, see Schmidt, Schäfer, and Berger (1980, 178ff.) and other parts in the New Testament, for instance, Matthew 12:22ff.

Discussion

The importance of baptism with regard to the use of *surb*, and the *surb hogi* (“holy spirit”), in particular, has already been pointed out. This sentence gives another example of this interrelation. The phrase “(...) in the name of the Father and of the Son and of the Holy Spirit” also is an additional example of the sedimentation of once meaningful combinations of attributes and concepts in the form of fixed expressions. This process indicates that the attribution of the “pure/purifying spirit” in particular contexts such as baptism morphed into a fixed *terminus technicus* (see also the discussion of the “concept” domain above).

Object

With a “sum” value of 14, the two lemmas ջրը (“water”) and հրը (“fire”) are the most prominent words in the “object” domain.

Ես մկրտեցի զձեզ ջրով, եւ նա մկրտեցէ զձեզ հոգևով սրբով: (Mark 1:8)

ἐγὼ ἐβάπτισα ὑμᾶς ὕδατι, αὐτὸς δὲ βαπτίσει ὑμᾶς ἐν πνεύματι ἁγίῳ.
I have baptized you with water, but he will baptize you with the Holy Spirit.

Պատասխանի ետ ամենեցուն՝ եւ ասէ. ես մկրտեմ զձեզ ջրով. բայց զայ զօրագոյն քան զհս, գորոյ չեմ բաւական բառնալ զլոշիկս. նա մկրտեցէ զձեզ հոգևով սրբով եւ հրով: (Luke 3:16)

ἀπεκρίνατο λέγων πᾶσιν ὁ Ἰωάννης· ἐγὼ μὲν ὕδατι βαπτίζω ὑμᾶς· ἔρχεται δὲ ὁ ἰσχυρότερός μου, οὗ οὐκ εἰμι ἰκανὸς λῦσαι τὸν ἱμάντα τῶν ὑποδημάτων αὐτοῦ· αὐτὸς ὑμᾶς βαπτίσει ἐν πνεύματι ἁγίῳ καὶ πυρί·

John answered them all, saying, "I baptize you with water, but he who is mightier than I is coming, the strap of whose sandals I am not worthy to untie. He will baptize you with the Holy Spirit and with fire.

Discussion

This passage is another example of the aforementioned importance of baptism and purification for the meaning of *surb*. It further illustrates the role of the “holy spirit” in purification processes such as baptism, namely that of the spiritual purification which complements the physical purification by water. The verses in Mark 1:8 and Luke 3:16 show that the “spiritual purification” is considered to be superior to the mere physical purification through water.

Title

With a “sum” value of 29, the lemma *υτηρ* (“lord”) is the most important word in the “title” domain.

Եւ եղէի ի հասսանէն նորա ի մի ի քաղաքացն, ահա այր մի լի բորոտութեամբ՝ իբրեւ ետես զՅիսուս, անկեալ ի վերայ երեսաց աղաչեաց զնա՝ եւ աւէ. *υτηρ՝* եթէ կամիս, կարող ես զիս սրբել: (Luke 5:12)
 Καὶ ἐγένετο ἐν τῷ εἶναι αὐτὸν ἐν μιᾷ τῶν πόλεων καὶ ἰδοὺ ἀνὴρ πλήρης λέπρας· ἰδὼν δὲ τὸν Ἰησοῦν, πεσὼν ἐπὶ πρόσωπον ἐδεήθη αὐτοῦ λέγων· κύριε, ἐὰν θέλῃς δύνασαι με καθαρίσαι.

While he was in one of the cities, there came a man full of leprosy. And when he saw Jesus, he fell on his face and begged him, “Lord, if you will, you can make me clean.”

Discussion

That the word *υτηρ* (“lord”) is mostly used as a designation of god has already been indicated in the previous sections. In the synoptic table of the Gospels, however, the word *υτηρ* is the only title in the “title” domain. This result is not surprising since most of the other titles such as bishop, archbishop, *katholikos*, etc. emerge subsequently to the increasing institutionalization of early Christian traditions, which took place in the centuries after the Gospels.

However, the proposed sentence in this passage is still of interest for a different reason. In this sentence, the notion *surb* appears in the form of the verb *սրբել* (*srbel*; *kat' arisai* in the Greek text). This verb is here used with reference to lepers and leprosy, and it is consequently translated as “to clean.” The words “leper” (*բորոտ*), as well as “leprosy” (*բորոտիւն*), figure among the top entries in the “concept” domain in the synoptic table. Another example of the use of *surb* in the context of lepers and leprosy is the following verse.

Եւ բազում բորոտք էին յԻսրայէլի առ Եղիսեբի մարգարէի. եւ ոչ ոք ի նոցանէ սրբեցաւ, բայց միայն նէեման Ասորի: (Luke 4:27)
 καὶ πολλοὶ λεπροὶ ἦσαν ἐν τῷ Ἰσραὴλ ἐπὶ Ἐλισαίου τοῦ προφήτου, καὶ οὐδεὶς αὐτῶν ἐκαθαρίσθη εἰ μὴ Ναϊμὰν ὁ Σύρος.

And there were many lepers in Israel in the time of the prophet Elisha, and none of them was cleansed, but only Naaman the Syrian.

That *surb* in the Gospels and in other sources has a latent meaning of (spiritual) purity has been repeatedly pointed out in this book and was also noted by others.²⁰ However, in the context of “leprosy” and “lepers,” this meaning of purity

20 The meaning of “pure” or “clean” does not only appear as a translation in the common dictionaries but was also pointed out by Stempel (1988).

goes beyond that of spiritual purity and also touches upon semantic fields related to physical sickness and healing.

Interestingly, the corresponding word in the Greek text of Luke 4:27 is *kat' arísai*, a word that is not directly related to *hágios* (which has usually been found as the corresponding Greek word of *surb* so far). Instead, *kat' arísai* is related to *kat' arós*, a word that has close associations with the semantic field of purity. That the Greek terminology linked to *kat' arós* in the Gospels is often translated as *surb* can be seen in the following overview that lists all the appearances of *surb* together with their corresponding words in the Greek version of the Gospels.

Without going into the details at this stage, it is noteworthy that *surb* was used to translate a variety of Greek terms, most prominently words related to *hágios* and *kat' arós*.²¹ The implications of this observation will be further discussed in the final part of this section.

21 For a more detailed discussion of the semantic field of *hágios*, see Chapter 2. The word *kat' arós* is very closely related to the field of purity in both a religious as well as non-religious way. For a good overview of the meaning and (religious) use of *kat' arós* as well as of other words such as *hagnós*, see Robertson (2013); Parker (2003).

Table 36: *Surb and the corresponding Greek terminology in the Gospel of Matthew.*

Arm	Greek	Ref
սուրբ	ἅγιω	Mt 3:11
սուրբ	ἅγιαν	Mt 4:5
սուրբ	καθαροὶ	Mt 5:8
սուրբ	ἁγιασθήτω	Mt 6:9
սուրբ	καθάρισον	Mt 23:26
սրբեա	καθαρόν	Mt 23:26
սուրբ	ἅγιαν	Mt 27:53
սուրբ	καθαροῖ	Mt 27:59
սրբոյ	ἁγίου	Mt 1:18
սրբոյ	ἁγίου	Mt 1:20
սրբոյ	ἁγίου	Mt 12:32
սրբոյ	ἁγίου	Mt 28:19
սրբեսցէ	ακαθαρῶ	Mt 3:12
զսրբութիւն	ἅγιον	Mt 7:6
սրբել	καθαρίσαι	Mt 8:2
սրբեաց	καθαρίσθητι	Mt 8:3
սրբեցաւ	ἑκαθαρίσθη	Mt 8:3
սրբեցէք	καθαρίζετε	Mt 10:8
սրբին	καθαρίζονται	Mt 11:5
զսրբէ	ἁγιάσας	Mt 23:17
սրբէ	ἁγιάζον	Mt 23:19
սրբէք	καθαρίζετε	Mt 23:25
սրբութեան	ἅγιω	Mt 24:15
սրբոց	ἁγίων	Mt 27:52

Table 37: *Surb* and the corresponding Greek terminology in the Gospel of Luke.

Arm	Greek	Ref
սուրբ	ἅγιον	Lk 1:35
սուրբ	ἅγιον	Lk 1:35
սուրբ	ἅγιον	Lk 1:49
սուրբ	ἁγίας	Lk 1:72
սուրբ	ἅγιον	Lk 2:23
սուրբ	ἅγιον	Lk 2:25
սուրբ	ἁγιασθήτω	Lk 11:2
սուրբ	καθαρά	Lk 11:41
սուրբ	ἅγιον	Lk 12:10
սուրբ	ἅγιον	Lk 12:12
սուրբդ	ἅγιος	Lk 4:34
սրբով	ἁγίου	Lk 1:15
սրբով	ἁγίου	Lk 1:41
սրբով	ἁγίου	Lk 1:67
սրբով	ἁγίῳ	Lk 3:16
սրբով	ἁγίου	Lk 4:1
սրբով	ἁγίῳ	Lk 10:21
սրբով	ἁγίων	Lk 1:70
սրբոց	ἁγίων	Lk 9:26
սրբութեան	καθαρισμοῦ	Lk 2:22
սրբութեան	καθαρισμοῦ	Lk 5:14
սրբոյ	ἁγίου	Lk 2:26
սրբոյ	ἅγιον	Lk 3:22
սրբել	καθαρίσαι	Lk 5:12
սրբեցաւ	ἐκαθαρίσθη	Lk 4:27
սրբեաց	καθαρίσθητι	Lk 5:13
սրբին	καθαρίζονται	Lk 7:22
սրբեք	καθαρίζετε	Lk 11:39
սրբեցան	ἐκαθαρίσθησαν	Lk 17:14
սրբեցան	ἐκαθαρίσθησαν	Lk 17:17

Table 38: *Surb and the corresponding Greek terminology in the Gospel of Mark.*

Arm	Greek	Ref
սուրբն	ἅγιος	Mk 1:24
սուրբ	ἅγιον	Mk 3:29
սուրբ	ἅγιον	Mk 6:20
սուրբ	ἅγιον	Mk 13:11
սրբով	ἅγιω	Mk 1:8
սրբով	ἅγιω	Mk 12:36
սրբեաց	καθαρίσθητι	Mk 1:41
սրբեցաւ	ἐκαθαρίσθη	Mk 1:42
սրբութեան	καθαρισμοῦ	Mk 1:44
սրբէ	καθαρίζων	Mk 7:19
սրբոց	ἁγίων	Mk 8:38

Table 39: *Surb and the corresponding Greek terminology in the Gospel of John.*

Arm	Greek	Ref
սուրբ	καθαρός	Jo 13:10
սուրբ	καθαροί	Jo 13:10
սուրբ	καθαροί	Jo 13:11
սուրբ	ἅγιον	Jo 14:26
սուրբ	καθαροί	Jo 15:3
սուրբ	ἅγιε	Jo 17:11
սուրբ	ἁγιάζω	Jo 17:19
սրբեալք	ἁγιασμένοι	Jo 17:19
սուրբ	ἅγιον	Jo 20:22
Սուրբ	ἁγίασον	Jo 17:17
սրբով	ἅγιω	Jo 1:33
սրբութեանն	καθαρισμόν	Jo 2:6
սրբութեան	καθαρισμοῦ	Jo 3:25
սրբեաց	ἁγίασεν	Jo 10:36

Attribute

With a “sum” value of 14, the lemma **անշեջ** (“unextinguishable”) is the most important word in the “attribute” domain.

Որոյ հեծանոցն ի ձեռին իրում. եւ սրբեցէ զկալ իւր, եւ ժողովեցէ զցորեանն ի շտեմարանս իւր. եւ զյարդն այրեցէ անշեջ հրով: (Matthew 3:12)

οὗ τὸ πτύον ἐν τῇ χειρὶ αὐτοῦ καὶ διακαθαριεῖ τὴν ἄλωνα αὐτοῦ καὶ συνάξει τὸν σίτον αὐτοῦ εἰς τὴν ἀποθήκην, τὸ δὲ ἄχυρον κατακαύσει πυρὶ ἀσβέστῳ.

His winnowing fork is in his hand, and he will clear his threshing floor and gather his wheat into the barn, but the chaff he will burn with unquenchable fire.

Discussion

The word **անշեջ** (“unquenchable” in English or *ásbestos* in the Greek text) appears in the vicinity of “fire” and also needs to be interpreted against the background of *surb* and its relation to purity. However, the overall frequency of **անշեջ** (“unquenchable”) is very low. The only reason why this attribute is in the synoptic table is because of its appearance in the results of the keyword analysis. Thus, **անշեջ** (“unquenchable”) cannot be interpreted as a hint at the overall meaning of *surb* but rather as a seldom attribute that happens to appear in the context of “fire.”

Person

With a “sum” value of 10, the name **յովսէփ** (Joseph) is the second most significant word in the “person” domain. The most important name is Jesus. As Jesus also appears in the “concept” category (namely with his designation as Christ), he overlaps with the categories of “person” and “concept.” Consequently, the analysis will focus on Joseph instead of Jesus.

Եւ առեալ զմարմինն յովսէփ՝ պատեաց սուրբ կտաւովք, (...) (Matthew 27:59)

Καὶ λαβὼν τὸ σῶμα ὁ Ἰωσήφ ἐνετύλιξεν αὐτὸ [ἐν] σινδόνι καθαρᾷ
And Joseph took the body and wrapped it in a clean linen shroud, (...)

Discussion

From both a micro- and macro-perspective, the appearance of Joseph in the context of *surb* is not significant either. Joseph is never called *surb*, but he only appears in the proximity of terms related to *surb*.

What is more interesting in this example is the allusion to both a non-ritualistic and non-religious use of *surb*. In the sentence, the term *surb* (in Greek text καθαρᾷ

from *kat' arós*) refers to an attribute for a shroud, thereby denoting its “clean” status, otherwise devoid of any complex associations.

Admittedly, a burial context itself might be regarded as connected to religion. Yet, the syntactical structure in this sentence and particularly the use of *surb* or καθαροῦ with σινδῶν (*sindôn*, Arm. սլաւեաց) implies a more neutral use as “physically clean.”

This rare instance of an underlying meaning of *surb* lacking any associations to religious, ethical, or spiritual purity (*holiness*) indicates that its connotation of purity does not derive from Christian concepts such as the “holy spirit.” In fact, given the basic use in this example, it stands to reason that the notion of *surb* in the sense of (physical) cleanliness had already existed in Armenian and subsequently acquired a more compelling meaning of religious, ethical, or spiritual purity (*holiness*). This new meaning was – alongside others – likely shaped by Christian notions, of which the “Holy Spirit” is an example.

The hypothetical origin of the later complex term *surb* in (physical) purity corresponds very well to the theoretical implications by (Douglas 2001) as well as the possible etymological roots of *surb* from idg. **kubh-* and **keu*, meaning “hell, leuchtend” (Stempel 1988; Morani 2001). Even though “hell, leuchtend” is usually attributed to the sky, it might also indicate a gleaming “whiteness” (Stempel 1988) of a very clean object. Thus, passages like the one discussed in this example can rightfully be taken as an indication of older strata of the meaning of *surb* that had already shifted towards a more complex meaning of purity and *holiness* by the time the Armenian texts were written down.

5.1.5.5 Final Discussion

The micro examination of the Armenian and the Greek version of the Gospels has revealed differences and similarities with other Armenian sources.

Some of the differences can be explained by the fact that the Armenian Gospels were the only source examined in this book that was not originally written in Armenian but is a translation of different neighboring languages such as Greek and Syriac. Furthermore, the Gospels were composed much earlier than the other three sources. The early dating also accounts for the rare mention of titles and persons in the immediate vicinity of *surb* in the Gospels. Saints, martyrs, and ecclesiastic hierarchies occurred later and only after the writing down of the Gospels.

Two main observations derive from the microanalysis of the Gospels.

Firstly, the expression and appearance of the “holy spirit” (*surb hogi*) is much more present in the Gospels than in most of the other Armenian sources examined so far. That being said, the “holy spirit” also appears as a fixed expression in the other sources.

Secondly, even though the other microanalyses already shed light on the close connection between *surb* and purity, this relationship is more pervasive in the Gospels. The Gospels also add new aspects to this discussion on purity. For instance, terms related to sickness (mostly leprosy) and healing are closely connected to the notion of *surb* in the Gospels. The Gospels also include an example of what is likely an older meaning of *surb* conveying physical cleanliness that is devoid of any religious or otherwise complex connotations. This, in turn, suggests that *surb* was initially connected to a vast field of purity that only later became more restricted in a primarily religious understanding.

This hypothesis is supported by the observation that two distinct Greek terms are commonly translated as *surb* in the Armenian version of the Gospels. Firstly, the term ἅγιος (*hágios*) (which is usually translated as “holy,” see Chapter 2) and secondly καθαρός (*kat' arós*), a word that is usually translated as “pure” or “clean.” These two Greek terms and their respective meanings have a complex history on their own, which has been outlined in Chapter 2, and that cannot be extensively discussed at this point.

The fact that both Greek terms are translated by *surb* makes the latter's connection to a semantic field of *holy*²² and to a semantic field of purity (i.e. physical, spiritual, ethical, religious purity) more evident. This naturally begs the question whether the notion of *surb* originally expressed ideas of purity and subsequently acquired the meaning of *holy* – for instance, in reaction to the newly emerging terminology of *holy* from the Christian context – or whether the two spheres of purity and holiness were coeval but were referred to by the same term. In other words, are we dealing with a shift or with homonyms? It is also possible that there was no distinction between purity and that which is *holy* (i.e., what is *holy* is pure and reciprocally).

The lack of earlier Armenian sources before the Christianization of Armenia precludes answering this question. However frustrating, this situation recalls what has been observed with respect to the Zoroastrian terminology in Chapter 2 that, according to the secondary literature, focused on notions of purity. The Armenian case might thus rightfully be regarded as being influenced by Persian and Zoroastrian concepts and traditions that did not possess a differentiated vocabulary for *holy* and *purity* either.

In the following part, passages from the Greek Gospels where words related to both *hágios* (ἅγιος)²³ and *kat' arós* (καθαρός) are not translated as *surb* are examined.

22 If taken for granted that words such as *hágios* refer to a field of *holy* as outlined in Chapter 2.

23 Unlike before, the following part will mostly include the transliterations of both the Armenian and the Greek terms. Yet, Greek or Armenian terms in other parts of this book might still be written in their original alphabet.

Passages where the Armenian terminology around *surb* is neither translated as *hágios* nor as *kat'arós* are also considered. This overview should reveal other Armenian and Greek terminologies related to the semantic fields of *holy* and *purity*.

The first part starts with two examples from the Gospel of Matthew.

Եւ կոչեցեալ առ ինքն զերկուսասանեսին աշակերտսն իւր, ետ նոցա իշխանութիւն այսոց պիղծոց հանել զնոսա, եւ բժշկել զամենայն ցաւս եւ զամենայն հիւանդութիւնս:

Καὶ προσκαλεσάμενος τοὺς δώδεκα μαθητὰς αὐτοῦ ἔδωκεν αὐτοῖς ἐξουσίαν πνευμάτων ἀκαθάρτων ὥστε ἐκβάλλειν αὐτὰ καὶ θεραπεύειν πᾶσαν νόσον καὶ πᾶσαν μαλακίαν.

And he called to him his twelve disciples and gave them authority over unclean spirits, to cast them out, and to heal every disease and every affliction. (Matthew 10:1)

Վայ ձեզ դպրաց եւ փարիսեցոց կեղծաւորաց. զի նման էք գերեզմանաց բռելոց, որ արտաքոյ երեւին գեղեցիկք, եւ ի ներքոյ լի են ոսկերօք մեռելոց, եւ ամենայն պիղծութեամբ:

Οὐαὶ ὑμῖν, γραμματεῖς καὶ Φαρισαῖοι ὑποκριταί, ὅτι παρομοιάζετε τάφοις κεκονιαμένοις, οἵτινες ἔξωθεν μὲν φαίνονται ὠραῖοι, ἔσωθεν δὲ γέμουσιν ὀστέων νεκρῶν καὶ πάσης ἀκαθαρσίας.

"Woe to you, scribes and Pharisees, hypocrites! For you are like whitewashed tombs, which outwardly appear beautiful, but within are full of dead people's bones and all uncleanness. (Matthew 23:27)

In the Gospel of Matthew, the two examples where *kat'arós* is not translated as *surb* are both related to the semantic field of impurity which, in the Greek terminology in this example, is represented by *akat'arós*. In the Armenian terminology, however, another word deriving from the stem *պիղծ-* (*plc-*) is used. In dictionaries, *պիլ* (*պիղծ*) is commonly translated as "unclean" or "stained" and is supposed to be an antonym of *surb*.

The term *akat'arós* (Arm. *պիլ*) in this example is firstly connected to "spirit" (Grk. *pneûma* and Arm. *այս*) and secondly used in a burial context to indicate the impurity of the dead. The Armenian term *այս* (*այս*) denotes a particularly evil spirit, a demon, and is different from the word *hogi* that is applied in the expression "holy spirit" (*surb hogi*). In the Greek text, however, the term *pneûma* applies to both Arm. *այս* and Arm. *hogi*.

In the Gospel of Mark six passages are found where the terms related to *kat'* and *hag-* are not translated as *surb* in the Armenian translation.

Եւ էր ի ժողովրդեանն նոցա այր մի, յորում այս պիղծ գոյր ի նմա. որ աղաղակեաց եւ ասէ:

Καὶ εὐθύς ἦν ἐν τῇ συναγωγῇ αὐτῶν ἄνθρωπος ἐν πνεύματι

ἀκαθάρτῳ καὶ ἀνέκραξεν

And immediately there was in their synagogue a man with an unclean spirit. And he cried out, (...) (Mark 1:23)

Եւ գարնանային ամենեքին, մինչեւ հօծել ընդ միմեանս՝ եւ ասել զինչ է այս նոր վարդապետութիւնս, զի իշխանութեամբ եւ այսոց պիտոց սասուէ, եւ հնազանդին սմա:

καὶ ἔθαμβήθησαν ἅπαντες ὥστε συζητεῖν πρὸς ἑαυτοὺς λέγοντας· τί ἐστὶν τοῦτο; διδαχὴ καινὴ κατ' ἐξουσίαν· καὶ τοῖς πνεύμασιν τοῖς ἀκαθάρτοις ἐπιτάσσει, καὶ ὑπακούουσιν αὐτῷ.

And they were all amazed, so that they questioned among themselves, saying, "What is this? A new teaching with authority! He commands even the unclean spirits, and they obey him." (Mark 1:27)

Գայ առ նա բորոտ մի, աղաչէր ի ծուր իջանէր՝ եւ ասէր. տէր՝ եթէ կամիս, կարող ես զիս բշչել:

Καὶ ἔρχεται πρὸς αὐτὸν λεπρὸς παρακαλῶν αὐτὸν [καὶ γονυπετῶν] καὶ λέγων αὐτῷ ὅτι ἐὰν θέλῃς δύνασαί με καθαρίσαι.

And a leper came to him, imploring him, and kneeling said to him, "If you will, you can make me clean." (Mark 1:40)

Եւ ընդ ելանելն նորա ի նաւէ անտի, պատահեաց նմա այր մի ի գերեզմանացն՝ զոր ուներ այս պիղծ:

καὶ ἐξελθόντος αὐτοῦ ἐκ τοῦ πλοίου εὐθύς ὑπήντησεν αὐτῷ ἐκ τῶν μνημείων ἄνθρωπος ἐν πνεύματι ἀκαθάρτῳ,

And when Jesus had stepped out of the boat, immediately there met him out of the tombs a man with an unclean spirit. (Mark 5:2)

Եւ կոչեաց առ ինքն գերկոտասանսն, եւ սկսաւ առաքել զնոսս երկուս երկուս. եւ տայր նոցա իշխանութիւն ի վերայ այսոց պիտոց:

Καὶ προσκαλεῖται τοὺς δώδεκα καὶ ἤρξατο αὐτοὺς ἀποστέλλειν δύο δύο καὶ ἐδίδου αὐτοῖς ἐξουσίαν τῶν πνευμάτων τῶν ἀκαθάρτων,

And he called the twelve and began to send them out two by two, and gave them authority over the unclean spirits. (Mark 6:7)

Իբրեւ ետես Յիսուս՝ եթէ կուտի ժողովորդն ի վերայ, սասուեաց այսոյն պիղծ՝ եւ ասէ. համրի՛ եւ խուլ այս՝ եւ տամ քեզ հրաման, ել ի դմանէ, եւ այլեւս՝ մի մտանիցես ի դա:

Ἴδὼν δὲ ὁ Ἰησοῦς ὅτι ἐπισυντρέχει ὄχλος, ἐπέτιμήσεν τῷ πνεύματι τῷ ἀκαθάρτῳ λέγων αὐτῷ· τὸ ἄλαλον καὶ κωφὸν πνεῦμα, ἐγὼ ἐπιτάσσω σοι, ἔξελθε ἐξ αὐτοῦ καὶ μηκέτι εἰσέλθῃς εἰς αὐτόν.

And when Jesus saw that a crowd came running together, he rebuked the unclean spirit, saying to it, "You mute and deaf spirit, I command you, come out of him and never enter him again." (Mark 9:24-25)

In the Gospel of Mark, the same pattern as in the Gospel of Matthew appears. Specific Greek terms stemming from *kat'arós* or the root *kat'* - are not translated with words based on the root *srb*- but with the Armenian term for “unclean/stained” (Arm. *pilc*). In five of the six examples in the Gospel of Mark where *surb* is not used as a translation for *akat'* -, the Armenian term *pilc* is applied. All of these examples are connected to evil or rather correctly unclean spirits (Grk. *pneûma* and Arm. *ays*).

Mark 1:40 includes an interesting exception where the Greek verb *kat'arísai* (“to clean, purify”) is not translated as Arm. *srbel* but as Arm. *bžškel*, a term that is closely connected to “healing.” Until today, a physician is called *bžišk* in modern Eastern Armenian.

In other parts of the Armenian translation of the Gospels, the verb *bžškel* is usually taken as a translation of the Greek term *t' erapeúein* (“to cure”), as in Matthew 1:10 or Luke 6:18. Even though the connection between Grk. *kat'* - and Arm. *bžšk*- rarely occurs in the Gospels, particularly when compared to the frequent connection between Grk. *akat'arós* and Arm. *plc*, it illustrates that the semantic fields of healing, purity, and *holy* were intertwined. It further indicates that this correlation was expressed with the same notions in the Armenian texts, namely the duality of *surb* and *pilc*.

The Gospel of Luke includes one verse where *surb* is used as a translation of a Greek term that is not related to *kat'arós* or *hágios*.

Սրբութեամբ եւ արդարութեամբ առաջի նորա՝ զամենայն ատուրս
լենաց մերոց:

ἐν ὁσιότητι καὶ δικαιοσύνῃ ἐνώπιον αὐτοῦ πάσαις ταῖς ἡμέραις
ἡμῶν.

in holiness and righteousness before him all our days. (Luke 1:75)

In this example, the Armenian noun *srbut'iwñ* is used to translate the Greek term *hosiós*, one of the three or four terms in Ancient Greek that are commonly associated with a semantic field of *holy*.²⁴

Additionally, there are several examples in the Gospel of Luke where words based on *kat'* - and *hag-* are translated with Armenian terms other than *surb*.

Իսկ եթէ ռուք որ շարքդ էք՝ գիտէք պարգևս բարիս տալ որդեց
ձերոց, որչափ եւս առաւել հայր ձեր յերկնից տացէ բարիս այն-
ցիկ, որ խնդրենն ի նմանէ:

εἰ οὖν ὑμεῖς πονηροὶ ὑπάρχοντες οἴδατε δόματα ἀγαθὰ διδόναι
τοῖς τέκνοις ὑμῶν, πόσω μᾶλλον ὁ πατήρ [ὁ] ἐξ οὐρανοῦ δώσει
πνεῦμα ἅγιον τοῖς αἰτοῦσιν αὐτόν.

If you then, who are evil, know how to give good gifts to your children, how much

24 See the section on the Greek terminology in Chapter 2 for more information.

more will the heavenly Father give the Holy Spirit to those who ask him!" (Luke 11:13)

Եւ էր ի ժողովրդեանս այր մի գոր ունէր՝ այս դիւի պիծոյ, եւ աղա-
ղակեաց ի ճայն մեծ՝

Καὶ ἐν τῇ συναγωγῇ ἦν ἄνθρωπος ἔχων πνεῦμα δαιμονίου ἀκαθάρτου καὶ ἀνέκραξεν φωνῇ μεγάλῃ·

And in the synagogue there was a man who had the spirit of an unclean demon, and he cried out with a loud voice (Luke 4:33)

Եւ եղէն զարմանալիք ի վերայ ամենեցուն. խօսելին ընդ միմեանս
եւ ասէին. զիչ է բանս այս, զի իշխանութեամբ եւ զօրութեամբ
սաստէ այսոց պիծոց՝ եւ ելանէն:

καὶ ἐγένετο θάμβος ἐπὶ πάντας καὶ συνελάλουν πρὸς ἀλλήλους λέγοντες· τίς ὁ λόγος οὗτος ὅτι ἐν ἐξουσίᾳ καὶ δυνάμει ἐπιτάσσει τοῖς ἀκαθάρτοις πνεύμασιν καὶ ἐξέρχονται;

And they were all amazed and said to one another, "What is this word? For with authority and power he commands the unclean spirits, and they come out!" (Luke 4:36)

Որ եկին լսել՝ ի նմանէ, եւ բժշկել ի հիւանդութենէ իրեանց. եւ
նշեցալքն յայսոց պիծոց բժշկէին:

οἱ ἦλθον ἀκοῦσαι αὐτοῦ καὶ ἰαθῆναι ἀπὸ τῶν νόσων αὐτῶν· καὶ οἱ ἐνοχλούμενοι ἀπὸ πνευμάτων ἀκαθάρτων ἐθεραπεύοντο,

who came to hear him and to be healed of their diseases. And those who were troubled with unclean spirits were cured. (Luke 6:18)

Քանզի հրամանս տայր այսոյն պիծոյ՝ ելանել ի մարդոյ անսի. զի
բարոյ ժամանակս յափշտակելալ էր զնա, եւ կապեր ի շրթայս՝ եւ
պահեր յերկաթս, եւ խզր զկապանսն. եւ վարեր ի դիւն յանա-
պատ:

παρήγγειλεν γὰρ τῷ πνεύματι τῷ ἀκαθάρτῳ ἐξελεθεῖν ἀπὸ τοῦ ἀνθρώπου. πολλοῖς γὰρ χρόνοις συνηρπάκει αὐτὸν καὶ ἐδεσμεύετο ἀλύσειν καὶ πέδαις φυλασσόμενος καὶ διαρρήσων τὰ δεσμὰ ἠλαύνετο ὑπὸ τοῦ δαιμονίου εἰς τὰς ἐρήμους.

For he had commanded the unclean spirit to come out of the man. (For many a time it had seized him. He was kept under guard and bound with chains and shackles, but he would break the bonds and be driven by the demon into the desert.) (Luke 8:29)

Եւ մինչ դեռ մատչէր, զարկոյց զնա դեւն, եւ շարժեաց: Սաստեաց
Յիսուս այսոյն պիծոյ, եւ բժշկեաց զմանկն. եւ ետ զնա ցհար
իւր.

ἔτι δὲ προσερχομένου αὐτοῦ ἔρρηξεν αὐτὸν τὸ δαιμόνιον καὶ συνεσπάραξεν· ἐπέτιμησεν δὲ ὁ Ἰησοῦς τῷ πνεύματι τῷ ἀκαθάρτῳ καὶ ἴασατο τὸν παῖδα καὶ ἀπέδωκεν αὐτὸν τῷ πατρὶ αὐτοῦ.

While he was coming, the demon threw him to the ground and convulsed him. But Jesus rebuked the unclean spirit and healed the boy, and gave him back to his father. (Luke 9:42-43)

With the single exception of Luke 11:13 where the “holy spirit” (Grk. *pneûma hágion*) is translated as Arm. *baris* (“good”), all the other examples follow the already familiar pattern by which the “unclean spirit” (Grk. *akat'arón pneûma*) is translated by Arm. *ays plcoc'* (“spirit of impurity/impure spirit”). It is also noteworthy that the Armenian verb *bžškel* (“to cure, heal”) appears in the direct proximity of *ays plc'* in the sense of “to heal/purify an unclean spirit.” This close relationship underlines once more the fact that *surb* was not only used as a translation of Greek words related to *holy* (such as *hágios*) or pure (*kat'arós*) but that it was closely linked to the semantic field of healing (Grk. *t' erapeúein*) as well.

The Gospel of John contains even more examples where the notion of *surb* was used as an equivalent for Greek words other than *hágios* or *kat'arós*.

Եւ Էր մերձ զառիկն Հրէից: Եւ բարսւմք ելին յԵրուսաղէմ ի գաւառէ աւստի յառաջագոյն քան զգառիկն, զի սրբսցեն զաւծիսս իրեանց:

Ἦν δὲ ἐγγὺς τὸ πάσχα τῶν Ἰουδαίων, καὶ ἀνέβησαν πολλοὶ εἰς Ἱεροσόλυμα ἐκ τῆς χώρας πρὸ τοῦ πάσχα ἵνα ἁγνίσωσιν ἑαυτοὺς.

Now the Passover of the Jews was at hand, and many went up from the country to Jerusalem before the Passover to purify themselves. (John 11:55)

The passage in John 11:55 exemplifies the use of *surb* as a verb for the translation of the Greek term *hagnós* (ἁγνός). *Hagnós* is yet another Greek term that is commonly associated with a semantic field of *holy* or purity. This observation is in line with the previous ones that *surb* was frequently used to translate different Greek terms related to either *holy* or purity. The term *hagiós*, however, is more widespread in the Gospels and in the context of *surb* as well.

Եւ ապա առեալ ջուր՝ արկ ի կոնք, եւ սկսաւ լուծալ զիսս աշակերտացն, եւ սրբէր դէշաւական զիր սիւսժեալ Էր:

εἶτα βάλλει ὕδωρ εἰς τὸν νιπτῆρα καὶ ἤρξατο νίπτειν τοὺς πόδας τῶν μαθητῶν καὶ ἐκμάσσειν τῷ λεντίῳ ᾧ ἦν διεζωσμένος.

Then he poured water into a basin and began to wash the disciples' feet and to wipe them with the towel that was wrapped around him. (John 13:5)

John 13:5 includes a rare example of the application of *surb* for the translation of the Greek verb *ekmássein* (“to wipe (off)” or “to wipe dry”). This passage can thus be regarded as one of the very few instances in the Armenian material that illustrate the link between a semantic field of physical (non-religious) purity and *surb*. This rare application of *surb* in the sense of physical purity (cleanliness) resonates with the –

probably older – more elementary understanding of *surb* conveyed in other parts of the Gospels (see discussion above) rather than with complex ethical, religious, or spiritual purity.²⁵ Matthew 27:59 provides another example of what might be an original signification of *surb*.

The remaining two passages in John 6:69 and John 15:2 are less relevant since they only include more examples of the translation of *kat'arós* as *surb* and one example (John 6:69) where the Greek term *hágios* is not translated in the Armenian text.

To summarize, the examination of the Gospels and a closer look at the relationship between *surb* and several Greek words commonly related to the notion of *holy* (such as *hágios*) and pure (such as *kat'arós*) have yielded the following results.

- There are other Armenian terms in addition to *surb* used to translate a specific Greek terminology commonly related to *holy* and pure. The first Armenian term is the verb բժշկել (*bžškel*), meaning “to cure, heal.” This term is often found in medical contexts in the Gospels where sicknesses such as leprosy need to be cured. The word *bžškel* was usually applied to translate the Greek verb *kat'arísai* (“to clean”), a term that was also commonly translated as Arm. *srbel*. This connection thus underlines the overlappings between the semantic fields of *surb*, purity, and healing that were only hardly visible in the other Armenian sources.
- Another relevant Armenian term that occurred in the context of *surb* in the Gospels was պիղծ (*pilc*). According to the standard dictionaries, the term *pilc* is linked to the semantic field of impurity and is commonly interpreted as the antonym of *surb*. Words based on *pilc* also appear in the other sources examined in this book, such as in the “History of the Armenians” by Agat'angelos (§143). In this account, *pilc* also functions as an attribute that stands for the opposite of *surb* and *yeraneli* (“blessed”).

Իսկ երանելին պարկեշտասերն Գայիանէ, սրբասնելովն Հռիփսիմեաւ հանդերձ եւ այլ ընկերաւքն իրեանց, յիշեալ զուխտն սրբութեան, զարինաւոր կրանիցն զգաստութեան սրբութիւնն յոր մտեալ էին՝ ողբային յանձինս իրեանց վասն նկարացոյց հրամանին, առ ի նկատել զպատկերս նոցա պիղծ եւ անաւրէն թագաւորին. (Agat'angelos §143)

But the blessed and chaste Gaiane, with the saintly Rhipsime and their other companions, remembered the covenant of holiness, the religious rule of chastity into which they had entered, and lamented amongst themselves over the impure and impious emperor's command to have their portraits painted.

25 See the discussions on the original meaning of *surb* in the previous part of this section.

- The word field around *pilc* and its relation to *surb* thus function similarly to the important *sacred* and *profane* distinction mentioned and discussed in Chapter 2. However, unlike the notion of *profane*, the word fields around *pilc* and its opposite term *surb* are more closely connected to the semantic field of purity. Despite this difference, the antagonistic use of *surb* and *pilc* proves that terminology of *separation-binding*, that was regarded as a crucial indication of the existence of a semantic field of *holy* (see also Chapter 2), existed in Armenian as well.
- The terminology related to *surb* was also applied to translate other Greek terms than *hágios* or *kat'arós*. Among others, *surb* was used for the translation of other Greek terms commonly related to *holy*, such as *hagnós* or *hosiós* and the verb *ekmássein* (“to wipe (off)”). The latter example is particularly interesting because it is a rare illustration of the relationship between *surb* and a basic understanding of purity in the sense of physical cleanliness.

5.2 English Corpora

The following part focuses on the examination of the *Holy/Sacred English Corpus* (HSEC) and the results of the *Sketch Engine* queries.

Starting with the examination of the HSEC, the corpus linguistic methods applied on the macrolevel are the same as those used in the previous study of the Armenian corpora. However, the HSEC examinations do not always include a keyword analysis because of the lack of an appropriate reference corpus. The examination of the English corpora will be based on “types” only due to two main reasons.

- Firstly, there are only small variations between “lemmas” and “types” in the English language.
- Secondly, the Twitter data is highly idiosyncratic and includes many misspellings; therefore, a potential lemmatization would only produce faulty results.

Before starting with the examination of the subcorpora of the HSEC, the first step is to investigate the similarities and differences in the use of the two lexical units “holy” and “sacred.” As mentioned in the introduction of this book, based on the secondary literature and given their common use in the study of religion, it is generally assumed that both words belong to a shared notion of *holy*. The sources in the HSEC, however, include the use of both terms. In some sources, the use of “sacred” is dominant (such as in HSEC *Eliade* and HSEC *Encyclopaedia Britannica*). In other sources, the focus lies on the word “holy” or even on other terms such as “numinous” (for instance, HSEC *Otto*). To verify that all these words indeed refer

to a common notion of *holy*, the first part of this chapter will search for potential differences in the use and meaning of “sacred” and “holy.”

5.2.1 Excursus: Holy vs. Sacred – Twitter Data

The following analysis is based on two different text corpora.

- The first corpus is the larger version of the HSEC Twitter Data introduced in Chapter 4. This version of the Twitter data includes 23,720 tweets (the Twitter data used in the examination below is a subset of this larger Twitter data). Among these tweets are 22,856 tweets (around 96% of the data) that contain the words “holy” or “holiness” and 939 tweets that mention “sacred.” Duplicates of the plain texts and the cleaned texts of the tweets have been removed from the data.²⁶ For the analysis of the use of “sacred” and “holy,” the Twitter data has not been classified into religious or non-religious tweets.
- The second corpus is the *English Web 2015 (enTenTen)* corpus that was used and queried with *Sketch Engine*.

The main examination methods to compare the use and meaning of “sacred” and “holy” are:

- a comparison of the most frequently used hashtags in the tweets including “holy” and those including “sacred,”
- the examination of absolute frequency word lists (no stop words included) of both subcorpora,
- a collocation analysis (6-LL, L3-R3; no function words),
- and a microanalysis of selected examples where “sacred” and “holy” appear together.

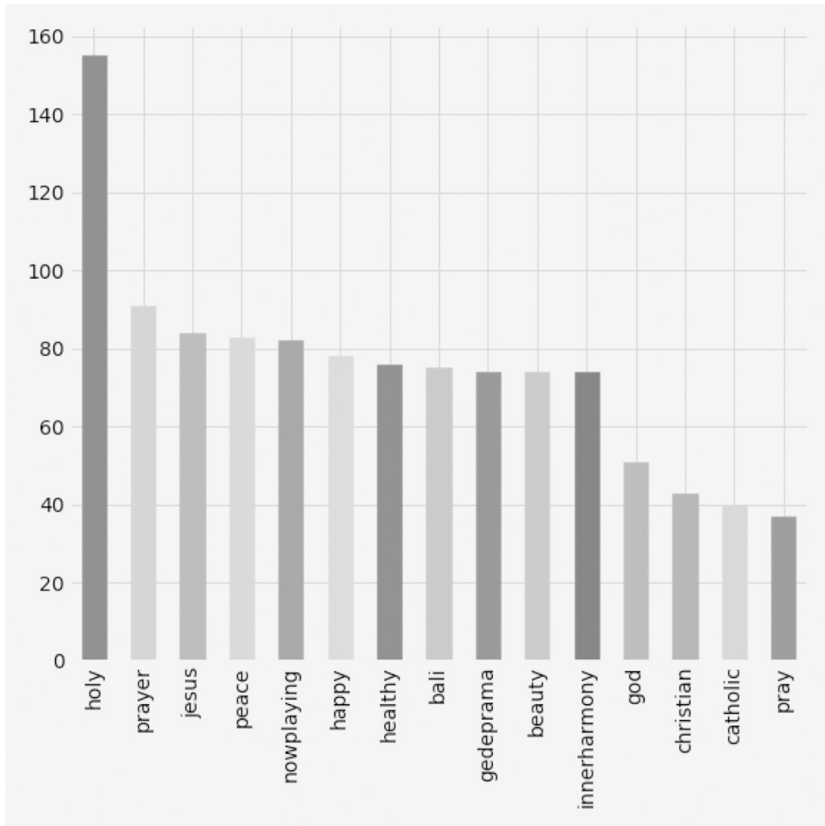
Most of these examinations were conducted with the help of *LancsBox*.

5.2.1.1 Tweets including Holy/Holiness – Hashtags

The fifteen most frequently used hashtags (see the corresponding figure 36) in tweets that include “holy” (or “holiness”) can be categorized in the following way.

26 Unfortunately, there might still be a lot of noisy data in the form of duplicates. For instance, some of the tweets have been retweeted and reposted over a couple of days or even weeks in almost the same form. Usually, the (automated) cleaning part of the data includes the dropping of duplicates. However, as some of these tweets differ in minor details (e.g. the date) from their preceding tweets, they are often not recognized as duplicates in the automated cleaning part and consequently not removed from the data.

Figure 36: 15 most common #hashtags in tweets including “holy/holiness.”



The hashtag #holy will not be considered in this categorization. The overall frequency (including #holy) is 1,117.

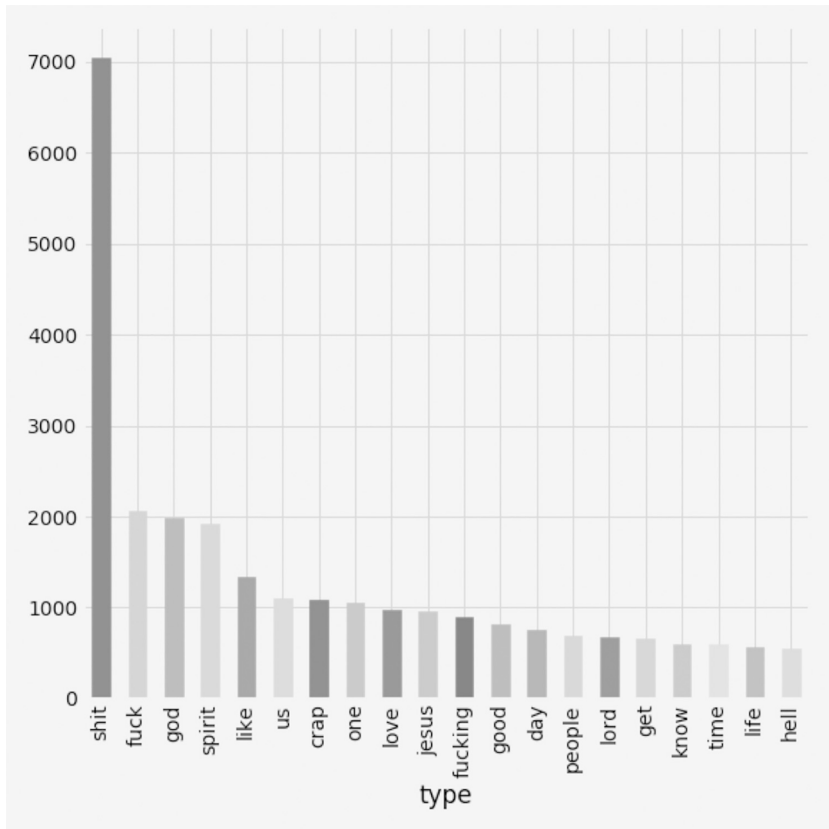
- Hashtags with a religious connotation (#prayer, #jesus, #christian, #god, #catholic, #pray) (AF 346, RF 31%)
- Hashtags with a non-religious connotation (including religioids) (#nowplaying, #peace, #happy, #healthy, #bali, #beauty, #innerharmony, #beauty, #gedeprema) (AF 616, RF 50%)

The majority of hashtags in the Twitter data with tweets that include “holy” and “holiness” are non-religious.²⁷ Many of the hashtags related to non-religious topics are connected to positively connotated terms that represent feelings or “states of mind,” such as peace, beauty, health, harmony, etc.

However, the two most widespread hashtags have a religious connotation (the first three if including #holy and interpreting it religiously).

5.2.1.2 Tweets including Holy/Holiness – Word Frequency List

Figure 37: Absolute frequency of words (no stop words) in tweets that include “holy” or “holiness” (top 20 excluding “holy”).



27 Note once more that the hashtag #holy was excluded from the analysis. Had the hashtag #holy been included in the “religious” section, the numbers would have been comparable.

Table 40: Absolute frequency (AF) table based on tweets that include “holy/holiness” (top 20 excluding “holy”).

Word	AF
shit	7039
fuck	2064
god	1977
spirit	1920
like	1339
us	1105
crap	1091
one	1051
love	972
jesus	954
fucking	889
good	820
day	760
people	695
lord	675
get	664
know	605
time	600
life	563
hell	552

There is a clear distinction between different groups that is fairly similar to the groups found in the examination of the HSEC Twitter Data below (the overall frequency of words in this sample is 26,335). The visualization is displayed in the corresponding figure 37.

- Words related to “Impurity” (shit, fuck, crap, fucking) (AF 9,019, RF 34%)
- Words related to “Religion” (god, spirit, jesus, hell, lord) (AF 6,078, RF 23%)
- Words related to (mostly positive) non-religious “Concepts” (good, love, life, one) (AF 3,406, RF 13%)
- Words related to “Persons” and “Titles” (people, lord) (AF 1,370, RF 5%)
- Words related to “Time” (day, time) (AF 1,360, RF 5%)
- Others, mostly verbs and adverbs (like, us, get, know, etc.)

Words related to religion only represent approximately 25% of the whole data. In fact, the large group of words that are used as insults and which are often related to notions of impurity outnumbers the religious use of *holy*. Since this phenomenon is also attested in the smaller HSEC Twitter Data sample, it will not be discussed further at this stage.

5.2.1.3 Tweets including Holy/Holiness – Collocation Analysis (6-LL, L3-R3, C100-NC100; no function words)

Table 41: Top 15 words collocation analysis (6-LL, 3L-3R, C100-NC100; no function words) HSEC Twitter Data large (“holy”).

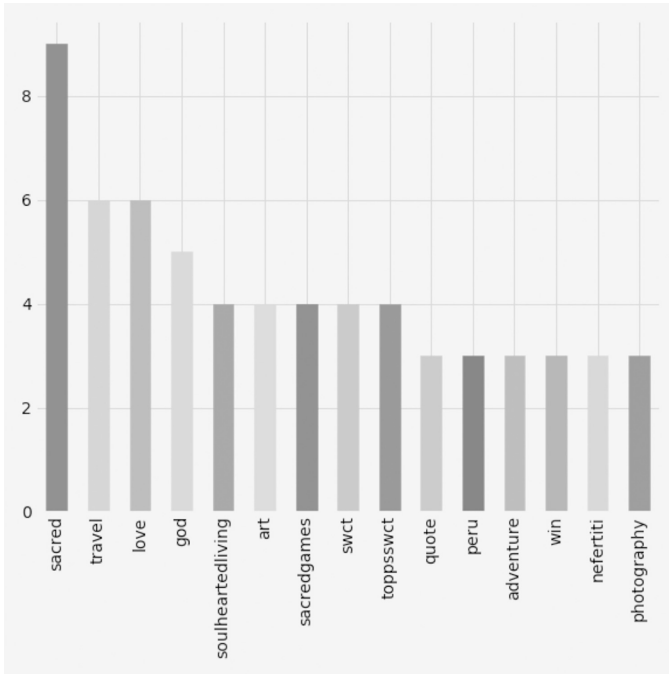
id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	shit	61475.86	9428	7039
2	r	fuck	15687.13	2594	2064
3	r	holy	14497.61	7155	23019
4	r	spirit	10848.82	1943	1920
5	r	crap	8932.95	1459	1091
6	r	cow	3561.33	586	440
7	r	hell	3180.83	571	552
8	r	moly	3127.64	508	366
9	r	fucking	3041.19	710	889
10	r	trinity	2575.04	447	392
11	r	ghost	2508.80	432	370
12	r	grail	2293.75	402	362
13	r	water	2085.76	377	370
14	r	god	1523.69	717	1977
15	r	smokes	1396.22	229	169

The overall collocation frequency in the collocation table is 27,558. The visualization of the collocation analysis can be found in the corresponding figure 38. The entries in the table from the collocation analysis can be categorized in the following way.

- Words related to “Impurity” (shit, fuck, crap, fucking) (AF 14,191, RF 52%)
- Words related to “Religion” (holy, spirit, hell, trinity, ghost, grail, god) (AF 11,667, RF 42%)
- Words related to “Objects” (cow, water, smokes) (AF 963, RF 4%)
- Misc (moly) (AF 508, RF 2%)

5.2.1.4 Tweets including Sacred – Hashtags

Figure 39: 15 most common #hashtags in tweets including “sacred.”



The fifteen most common hashtags (see the corresponding figure 39) that are used in tweets that mention “sacred” can be categorized in the following way. The absolute frequencies are much lower compared to the numbers in the examination of tweets that include “holy.” This is due to the fact that the sample itself is much smaller (see Chapter 3 and the introduction of this section). Similarly to the approach in the section above, the hashtag “sacred” is left out in the following categorization. The overall frequency is 64 (including the hashtag #sacred).

- Hashtags with a religious connotation (#god) (AF 5, RF 8%)
- Hashtags with a non-religious connotation (the rest except for #sacred) (AF 50, RF 78%)

There is an overwhelming number of hashtags used in tweets that include “sacred” which are not related to religion. It is noteworthy that many of the non-religious hashtags are linked to traveling, travel locations, and journeys (for instance, #travel,

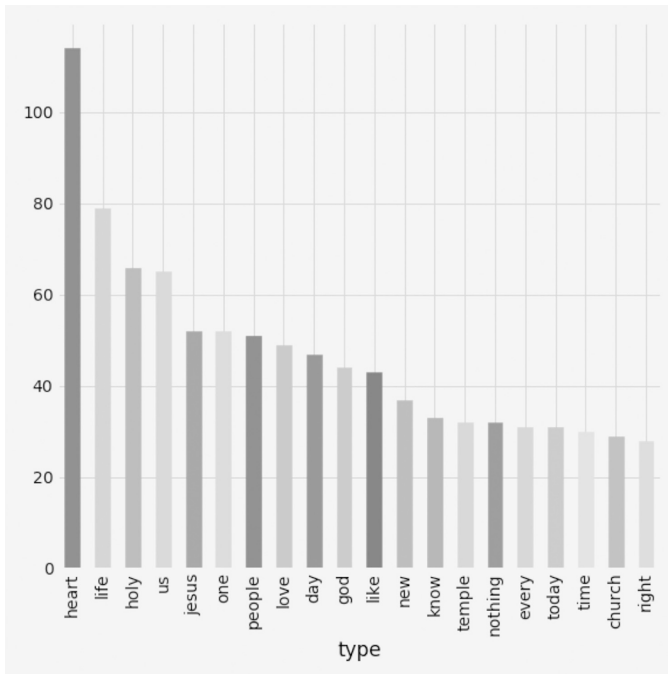
#peru, #adventure, #photography, etc.). This might express the importance of the connection between “sacred” and places. A significant number of tweets are connected to pop culture and video games (such as Star Wars in #swct). A very interesting hashtag is the hashtag #art, which might hint at a sacralization process in the form of *Kunstreligion*, a phenomenon briefly discussed in Chapter 2.

Be that as it may, the top hashtags indicate that the tweets that include “holy” or “holiness” have stronger religious affiliations than the tweets mentioning “sacred.”

5.2.1.5 Tweets including Sacred – Word Frequency List

The visualization of the word frequencies is found in the corresponding figure 40.

Figure 40: Absolute frequency of words (no stop words) in tweets that include “sacred” (top 20).



The following groups can be formed from the entries in the word frequency list (overall frequency: 945):

- Words related to (positive) non-religious “Concepts” (life, heart, new, one, love, nothing, right, every) (AF 422, RF 44.7%)
- Words related to “Religion” (holy, god, jesus, church, temple) (AF 223, RF 23.6%)
- Words related to “Persons” (people, us) (AF 116, RF 12.3%)
- Words related to “Time” (day, time, today) (AF 108, RF 11.4%)
- Other words (know, like) (AF 76, RF 8%)

The Twitter data word frequency lists shows that there seem to be more similarities than differences between the application and meaning of “holy” and “sacred.”

For instance, the positive concepts mentioned in both groups as well as the religious terminology are relatively similar and also seem to have an equal distribution. However, “sacred” appears to be less frequently associated with persons and notions of “impurity.” Instead, the word “sacred” is more frequently linked to places such as “temples” and “church” as well as “time.” Interestingly “holy” appears among the most frequent words that occur in tweets that mention “sacred.” This observation, as well as the considerable overlap between the most frequent groups, strengthens the initial hypothesis that both notions have a similar function despite all their differences. They thus form a common notion of *holy* as defined in this book.

Table 42: Absolute frequency (AF) table based on tweets that include “sacred” (top 20).

Word	AF
heart	114
life	79
holy	66
us	65
jesus	52
one	52
people	51
love	49
day	47
god	44
like	43
new	37

know	33
temple	32
nothing	32
every	31
today	31
time	30
church	29
right	28

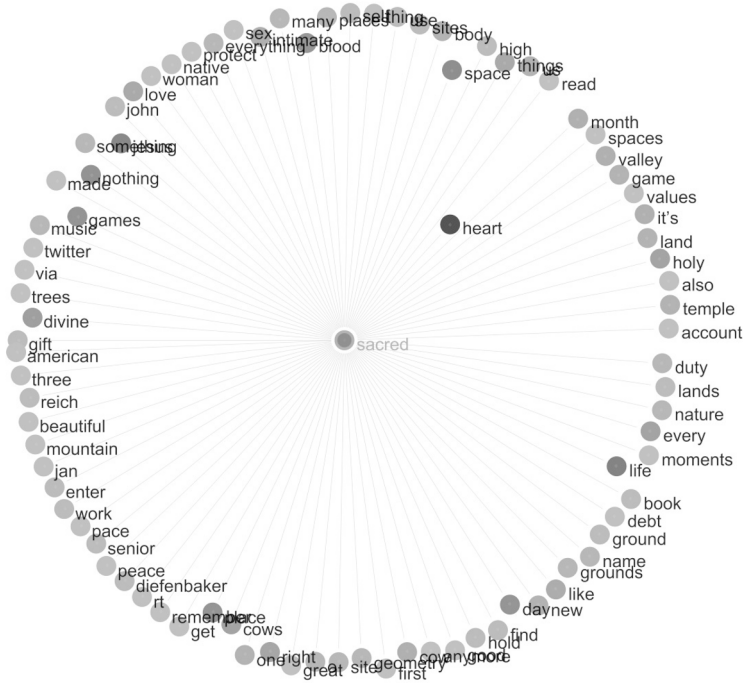
5.2.1.6 Tweets including Sacred – Collocation Analysis (6-LL, L3-R3, C5-NC5; no function words)

Table 43: Top 15 words collocation analysis (6-LL, 3L-3R, C5-NC5; no function words)
HSEC Twitter Data large (“sacred”).

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	heart	545.39	98	114
2	r	space	131.38	24	28
3	r	games	120.31	20	21
4	r	jesus	109.97	28	52
5	r	blood	104.55	19	22
6	l	life	102.77	32	79
7	r	place	95.42	20	28
8	l	nothing	94.42	21	32
9	r	cows	85.36	13	12
10	r	day	66.37	20	47
11	r	divine	63.29	15	25
12	l	intimate	57.62	9	9
13	r	cow	47.35	9	11
14	r	geometry	46.96	7	6
15	r	duty	44.80	7	7

The overall collocation frequency in the collocation table is 342. The visualization of the collocation analysis can be found in the corresponding figure 41. The entries in the table from the collocation analysis can be categorized in the following way.

Figure 41: Top 15 words collocation analysis (6-LL, L3-R3, C5-NC5; no function words) HSEC Twitter Data large (“sacred”) (graph) (c) LancsBox.



- Words related to non-religious “Concepts” (heart, games, life, nothing, geometry, duty) (AF 185, RF 54.1%)
- Words related to “Places” (space, place) (AF 44, RF 12.9%)
- Words related to “Religion” (jesus, divine) (AF 43, RF 12.6%)
- Words related to “Objects” (cow(s), blood) (AF 41, RF 12%)
- Words related to “Time” (day) (AF 20, RF 5.9%)
- Other words (intimate) (AF 9, RF 2.6%)

The collocation analysis illustrates that the term “sacred” is rarely used in the proximity of words that are related to religion. Most of the words found in the collocation analysis are various concepts, among them “heart”²⁹ and “games.” A notice-

29 The “heart” is frequently used together with “sacred” in place names or institutional names, such as hospitals or churches.

able difference from the collocation analysis of “holy” is the presence of the “space” and “time” categories. The observation that some of the words that appear in the “concepts” category (such as “heart”) are also often used as names for places or institutions (hospitals, churches, etc.) further stresses their importance.

If we consider that most of the words in the “holy” collocation analysis are used together with “holy” as fixed expressions or exclamations in non-religious contexts, it can be concluded that the everyday use of “sacred” still has a strong religious (or, at least, religioïd) potential. It must, however, be said that words directly related to religion are only rarely found in its direct vicinity. For example, place names such as “Sacred Heart” (for a hospital) or expressions such as “life is sacred” match the concept of religioïds discussed in Chapter 2. They are thus also more likely to regain a religious meaning than most of the expressions seen in the collocation analysis of “holy.” Yet, it should be recalled that the sample with tweets including “sacred” is much smaller than the sample with tweets including “holy.”

5.2.1.7 Holy/Sacred (Microanalysis)

The final part of the Holy/Sacred analysis of the Twitter data discusses tweet examples where “holy” and “sacred” are used together. The examples all stem from the collocation analysis of “sacred.” Even though “holy” is only ranked 93 in the collocation analysis with a score of 22.16 (collocation frequency 13 and overall frequency 66), the thirteen collocates still provide a reasonable basis for the analysis. Unlike the microanalyses in the other sections, these thirteen sentences will be briefly discussed all at once.³⁰

This short discussion should offer the reader an overview of some everyday use cases of *holy* and allow to assess whether the terminologies of “sacred” and “holy” are used synonymously (while referring to a common semantic field of *holy*). The following tweets are based on the cleaned version of the tweet texts. Thus, smilies, links, most of the punctuation, etc. are missing.

- “Each day is sacred each day is holy some days the sacred and the holy are easy to see some days it feels as though the sacred and the holy are no where to be found but the truth is we are.”
- “I desire to experience god’s presence in a special way today so I will be engaging in some holy sacred flirting heeey god yes you god you are my god yes mine and there is none like you nothing compares to you you my god take my breath away.”
- “Why do loads of arabs make loads of things holy and sacred it is terrifying.”

30 Some sentences are left out because they are either duplicates of other tweets except that a few words were changed or because they are too noisy.

- “Pakistan may as well just amend the blasphemy law to add cjp name to the list of sacred holy personalities that no one should criticise or question.”
- “jesus christ head in thorns religious sacred holy cross medal vatican church god.”
- “Ja ja hail germany against galaxiusmons crimes or in your words universal human rights yeah make sure you’s let countries know brief federation upon returning to the holy sacred vaterland herr sir.”
- “Let the lord bless us on this most sacred most holy of days: the day taylor swift’s reputation stadium tour goes onto netflix.”
- “India is the sacred land of holy rishis munis and divine incarnations like sant shri asaram babu ji we should learn from their life to respect parents.”
- “There is a possibility that sagittarians may go out on a pilgrimage to a very holy and sacred place close to them.”

A closer look at the sentences above reveals the following:

- The terms “sacred” and “holy” are very often used together as attributes for the same words. For instance, they are used to define “day,” “flirting,” “personality,” “vaterland,” and “place.” This is worth mentioning since “places,” as well as “time,” seemed to be more common in the context of “sacred” in the macroanalysis. The sentences make clear that the word “holy” can also be used for “places” and “time;” however, the term “sacred” seems to be used more frequently in these contexts.
- The selected examples also show that the terminology of “holy” and “sacred” is not restricted to either non-religious or religious use cases in the context of Christianity. Instead, it is linked to a greater variety of religious traditions (Hinduism, Islam, esotericism). Even though Christian traditions are still very present among the top entries in the macroanalysis, the middle-placed entries highlight the importance of other religious traditions (where words such as “quran” regularly appear). The application of *holy* in the context of non-Christian religions in the Twitter data is an interesting phenomenon in light of the debate of the influence of applications of Western notions in non-Western religions.³¹ The results of the Twitter analysis could be interpreted as the outcome of acculturation processes or, rather, as the demonstration that ideas of *holy* existed in non-Christian religions. From the perspective of the believers, these ideas correspond to the English notion of *holy*.

31 As an example, see the debates in the field of postcolonial studies and Talal Asad’s critique of Western concepts of secularism Asad (1997, 2003); Dhawan and Varela (2015, 54ff.).

- It is often difficult to determine whether a tweet is written from an inner-religious viewpoint or whether the topos of religion (evoked by using its terminology) is instead used for other purposes, for instance, irony, politics, evoking a particular atmosphere, etc. A telling example in this context is the second tweet in the list.

Overall, the parallel use of “holy” and “sacred” in the analysis of the Twitter data supports the assumption that the meanings of both words overlap in many respects and that they relate to a common semantic field of *holy*. Nevertheless, there are significant differences in their use cases as well. For instance, the word “sacred” is more frequently used in the context of places and time, whereas “holy” is often applied together with notions of impurity in fixed expressions, such as “holy hell” or “holy shit.” However, when both words co-occur, their common function as attributes referring to a semantic field of *holy* is evident.

A more extensive data set (namely the *English Web 2015* corpus) will now be queried with the help of *Sketch Engine* to assess this preliminary conclusion.

5.2.2 Excursus: Holy vs. Sacred – Sketch Engine

Sketch Engine’s “Word Sketch” feature is mostly applied in the following analysis of the use of “sacred” and “holy.” The “Word Sketch” option offers a comprehensive overview of the full results of the *Sketch Engine* queries.

For instance, it allows searching for the most common verbs, nouns, etc. that are frequently used in the context of the search term. Furthermore, the “Word Sketch Difference” feature facilitates the comparison of the results of one query to the results of a second query. The latter is very useful when looking for differences and similarities in the use of “sacred” and “holy.” The documentation on the official *Sketch Engine* website provides more information about the “Word Sketch (Difference)” feature(s).

The *Sketch Engine*’s “Word Sketch” feature offers a detailed overview of several characteristics of the search term and its use in the selected corpus (in this case, *English Web 2015* corpus with over 15 billion words). For this examination, only a few selected columns from the overall “Word Sketch” results are used. The selected columns are listed in the following overview, including their abbreviations that are used in the tables below.

For the nouns, the following columns are used:

- “possessors of holy/sacred” (N1)
- “holy/sacred is a ...” (N2)
- “holy/sacred and/or” (N3)
- “nouns modified by holy/sacred” (N4)
- “verbs with holy/sacred as object” (N5)

For the adjectives, the following columns are used:

- “subjects of ‘by holy/sacred’” (A1)
- “verbs complemented by holy/sacred” (A2)
- “modifiers of holy/sacred” (A3)
- “holy/sacred and/or” (A4)
- “nouns modified by holy/sacred” (A5)

5.2.2.1 The Use of Holy in the English Web 2015 (enTenTen15) Corpus - Word Sketch Query

The top ten results in each section of the *Sketch Engine* query are displayed in the following table. Neither the frequencies of the collocates nor the scores (here: based on *LogDice*) are listed in the table. *Sketch Engine*’s *English Web 2015* corpus search yielded 696,536 appearances of “holy as a noun” and 316,798 appearances of “holy as an adjective.” This result, however, is misleading, because *Sketch Engine* interprets appearances of “holy” with a capital letter (such as “Holy See”) as nouns. This is only partly correct but does not have to concern the value of the *Sketch Engine* results for the topic of this book. This book is interested in the general use of the terminology, regardless of whether it is used as an attribute, as a noun, as part of a fixed expression, or as a verb.

In the following, the results of the first query (“holy”) are displayed and divided into two tables. The first table represents “holy as a noun” and the second “holy as an adjective.” The full column names are provided in the list above.

Table 44: Sketch Engine “Word Sketch” search for “holy” (as noun) top results (selection; top 10 collocates; no scores/frequencies displayed).

N1	N2	N3	N4	N5
temple	almighty	holy	spirit	anoint
lord	lord	catholic	communion	disown
christ	name	synod	ghost	petition
god	god	righteous	see	despise
israel	light	trinity	trinity	tend
-	something	cavour	father	sanctify
-	-	sacred	land	desecrate
-	-	almighty	grail	cry
-	-	r-cheney	cross	unleash
-	-	apostolic	quran	encounter

Table 45: Sketch Engine “Word Sketch” search for “holy” (as adjective) top results (selection; top 10 collocates; no scores/frequencies displayed).

A1	A2	A3	A4	A5
god	present	thrice	righteous	grail
sabbath	declare	infinitely	holy	crap
jerusalem	call	ineffably	thy	shit
commandment	live	immutably	sacred	scripture
ye	keep	intrinsically	muslim	shrine
cow	consider	eminently	pure	cow
temple	show	am	bible	relic
thy	make	supremely	catholic	communion
o	visit	perfectly	blameless	trinity
mary	see	altogether	islamic	quran

5.2.2.2 The Use of Sacred in the English Web 2015 (enTenTen15) Corpus - Word Sketch Query

Using Sketch Engine’s “Word Sketch” search engine, the word “sacred” is estimated to be used as an adjective 288,080 times and 124,635 times as a noun. Similarly to the examination above, the top ten results of the Sketch Engine query are displayed in the following table. Neither the frequencies of the collocates nor the scores (here:

LogDice) are shown in the tables. The full column names are provided in the list above.

Table 46: Sketch Engine “Word Sketch” search for “sacred” (as noun) top results (selection; top 10 collocates; no scores/frequencies displayed).

N1	N2	N3	N4	N5
titian	sperm	profane	heart	re-imagining
nothing	facts	scared	scripture	reinvent
-	salmon	secular	harp	globalize
-	children	ecofeminism	scriptures	figure
-	hills	sequential	texts	relocate
-	nothing	eliade	sites	realise
-	color	majesty	liturgy	recover
-	earth	feminine	stones	encounter
-	child	holy	places	reclaim
-	body	sacred	geometry	sense

Table 47: Sketch Engine “Word Sketch” search for “sacred” (as adjective) top results (selection; top 10 collocates; no scores/frequencies displayed).

A1	A2	A3	A4	A5
cow	hold	indescribably	secular	cow
mountain	consider	thrice	profane	scripture
life	destroy	supremely	holy	text
earth	touch	intrinsically	choral	geometry
marriage	regard	something	inviolable	grove
land	declare	nothing	hindu	relic
temple	protect	doubly	solemn	ritual
nothing	call	inherently	invisible	mountain
creation	deem	religiously	ancient	duty
tree	contain	supposedly	divine	rite

5.2.2.3 Holy/Sacred with Sketch Engine's Word Difference Tool – A Comparison

It is interesting to note that many of the words that appear in the *Sketch Engine* queries also occurred in the previous macroanalyses of the Twitter data. For example, the connections between “holy” and “spirit,” “sacred” and “heart,” or “sacred” and spatial terms are present in the *Sketch Engine* results as well as in the tables of the Twitter data examined in the previous section.

However, there are also some noticeable differences. For instance, the very prominent connection between “holy” and words related to “impurity” (Twitter data) is hardly reflected in the *Sketch Engine* results. Yet, this is not very surprising since this relation is mostly encountered in colloquial expressions in everyday communication. This is a layer that only constitutes a small subset of the overall structure of the *English Web 2015 (enTenTen15)* corpus.

The tendency that “sacred” is frequently used with words not directly related to religion (such as “scripture,” “cow,” “life,” “earth,” “tree,” or “mountain”) whereas “holy” is often found in the context of words closely related to religion (“god,” “sabbath,” “Christ,” or “lord”) is also present in the *Sketch Engine* results. This impression is supported by the “x is a (...)” (N2) column where “holy” is related to terms such as “name,” “god,” “light” whereas “sacred” is mostly connected to material objects such as “sperm,” “earth,” “hill,” or “body.”

Additional differences regarding the use cases of “holy” and “sacred” can be further examined with the help of *Sketch Engine*'s “Word Sketch Difference” feature, which traces the differences between the use cases of two different words.

The aforementioned observation that many core religious words are commonly used together with “holy” (such as “spirit,” “catholic,” or “god”) whereas “sacred” frequently occurs in the context of non-religious materials or spatial terms is also revealed by the “Word Sketch Difference” query. Also worth mentioning is the fact that antonyms of the notion of *holy* that are crucial for the *separation-binding* dimension mentioned in Chapter 2 (for instance, “profane” or “secular”) are almost exclusively found in the context of “sacred” in the *Sketch Engine* queries. Thus, the use of “sacred” seems to imply a more ambivalent quality compared to the use of “holy.”

Despite the differences in the use cases of “holy” and “sacred,” the fact that both words are often found in each other's context of one another underlines that they can rightfully be subsumed under a common semantic field of *holy*.

5.2.1. Summary Holy vs. Sacred

This short digression on the differences and similarities of the use of “sacred” and “holy” has raised several ambiguities.

On the one hand, there are apparent differences in the applications of “sacred” and “holy” in both the *Sketch Engine* queries and in the Twitter data analysis. Depending on the genre of the text, the word “holy” is frequently found in the context of words related to impurity and in the context of words pertaining to religion (“god,” “catholic,” or “spirit”). The term “sacred,” however, is more frequently (though not exclusively) connected to spatial and material terms and to words that are supposed to denote the opposite of *holy*, such as “secular” or “profane.”

On the other hand, the frequent use of “holy” in expressions and exclamations such as “holy cow” has shown that the term “holy” is regularly used in contexts that are neither religious nor religioid. The term “sacred,” however, is often used in situations that can be interpreted as religioid in a strict sense. Some examples are terms such as “sperm,” “children,” or “creation.” In these debates related to controversial discussions such as sexual preservation and the “creation” of humans, the “sacred” functions more obviously as a religioid in the sense of a *Diskurstopper* (see Chapter 2) compared to many of the non-religious use cases of “holy” mentioned above.

Regardless of the differences, there are strong indications that both terms belong to the same field of *holy*. This assumption is not only based on the macro-analysis, where both notions often appear in the context of each other. It is also based on the microanalysis of the tweets, where both terms were almost used interchangeably.

Therefore, the initial assumption that both terms belong to the same semantic field of *holy* can be regarded as accurate. Consequently, including sources that focus either on “sacred” (such as HSEC Eliade) or on “holy” (such as HSEC Otto) to research a more abstract semantic field of *holy* is a fruitful approach and will be adopted in the following sections.

5.2.2. HSEC Twitter Data (Full)

The examination of the HSEC data starts with the HSEC Twitter Data. The Twitter data can be subdivided into:

- the full Twitter data (including the religious and the non-religious tweets),³²
- the religious Twitter data,³³
- and the non-religious Twitter data.

32 As described in Chapter 3 on the methodology, the full Twitter data means in the context of the examination in this part the smaller sample. The larger sample was only applied in the analysis of the use of “holy” and “sacred.”

33 Regarding the classification of the data, please see Chapter 4.

This section will be divided into three parts as well. The focus of the analysis lies in the examination of the separated religious and non-religious Twitter data. Therefore, the microanalysis of example sentences will only be part of these examinations and not of the full Twitter data set.

5.2.2.1. Word Frequency List (AF)

The 15 most common words (“types”) in the subcorpus HSEC Twitter Data (full corpus; no stop words) are shown in following table and bar plot (see the corresponding figure 42).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. Among the fifteen most common words are (no *holy*; overall frequency: 11,166):³⁴

- Words related to non-religious “Concepts” with a strong focus on impurity (fuck, shit, crap, fucking) (AF 5,848, RF₁ 3%, RF₂ 52.4%).³⁵
- Words related to religious “Concepts” (God, spirit, jesus, lord) (AF 2,633, RF₁ 1.4%, RF₂ 23.6%)
- Words related to other non-religious “Concepts” (love, good, one) (AF 1,391, RF₁ 0.7%, RF₂ 12,5%)
- Verbs such as “to get” or “to know” (AF 614, RF₁ 0.3%, RF₂ 5.5%)
- Words related to “Time” (day, one) (AF 359, RF₁ 0.2%, RF₂ 3.2%)
- Words related to “Persons” (AF 321, RF₁ 0.2%, RF₂ 2.9%)

The use of “holy” in the Twitter data includes two entirely different associative contexts.

34 Note that many words that usually appear in the frequency list such as “holy” or “like” have been removed.

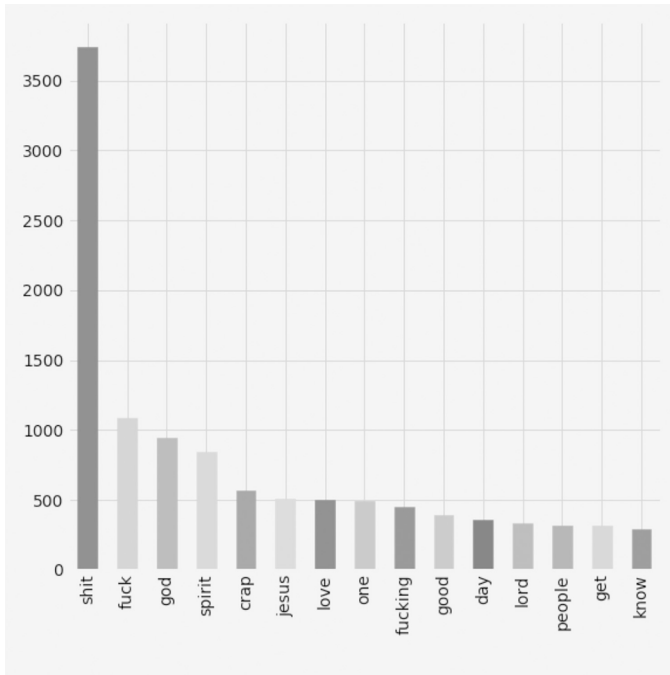
35 The relative frequency RF₁ is based on the total number of running words in the HSEC Twitter Data, including stop words (194,987 words). The second relative frequency RF₂ is calculated by “absolute frequency of the word” divided by “sum of all absolute frequencies in the top 15 table.”

Table 48: Most frequent words (top 15) in HSEC Twitter Data (full).

Type	Frequency
shit	3736
fuck	1089
god	947
spirit	843
crap	570
jesus	512
love	502
one	493
fucking	453
good	396
day	359
lord	331
people	321
get	320
know	294

On the one hand, there is a high frequency of words related to impurity, such as “shit,” “fuck,” and “crap.” Many of these words are either connected with excrements or sexuality. On the other hand, several terms have a religious connotation (“god,” “Jesus,” “lord,” “spirit,” etc.) or a more positive meaning (“love,” “good,” etc.). Individual persons are only barely mentioned (except for Jesus), and titles are entirely absent (except for “lord” which is a synonym for god). This indicates that the contemporary terminology of *holy* (except for “saint”) in the context of individual names is relatively unimportant. Instead, the words connected to *holy* are regularly applied to abstract concepts such as “love” and a very generic religious terminology (“god”). However, this hypothesis needs to be examined in light of the other English corpora as well.

Figure 42: Most frequent words (top 15) and their absolute frequencies in HSEC Twitter Data (full).



In order to find out if these different words also indicate various use cases of “holy,” the Twitter data is further classified into two distinct categories (“religious” and “not religious”), which will be analyzed in more detail after this section. The classification as “religious” or “not religious” has been made based on a machine learning algorithm.³⁶ The examination of the classified data will be discussed after this section on the full Twitter data.

5.2.2.2. Collocation Analysis

The top 15 entries of the HSEC Twitter Data (full) collocation analysis of the word “holy” (6-LL, L3-R3, C45-NC45; function words removed) can be found in the following table. The visualization is displayed in the corresponding figure 43.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic

36 For more information, see the machine learning excursus in Chapter 4.

tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 14,165).

Table 49: Top 15 words collocation analysis “holy” (6-LL, L3-R3, C45-NC45; function words removed) in HSEC Twitter Data (full).

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	shit	32914.1	5013	3736
2	r	fuck	8520.89	1396	1089
3	-	holy	7903.97	3760	11783
4	r	crap	4763.44	773	570
5	r	spirit	4714.77	847	843
6	r	cow	1782.98	293	220
7	r	hell	1718.95	307	294
8	r	moly	1649.84	269	197
9	r	fucking	1458.21	350	453
10	r	ghost	1369.19	234	197
11	r	trinity	1278.97	221	192
12	r	water	1139.95	206	203
13	r	grail	1044.98	185	172
14	r	communion	877.59	145	110
15	r	mother	826.87	166	178

- Words related to non-religious “Concepts” of impurity/sexuality” (shit, fuck, crap, fucking) (AF 7,532, RF 53.2%)
- Words related to religious “Concepts” and “Objects” (spirit, ghost, grail, trinity, hell, holy) (AF 5,699, RF 40.2%)
- Words related to non-religious “Objects” (water, cow) (AF 499, RF 3.5%)
- Words related to non-religious “Concepts” (mother, moly) (AF 435, RF 3.1%)

The high frequency of words related to impurity and religion in the word frequency lists can also be observed in the collocation analysis. Yet, there are some minor differences. For instance, the collocation analysis results also include objects such as “water” or “cow.” Some of the terms that appear in the “object” domain, such as “cow,” are usually part of fixed expressions as well. They serve a similar purpose as

Table 50: Synoptic table of the HSEC Twitter Data (full).

id	lemma	wf_t	cf_t	sum	domain	semantic_field
1	shit	1;3736	1;5013	100	object	shit_B1
2	god	1;947	0;0	9	concept	god_S9
3	fuck	2;1542	2;1746	36	concept	fuck_S3.2/B1
4	spirit	1;843	1;847	17	concept	spirit_S9
5	crap	1;570	1;773	13	object	crap_A5.1-
6	love	1;502	0;0	4	concept	love_E2+
7	one	1;493	0;0	4	concept	one_N1
8	jesus	1;512	0;0	4	person	jesus_Z1m
9	good	1;396	0;0	2	concept	good_A5.1+
10	day	1;359	0;0	2	time	day_T1.3
11	lord	1;331	0;0	2	title	lord_S7.1+/S2.2m
12	get	1;320	0;0	2	verb	get_A9+
13	people	1;321	0;0	2	person	people_S2mfc
14	know	1;294	0;0	1	verb	know_X2.2+
15	ghost	0;0	1;234	1	concept	ghost_S9mfn
16	grail	0;0	1;185	0	object	grail_O2
17	trinity	0;0	1;221	0	concept	trinity_Z3c
18	cow	0;0	1;293	1	object	cow_L2fn
19	mother	0;0	1;166	0	concept	mother_S4f
20	hell	0;0	1;307	1	concept	hell_S9
21	holy	0;0	1;3760	42	concept	holy_S9
22	water	0;0	1;206	0	object	water_O1.2
23	moly	0;0	1;269	1	concept	moly_Z99
24	communion	0;0	1;145	0	concept	communion_S9

5.2.3. HSEC Twitter Data (Religious)

5.2.3.1. Word Frequency List (AF)

The Twitter data classified as “religious” includes 2,660 tweets (RF 22,68%). The religious Twitter data consists of 74,071 running words and 9,712 types (stop words included).³⁷ The 15 most common words (“types”) in the subcorpus HSEC Twitter

37 Note, however, that there is a lot of noisy data left. Consequently, these numbers only serve as a rough approximation.

Data (religious) (no stop words included) are displayed in following table and bar plot (see the corresponding figure 44).

Table 51: Most frequent words (top 15) in HSEC Twitter Data (religious).

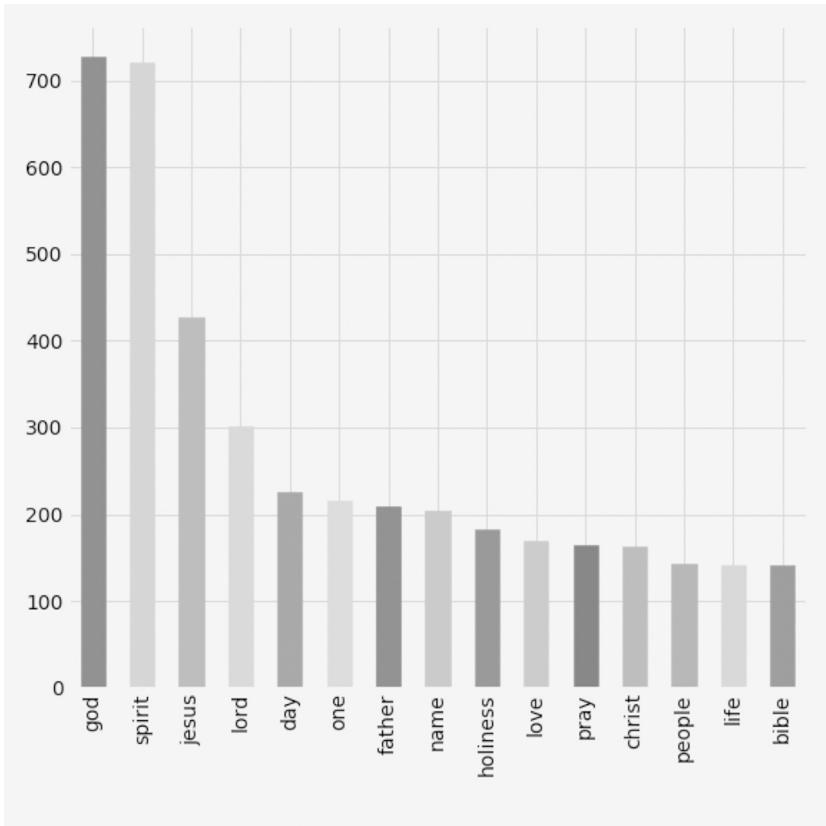
Type	Frequency
god	727
spirit	720
jesus	427
lord	302
day	225
one	215
father	209
name	204
holiness	183
love	169
pray	164
christ	163
people	143
life	142
bible	142

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The RF_1 is calculated on the basis of the 74,071 running words in the HSEC Twitter Data (religious) (including stop words). The RF_2 is calculated on the basis of the overall absolute frequency of the top 15 words (taken from the HSEC Twitter Data (religious) sample; without stop words), which is 4,135.

- Verbs and nouns related to religious “Concepts” (god, spirit, jesus, holiness, Christ, pray, bible) (AF 2,828, RF_1 3.8%, RF_2 68.4%)
- Words related to non-religious “Concepts” (love, one, father, name, life) (AF 939, RF_1 1.3%, RF_2 22.7%)
- Words related to “Time” (day) (AF 225, RF_1 0.3%, RF_2 5.4%)

The importance of terms related to religion in the HSEC Twitter Data (religious) is not surprising. It is also evident that words related to impurity are not present in

Figure 44: Most frequent words (top 15) and their absolute frequencies in HSEC Twitter Data (religious).



the data either. Thus, the use of words related to impurity seems to demarcate a religious and non-religious use of *holy* in the case of its contemporary application.

5.2.3.2. Collocation Analysis

The top 15 entries of the HSEC Twitter Data (religious) collocation analysis of the word “holy” (6-LL, L3-R3, C15-NC15; function words removed) can be found in the following table. The visualization is displayed in the corresponding figure 45.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The ab-

solute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 1,854).

Table 52: Top 15 words collocation analysis “holy” (6-LL, L3-R3, C15-NC15; function words removed) in HSEC Twitter Data (religious).

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	spirit	4167.21	666	720
2	r	ghost	909.88	136	131
3	r	communion	765.55	114	108
4	r	quran	472.34	82	107
5	r	water	463.55	81	93
6	r	bible	403.50	89	142
7	r	mother	344.87	61	71
8	r	shit	342.38	52	52
9	r	god	322.47	157	727
10	r	land	318.31	59	73
11	r	name	298.30	86	204
12	r	mary	242.66	61	118
13	l	jesus	211.20	98	427
14	r	amen	199.45	57	133
15	l	power	193.00	50	101

- Words related to religious “Concepts” and “Persons” (spirit, ghost, quran, bible, communion, prophet, mary, amen, jesus) (AF 1,465, RF 79%)
- Words related to non-religious “Concepts” (mother, name, power) (AF 197, RF 10.6%)
- Words related to “Objects” (water) (AF 81, RF 4.4%)
- Word related to “Space” (land) (AF 59, RF 3.2%)
- Words related to non-religious “Concepts” of impurity (shit) (AF 52, RF 2.8%)

Since the machine learning algorithm was trained with a manually annotated set of data, the frequent occurrence of words that are commonly associated with religion is an expected result. The somewhat surprising appearance of “shit,” however, is due to wrongly classified tweets by the machine learning classifier.³⁸ Words worth

³⁸ Regarding the overall performance of the selected machine learning classifier and the corresponding problems, see the machine learning excursus in Chapter 4.

Table 53: Synoptic table of HSEC Twitter Data (religious).

id	lemma	wf_t	cf_t	sum	domain	semantic_field
1	shit	0;0	1;52	0	object	shit_B1
2	god	1;727	1;157	52	concept	god_S9
3	name	1;204	1;86	19	concept	name_A1
4	spirit	1;720	1;666	100	concept	spirit_S9
5	christ	1;163	0;0	10	concept	christ_Z1m
6	love	1;169	0;0	10	concept	love_E2+
7	one	1;215	0;0	14	concept	one_N1
8	jesus	1;427	1;98	28	person	jesus_Z1m
9	mother	0;0	1;61	2	concept	mother_S4f
10	day	1;225	0;0	14	time	day_T1.3
11	lord	1;302	0;0	20	title	lord_S7.1+/S2.2m
12	father	1;209	0;0	13	concept	father_S4m
13	people	1;143	0;0	22	person	people_S2mfc
14	pray	1;164	0;0	9	verb	pray_S9/Q2.2
15	ghost	0;0	1;136	7	concept	ghost_S9mfn
17	land	0;0	1;59	2	place	land_W3
18	holiness	1;183	0;0	11	concept	holiness_S9
19	life	1;142	0;0	8	concept	life_L1+
20	bible	1;142	1;89	3	concept	bible_S9/Q4.1
21	water	0;0	1;81	2	object	water_O1.2
22	quran	0;0	1;82	4	concept	quran_S9
23	communion	0;0	1;114	2	concept	communion_S9
24	mary	0;0	1;61	1	person	mary_Z1f
25	power	0;0	1;50	0	concept	power_S7.1+
26	amen	0;0	1;57	0	concept	amen_Z4

As explained in Chapter 3, the selection of example sentences for the micro-analysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *holy* appears in the proximity of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

The following domains appear in the HSEC Twitter Data (religious) (ordered by their summarized “sum” values):

Table 54: Categories in the HSEC Twitter Data (religious) in descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	259
Person	41
Title	18
Time	13
Verb	8
Object	2
Place	0

Concept

With a “sum” value of 100, the most frequent word in the “concept” domain is the “spirit.” The “spirit” is also the top entry in the synoptic table of the HSEC Twitter Data (religious).

If you don't know if you've been baptized in the Holy Spirit, you haven't been. Come learn Sunday [LINK] When I was at work today I got this huge rush of happiness even though my back is broke and I have terrible cramps. I know that The Holy Spirit was there at that moment filling my heart with joy.

NowPlaying Holy Spirit Fall On Us by Michael Palattao

Jesus isn't just a little ray of truth. He is The Truth. This tells us that if we want truth instead of halftruth, we must look to Him. Its arrogant to imagine we'll always understand Him, but through the gift of His Holy Spirit we can obediently follow Him and learn from Him.

Discussion

The use of “holy” in combination with “spirit” in the HSEC Twitter Data (religious) shows many similarities with the use of “spirit” (or *hōgi*, *ays*) in the analysis of the Armenian data (and the Gospels in particular). For instance, the “spirit” in the Twitter data is also frequently mentioned in connection to baptism. In addition, tweets where the “holy spirit” is mentioned often deal with traditional religious topics or constellations that also play a specific role in the Armenian sources (at least implicitly). Both refer to classical (Christian) constellations of the “holy trinity” or represent religious actions such as (personal) prayers or thanksgiving as closely related to the holy spirit (see the second example sentence in the Twitter data). The fol-

lowing contemporary tweet provides an insightful example of a classical religious discussion that could have also taken place during the Christological councils of the fourth and fifth centuries CE.

I agree. The Holy Spirit is the Spirit of God. Very real and eternal And Holy. I question the term person, however. Is the spirit of a man a separate person from that man? Is your spirit a separate person from you? Thanks for your response, sir.

The “holy spirit” in the HSEC Twitter Data (religious) is mostly described as a strong religious agent that manifests through his actions. However, the actions are less concerned with purification or punishment (such as in the Armenian sources) than with learning and personal guidance for individuals.

I'm inspiring myself with guidance from the Holy Spirit.

The motif of guidance which runs through the tweets is a common feature of this genre of data. While companies, as well as religious institutions, use Twitter to communicate with their followers and promote events, it is mostly applied by individuals to share their experiences and thoughts with a broader community. The language register in these cases is, however, more informal.

The “holy spirit” is also frequently mentioned in song titles (that often still have a religious meaning), which are then displayed on Twitter to share them with friends.

Overall, the “holy spirit” is a reliable indication of a religious (or, at least, strong religioid) context in both the Armenian and the English data, as it is almost exclusively used in religious tweets. This assumption is underlined by the observation that although the “spirit” also appears in the synoptic table of the HSEC “Twitter Data” (non-religious), it only has a “sum” value of 0 in the non-religious context. An example of the appearance of the “holy spirit” in a non-religious tweet is the following example.

Finally, the best part is that you need not sign up for this service. I am like the holy spirit of God in that I am present on every back deck. I am perpetually on the edge of being wine drunk and im righteously angry all the time. if you hate threads pls unfollow thank u.

Even though the immediate context of this example does not have a religious connotation, the metaphorical use of the “holy spirit of God” in the sense of omnipresence is a strong indication of a religioid use case.

Overall, the use of “holy spirit” is difficult to compare with the Armenian sources, and this for obvious reasons: they are different text genres that stem from a specific cultural and historical contexts. This needs to be taken into consideration when the data is interpreted. That being said, one can recognize some level of correspondence between the two sets of data. In the Armenian and in the English data, the “holy spirit” is closely related to god and acts as a mighty actor that can

exert both a destructive power and a guiding function. Consequently, one might argue that the "holy spirit" and the *surb hogi* build an important diachronic link between the understanding of *holy* and *surb*.

Person

With a "sum" value of 35, the most important name in the "person" domain is "Jesus."

Hail Mary full of grace the Lord is with thee. Blessed art thou amongst women & blessed is the fruit of thy womb Jesus. Holy Mary Mother of God pray for us sinners now and at the hour of our death. Amen.

Discussion

That Jesus and Mary are the most frequently mentioned persons in religious tweets that mention "holy" is, at first sight, not very surprising. However, when looking back at the Armenian data, the saints (or martyrs) were often much more present than Jesus or Mary. The saints (plural), as well as single saints, also appear in the HSEC "Twitter Data" (religious), but they are, compared to other terms, less frequent. A review of the word frequency table reveals that the word "saints" only has a frequency of 19 and "saint" a frequency of 26. The divergences arising from the distinct genres aside, the discrepancy points to the lesser importance of saints within contemporary religiously connotated tweets that mention "holy," "holiness," or "sacred."

Title

The only word that appears in the synoptic table in the "title" domain is "lord" with a "sum" value of 18. As "lord" is almost exclusively used for god, this word will not be discussed in this section. The word "god" is also the second most important word in the "concept" domain, with a "sum" value of 62. Together with the relatively high frequency of "lord," this underlines that all three parts of the Christian trinity, namely god as well as Christ (see "Jesus" in the "Person" domain) and the "holy spirit" ("concept" domain), are significant in religiously classified tweets in the HSEC "Twitter Data."

The repeated occurrence of these three concepts from a macro-perspective is also observable in the Armenian data. It thus illustrates a religious continuity between ancient and contemporary religious use cases of *holy* in the English and the Armenian data. The pitfalls of such diachronic comparisons have already been mentioned several times. Furthermore, the affinity between the Armenian and English applications of these concepts is not surprising when considering the common Christian background of both cultures. Notwithstanding the difficulties of comparing texts of such different genres that stem from distinct historical and cultural contexts, it is interesting to note that the concepts of "god," "spirit," and "Christ"

remain central in both the English and the Armenian texts. These concepts connect to the semantic fields of *holy* and *surb* respectively and – I content – hint at historically overarching notion of *holy*.

Verb

With a “sum” value of 9, the most significant verb in the “verb” domain is “pray.”

Pray the Holy Rosary that the Church will be delivered from this devil [LINK]
 Holy Mary Mother of God pray for us sinners now and at the hour of our death.
 Amen.
 May the God of hope fill you with all joy and peace as you trust in him so that you may overflow with hope by the power of the Holy Spirit Romans [LINK]

Discussion

The relatively high frequencies of verbs such as “may” (even though it is not in the synoptic table) and “pray” in the HSEC Twitter Data (religious) need to be examined together. Both verbs confirm the tendency highlighted in the other examinations of this section, namely that many tweets are formulated as individual prayers asking for guidance and help.

Time

With a “sum” value of 13, the most frequent word in the “time” domain is the word “day.”

Holy. You're holy. Thank You God for this Holy Day
 Each day is sacred. Each day is holy. Some days the sacred and the holy are easy to see. Some days it feels as though the sacred and the holy are no where to be found (...)
 We had a bless time last night at church thanks to the Holy Spirit. Have a bless day today. I totally agree to that. That is power of the holy spirit in the our everyday lives. This reinforcers what prof Horton said (...)

Discussion

The importance of the term “day” also results from the fact that many of the religious tweets are prayers that either thank or ask for advice and help for a specific day. There are also tweets that are related to particular holy days (i.e., feasts) such as Sundays. Yet, the term “day” itself is only seldom used in the direct proximity of “holy.”

This is the day that the Lord has made. Let us rejoice and be glad in it. Praise the Lord O my soul and all that is in me. Praise his holy name. JESUS LOVES YOU HOW MUCH DO YOU LOVE HIM.

However, the majority of tweets that include “day” seem to be concerned with “everyday” days. The observation that the term “today” has a frequency value of over 100 underlines this assumption.

Space

With a “sum” value of 0, the most prominent word in the “space” domain is the word “land.” The “space” domain only includes the word “land” with a rather low overall value. This shows that spatial terms are of relatively little relevance in the HSEC Twitter Data (religious).

Join us Friday at seven on The Fred Holland Morning Show as Doctor Kevin Dieckilman, Founding President and Chief Executive Officer of One Heart for Israel, joins us by phone from the Holy Land.

Discussion

The term “land” often appears in the context of “holy” in the expression of the “holy land” (which is usually Israel). The problem of the attribution of such use cases as “religious” has already been discussed in Chapter 4. The tweet quoted in this section is a telling example of such use cases where the term “holy land” has a very strong religioid or even religious meaning. The organization “One Heart for Israel” has a caption that states “Connecting Christians & Jews to the Homeland” and advertises “Pastors Tours.” It claims to have the following goal:

One of the primary goals for One Heart for Israel is to equip pastors and church leaders with the full spiritual understanding of God’s Word in the context of the people, places and cultures in which it was lived.³⁹

This example tweet shows why I have decided to keep the use of “holy land” as an indication of a religious (or at least religioid) application of the terminology. The term “holy land” is applied to refer to Israel to highlight the religious significance of the territory from an inner-religious perspective. Another hint at the religious/religioid understanding of “holy land” appears in the fictional case of writing an academic (non-theological) article about the region where the modern state of Israel is located. In this context, a scholar would try to avoid using the terminology “holy land” because it is intuitively interpreted as indicating something that goes beyond a mere signification of the territory or geographical region by referring to religion.

Object

With a “sum” value of 2, the central word in the “object” domain is the word “water.”

39 *One Heart for Israel website*: <http://oneheartforisrael.com/project/pastors-tour/> (last accessed: 11/05/2021) .

I need some holy water

If prayers disgust you maybe drink some holy waterget rid of that evil inside you

[Link]

Discussion

The term “water” is also present in the Armenian data (i.e., the Gospels) and mostly used with reference to baptism. Yet, the term “water” is used differently in the HSEC Twitter Data (religious) compared to the Armenian *pendant*.

Firstly, in the context of tweets, the water is only rarely used to “wash” or to “clean” the body or the spirit. It is, however, often drunken or used to clean places. Secondly, the “water” in combination with “holy” is much more present in the HSEC Twitter Data (non-religious) than in the religious Twitter sample. The reason for this is that the expression of “holy water” is very often used ironically in the sense of “cleaning my eyes with holy water because I saw a very disturbing video” (this sentence is a fictional example and not part of the Twitter data). The non-religious use cases will be discussed in the following part on the HSEC Twitter Data (non-religious).

Overall, the appearance and meaning of “water” in the context of *holy* demonstrates that there is neither a fixed lexicographical meaning of words nor a fixed semantic localization in the sense of “religious” or “non-religious.” The ascription of meaning always depends on the context and can range from using “holy water” in a religious ritual (baptism) or in a situation that is absurd, funny, or disgusting. In the latter case, the application of “holy water” is metaphorical.

5.2.3.4. Final Discussion

The examination of the religious use of *holy* in the HSEC Twitter Data (religious) has revealed interesting features, particularly in comparison with the analysis of the Armenian corpus in the previous subchapter.

- The frequent use of words such as “Jesus,” “god,” and “spirit” indicates the constant adoption of typical (Christian) religious topics and actors that mostly stem from a Christian context and are thus also present in the Armenian data. However, the appearance of the “quran” in the synoptic table also proves that the terminology of *holy* is not restricted to Christianity but includes other religious traditions as well. At first, this might be obvious, but the ongoing discussions in the study of religion about the (il-)legitimacy of using allegedly Christian/Western terminology in the context of non-Christian/non-Western religious traditions demonstrate that this matter is more complicated than it might first appear. The tweets in the present Twitter sample show that the terminology is partially adopted from an inner-religious perspective of other religious tra-

ditions. Even though this is not the topic of this book, this observation can rightfully be regarded as an indication that a common field of *holy* exists across different (historical) traditions.

- The prevalence of the “Holy Spirit” is closely related to the importance of prayers in many tweets. This observation is also confirmed by the high frequency of different words related to god (“god,” “father,” “lord”), Christ, and Mary, who are all often directly addressed, as well as by the appearance of “amen” and “to pray” in the synoptic table. These variations highlight the individual and personal character of many tweets, which significantly differs from the historical/theological sources discussed in the examination of the Armenian data.
- The “person” domain is relatively unimportant in the religious tweets and mostly includes biblical persons such as Mary and Christ. Saints are sometimes mentioned as well, but they have a much lower frequency compared to other persons or to saints in the Armenian data. This observation might hint at the lower importance of “holy persons” in the contemporary use of *holy* compared to the ancient use of terms related to *holy*.

5.2.4. HSEC Twitter Data (Non-Religious)

5.2.4.1. Word Frequency List (AF)

The Twitter data classified as “non-religious” includes 9,066 tweets (RF 77.32%). The non-religious Twitter data includes 120,916 running words and 15,579 types (stop words included).⁴⁰ The 15 most common words (“types”) in the subcorpus HSEC Twitter Data (non-religious) (including stop words) are shown in following table and bar plot (see the corresponding figure 46).

The RF_1 is calculated on the basis of the 120,916 running words in the HSEC Twitter Data (non-religious) (including stop words). The RF_2 is calculated on the basis of the overall absolute frequency of the top 15 words taken from the HSEC Twitter Data (non-religious) (without stop words), which is 8,404.

40 Note, however, that there is a lot of noisy data left. Thus, these numbers only serve as a rough approximation.

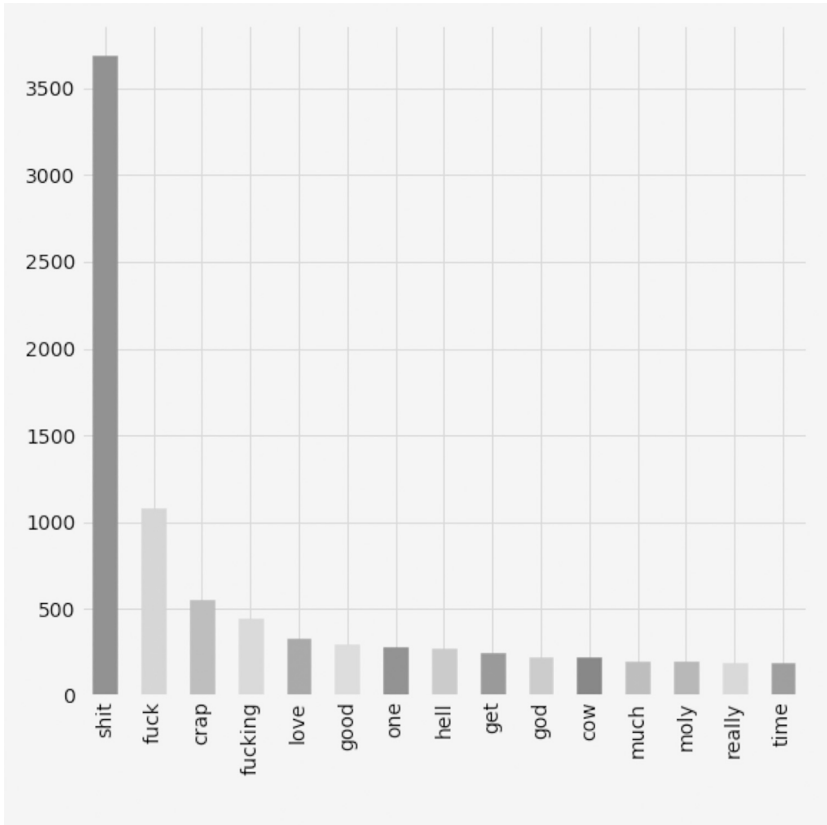
Table 55: Most frequent words (top 15) in HSEC Twitter Data (non-religious).

Type	Frequency
shit	3684
fuck	1078
crap	555
fucking	445
love	333
good	298
one	278
hell	275
get	249
god	220
cow	218
much	201
moly	195
really	188
time	187

- Words related to non-religious “Concepts” of impurity (fuck, shit, crap, fucking) (AF 5,762, RF₁ 4.8%, RF₂ 68.6%)
- Words related to (mostly positive) non-religious “Concepts” (love, good, one, much, moly) (AF 1,305, RF₁ 1.1%, RF₂ 15.5%)
- Words related to religious “Concepts” (hell, god) (AF 495, RF₁ 0.4%, RF₂ 5.9%)
- Misc. words such as adverbs and verbs (get, really) (AF 437, RF₁ 0.4%, RF₂ 5.2%)
- Words relate to “Objects” (cow) (AF 218, RF₁ 0.2%, RF₂ 2.6%)
- Words related to “Time” (time) (AF 187, RF₁ 0.2%, RF₂ 2.2%)

The results of the HSEC Twitter Data (non-religious) word frequency analysis support the hypothesis mentioned above that the word fields related to impurity can rightfully be taken as an indication of a non-religious use of *holy*. Terms related to impurity build the most significant part of the top entries in the word frequency analysis table. Interestingly, there are also words related to religion, such as “god” and “hell.” Their appearance might partially be explained by the wrong classification of religious tweets as “non-religious;” however, the terms “god” or “hell” in combination with “holy” frequently appear in a non-religious but exclamative sense as

Figure 46: Most frequent words (top 15) and their absolute frequencies in HSEC Twitter Data (non-religious).



well. This is another demonstration of the necessity to take the broader context of use into account when the semantics of words are interpreted.

5.2.4.2. Collocation Analysis

The top 15 entries of the HSEC Twitter Data (non-religious) collocation analysis of the word “holy” (6-LL, L3-R3, C35-NC35; function words removed) can be found in the following table. The corresponding visualization is displayed in figure 47.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 13,547).

Table 56: Top 15 words collocation analysis “holy” (6-LL, L3-R3, C35-NC35; function words removed) in HSEC Twitter Data (non-religious).

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	shit	33659.33	5232	3684
2	-	holy	9018.08	3808	9104
3	r	fuck	8392.33	1453	1078
4	r	crap	4589.05	790	555
5	r	cow	1690.64	298	218
6	r	hell	1631.99	309	275
7	r	moly	1584.42	276	195
8	r	fucking	1462.42	368	445
9	r	trinity	1154.36	212	173
10	r	grail	952.07	182	165
11	r	spirit	711.44	136	123
12	r	smokes	685.58	121	88
13	r	water	609.91	118	110
14	r	mother	573.33	112	107
15	r	night	504.30	132	165

- Words related to non-religious “Concepts” of impurity/sexuality (shit, fuck, crap, fucking) (AF 7,843, RF 57.9%)
- Words related to religious “Concepts” (holy, grail, trinity, hell, spirit) (AF 4,647, RF 34.3%)

- Words related to “Objects” (water, cow, smokes) (AF 537, RF 4.0%)
- Words related to other non-religious “Concepts” (moly, mother) (AF 388, RF 2.9%)
- Words related to “Time” (night) (AF 132, RF 1.0%)

The previous results of the word frequency analysis partly also hold for the collocation analysis.

However, in addition to the overall dominance of words related to impurity in the collocation analysis, the group of words related to religion is much larger than the group in the word frequency list (34% to 4%). This group includes some noisy data as well. The frequent use of words closely related to religion such as “grail” or “trinity” and “holy” is newly accounted for by their metaphorical, ironical, and exclamative application.⁴¹

5.2.4.3. Synoptic Table – Sentences Selection and Discussion (Microanalysis)

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses as well as their corresponding values and annotation has been created.

Table 57: Synoptic table of HSEC Twitter Data (non-religious).

id	lemma	wf_t	cf_t	sum	domain	semantic_field
1	shit	1;3684	1;5232	100	object	shit_B1
2	fuck	2;1523	2;1821	36	concept	fuck_S3.2/B1
3	crap	1;555	1;790	14	object	crap_A5.1-
4	love	1;333	0;0	2	concept	love_E2+
5	one	1;278	0;0	1	concept	one_N1
6	good	1;298	0;0	2	concept	good_A5.1+
7	get	1;249	0;0	1	verb	get_A9+
8	hell	1;275	1;309	5	concept	hell_S9

As explained in Chapter 3, the selection of example sentences for the micro-analysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and

41 “The holy grail of something” as a metaphor that something is precious/valuable/important but still not discovered yet.

9	god	1;220	0;0	1	concept	god_S9
10	cow	1;218	1;298	4	object	cow_L2fn
11	much	1;201	0;0	1	concept	much_A13.3
12	really	1;188	0;0	0	misc	really_A13.3
13	smokes	0;0	1;121	0	object	smokes_F3
14	holy	0;0	1;3808	41	concept	holy_S9
15	grail	0;0	1;182	0	object	grail_O2
16	trinity	0;0	1;212	1	concept	trinity_Z3c
17	moly	1;195	1;276	4	concept	moly_Z99
18	spirit	0;0	1;136	0	concept	spirit_S9
19	night	0;0	1;132	0	time	night_T1.3
20	water	0;0	1;118	0	object	water_O1.2
21	time	1;187	0;0	0	time	time_T1
22	mother	0;0	1;112	0	concept	mother_S4f

discussed in the section below. In case there are several sentences where *holy* appears in the proximity of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

The following domains appear in the HSEC Twitter Data (non-religious) (ordered by their summarized “sum” values):

Table 58: Categories in the HSEC Twitter Data (non-religious) in descending order according to their “sum” values in the synoptic table.

Domain	Sum
Object	118
Concept	94
Verb	1
Time	0
Misc	0

Concept

With a “sum” value of 41, the most important word in the “concept” domain is the term “holy.”

HOLY HOLY HOLY SHIT.

Now playing David Bowie Holy Holy Mono Single Version

Discussion

That “holy” is the top entry in the “concept” domain is due to the high frequency in the collocation analysis, meaning that the term “holy” is often found in the context of another “holy.” The use of multiple “holy” mostly functions as an amplifier to express an intense surprise or exclamation. Furthermore, the more religious or religioid meaning of the expression “holy holy holy” as in “Holy, holy, holy is the Lord of hosts; the whole earth is full of his glory!” (Isaiah 6:3) also appears in song titles that are then shared via tweets and thus appear in the (non-religious) Twitter data.

As *LancsBox* assumes that the overall HSEC Twitter Data (non-religious) is one coherent document, the frequent appearance of “holy” in the neighborhood of another “holy” is also due to the partially very short tweets. *LancsBox* evaluates two separate (in this case, fictional) tweets such as “Holy shit!” (tweet 1) and “Holy crap!” (tweet 2) as belonging together if they appear one after the other in the collection of tweets. Thus, a collocation analysis with a range of “R3-L3” considers that the occurrence of the appearance of “holy” in these two tweets belongs together from a collocation perspective, even though they stem from two different tweets. Therefore, the frequent occurrence of “holy” and particularly its top position in the collocation analysis has to be treated with caution. Still, the repetition of “holy” in one tweet is also observable in other contexts. Thus, the appearance in the tables is not only due to the skewed results of the collocation analysis. The other top entries in the “concept” domain, such as “fuck,” are discussed together with several terms of impurity as part of the “object” domain.

Object

With a “sum” value of 100, the most significant word in the “object” domain is “shit.” “Shit” is thus also the word in the synoptic table with the highest overall frequency.

Discussion

The use of “holy” together with “shit” is omnipresent in the (non-religious) Twitter data and does probably not need any concrete examples. The combination of “holy” and “shit” is the most frequent expression in a broader category of *holy* plus words related to impurity. Other examples of this field in the synoptic table of the HSEC Twitter Data (non-religious) are “crap,” “fuck,” and “cow.”

These expressions are mostly used in a very colloquial language to express surprise, disgust, or other (strong) emotions. What is most interesting about this exclusively non-religious use of “holy” is that it works via a contradiction that brings together the semantic field of *holy* and that of impurity. By associating *holy* with purity, the (non-religious) Twitter data takes up a specific meaning of *holy* that was visible in the Armenian data as well. Yet, the relationship between *holy*, purity, and impurity

seems to have vanished almost entirely from the contemporary religious use of *holy*, at least in the sources examined in this book. This makes it even more noteworthy that the relationship between *holy* and purity survived in expressions of everyday language that are (mostly) no longer connected to religion.

Time & Verb

With a “sum” value of 0, the most important words in the “time” domain are “night” and “time.” With a “sum” value of 1, the most significant term in the “verb” domain is “to get.”

The words in both domains do not add much to the discussion on the use and meaning of *holy* in non-religious tweets. However, the combination of “holy” and “night” exists in the religious tweets as well and commonly denotes the night before Christmas.

The verb “to get” and the noun “time” are also frequently applied in common expressions such as “it is time to” or “this time.” Because of their general relevance in the English (everyday) language, they consequently appear in the context of *holy* as well.

Because of their low “sum” value and since the two domains only include few words, I will not discuss specific example sentences in this part of the examination.

5.2.4.4. Final Discussion

The results of the analysis of the use of *holy* in the HSEC Twitter Data (non-religious) is straightforward. In the majority of cases, the term “holy” is used in short expressions to convey surprise, disgust, or any intense emotions. These expressions mostly consist of the attribute “holy” and a noun such as “shit,” “hell,” “fuck,” etc., which seems to contradict the religious meaning of “holy.” The juxtaposition of two terms that are opposed to each other creates a paradox, which then enhances the already mentioned intense emotion or the state of surprise. Interestingly, most of the words used as counterparts to the term “holy” stem from or are at least closely related to the semantic field of impurity. This semantic field also played an essential role in the Armenian corpora but seems to be less relevant in the contemporary religious use of *holy*.

5.2.7 HSEC Canon Law

5.2.7.1 Keyword Analysis

In the case of the HSEC Canon Law, it is possible to conduct an analysis of keywords with the sentences that include “holy” as a node corpus and the overall Canon Law corpus as a reference corpus.

Of the approximately 3,927 sentences in the cleaned HSEC Canon Law corpus, only 112 sentences include either the word “holy” or “holiness.” 137 sentences include

the word “sacred.” The potential differences in the use of “holy” compared to the use of “sacred” in the HSEC Canon Law are briefly discussed within the collocation analysis.

The keyword analysis, as well as most of the other examinations of the HSEC Canon Law data, focuses on the use of words related to “holy” or “holiness.” Possible differences between “sacred” and “holy” have already been discussed in the introductory excursus of Chapter 5. Despite some minor differences in the use cases of “holy” and “sacred,” the hypothesis that both belong to a common field of *holy* and are often used interchangeably was corroborated by the results of the analysis in the first part of Chapter 5.

The focus in this book lies on the question of the meaning of an abstract notion of *holy*, to which words such as “holy” and “sacred” belong. The corpus analysis in the following part focuses on those terms that are most relevant in the individual sources. Usually, the study of the word “holy” is preferred. However, there are some sources (for instance, HSEC Eliade) in which the term “holy” only rarely appears. In these cases, the application of the word “sacred” is examined instead.

In the case of the HSEC Canon Law, the frequency of the term “holy” is high enough to focus the examination on this term. Yet, the fact that “sacred” is used as well needs to be considered, too.

Figure 48: Keywords and lockwords in the HSEC Canon Law (only sentences incl. “holy”) (+) and HSEC Canon Law (full) (-).

1/706	Keywords +	1/1654	Lockwords	1/2460	Keywords -
1	holydays	1	abbess	1	judge
2	holyday	2	abducted	2	judgment
3	communion	3	abductor	3	power
4	sundays	4	abducts	4	lawfully
5	sunday	5	ability	5	party
6	feast	6	aborted	6	care
7	participate	7	absolved	7	proper
8	penitents	8	abundantly	8	matters
9	catechetical	9	abused	9	manner
10	see	10	abusing	10	rights

It is evident that the keywords in the smaller HSEC Canon Law sample (see the corresponding figure 48) that includes only sentences with “holy” are frequently related to the “time” domain by referring to different types of holidays. Another

important term in the subcorpus is “see,” which is “the see” (and not the verb “to see”). This demonstrates that, at least according to the analysis of keywords, *holy* often seems to be used in fixed expressions that stand for established (religious) institutions and dates such as “the holy see” (that stands for the Pope) or “holy days” > “holiday.”

5.2.7.2 Word Frequency List (AF)

Full Corpus

The full HSEC Canon Law data consists of 115,630 running words and 4,822 types (stop words included). The 15 most common words (“types”) in the HSEC Canon Law (full) (no stop words) are shown in following table and bar plot (see the corresponding figure 49).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The RF_1 is calculated based on the 115,630 running words in the HSEC Canon Law (full) (including stop words). The RF_2 is calculated based on the overall absolute frequency of the top 15 words (taken from the HSEC Canon Law (full) sample; without stop words), which is 5,610.

- Words related to religious or non-religious “Concepts” that mostly represent (legal) institutions (law, church, diocesan, judge, authority, office, see) (AF 2,559, RF_1 2.2%, RF_2 45.6%)
- Verbs as well as different attributes (may, unless, must, accordance, one, ordinary) (AF 2,115, RF_1 1.83%, RF_2 37.7%)
- Words related to “Persons” and “Titles” (person, bishop) (AF 936, RF_1 0.81%, RF_2 16.7%)

Table 59: Most frequent words (top 15) in HSEC Canon Law (full).

Type	Frequency
law	577
bishop	518
church	461
may	425
person	418
diocesan	387
unless	386
must	346
accordance	339
one	318
judge	310
ordinary	301
authority	286
office	270
see	268

The word frequency list of the HSEC Canon Law (full) corpus reveals what one would expect from a corpus that is supposed to represent a set of laws regulating fundamental aspects of a religious institution. Yet, it is remarkable that there are only very few words with a direct religious connotation; instead, most terms refer to institutional aspects and persons. It is important to keep in mind that words related to *holy* only seldom appear in the HSEC Canon Law (full) when one interprets the results of this examination.

Sentences with "Holy" Only

The HSEC Canon Law (holy) data with sentences that include "holy" only consists of 4,175 running words and 960 types (stop words included). The 15 most common words ("types") in the HSEC Canon Law (holy) (no stop words) are shown in following table and bar plot (see the corresponding figure 50).

Figure 49: Most frequent words (top 15) and their absolute frequencies in HSEC Canon Law (full).

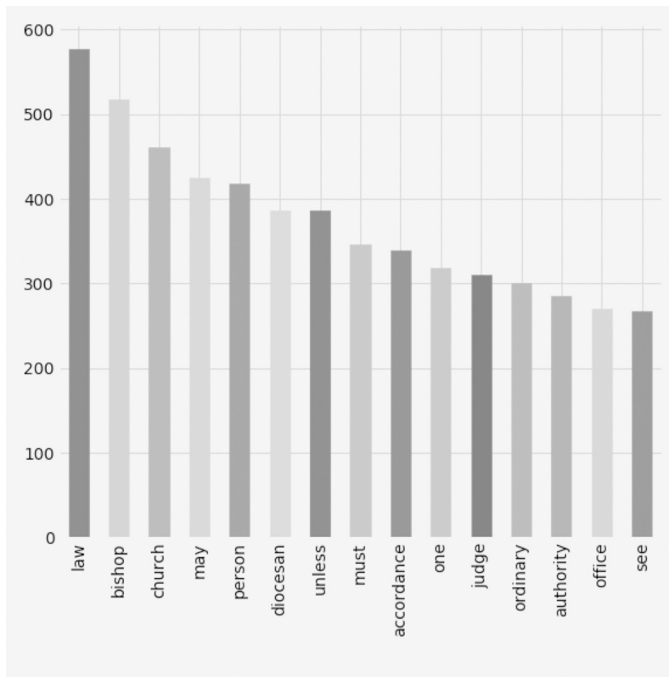
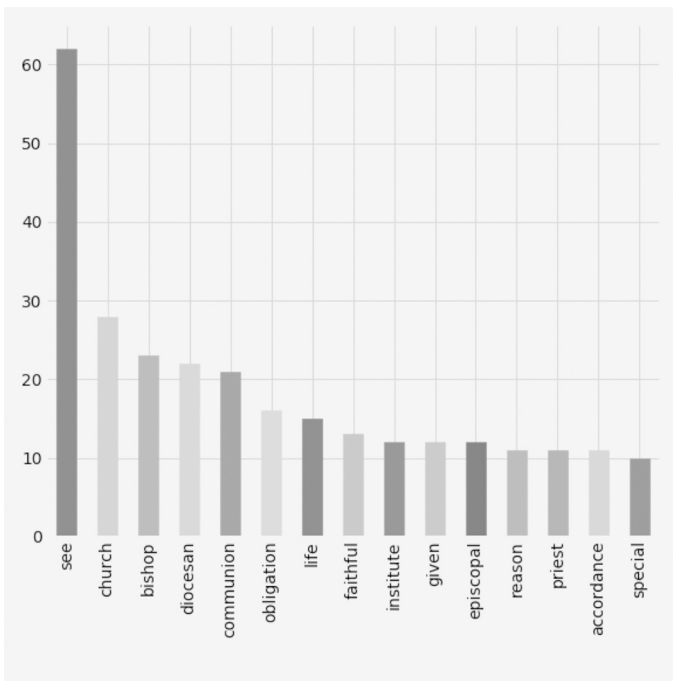


Table 60: Most frequent words (top 15) in HSEC Canon Law (holy).

Type	Frequency
see	62
church	28
bishop	23
diocesan	22
communion	21
obligation	16
life	15
faithful	13
institute	12
given	12
episcopal	12

reason	11
priest	11
accordance	11
special	10

Figure 50: Most frequent words (top 15) and their absolute frequencies in HSEC Canon Law (holy).



The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The RF₁ is calculated based on the 4,175 running words in the HSEC Canon Law (full) (including stop words). The RF₂ is calculated based on the overall absolute frequency of the top 15 words (taken from the HSEC Canon Law (full) sample; without stop words), which is 279.

- Words related religious “Concepts” that represent institutions (see, church, diocesan, institute) (AF 124, RF₁ 13%, RF₂ 44.4%)

- Words related to other religious and non-religious “Concepts” (communion, obligation, life, faithful, given, reason, accordance, special) (AF 109, RF₁ 2.6%, RF₂ 39.1%)
- Words related to “Titles” (episcopal, bishop, priest) (AF 46, RF₁ 1.1%, RF₂ 16.5%)

Similarly to the previous word frequency list, the second list also includes many words that are closely related to institutions. This is explained by the strong overall institutional and legal character the Canon Law.

In this source, persons and titles play a much more important role than in the Twitter data. With respect to sentences including “holy,” one can observe a significant difference between the word lists of the overall corpus and the subcorpus with sentences that include “holy,” namely the occurrence of words related to concepts such as “life,” “special,” or “obligation.” These words are more present than in the smaller sample of the HSEC Canon Law data. Many of these concepts, such as “life” or “communion,” are also commonly found in other word lists in this book.

Due to the unrepresentative character of the first word frequency list for the application and meaning of *holy*, only the results from the second word frequency list of the subcorpus HSEC Canon Law (*holy*) are considered in the synoptic table.

5.2.7.3 Collocation Analysis

Holy

The following top 15 entries of the HSEC Canon Law collocation analysis of the term “holy” (6-LL, L2-R2, C3-NC3; function words removed)⁴² can be found in the following table. The term “holy” only appears 93 times in the overall HSEC Canon Law (RF 0.08%). The visualization is displayed in the corresponding figure 51.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 136).

- Words related to non-religious “Concepts” and “Titles” (see, minister, approval, permission) (AF 69, RF 51%)
- Words related to religious “Concepts” (communion, spirit, sacrifice, holy, eucharist) (AF 40, RF 29%)

42 As the HSEC Canon Law, as well as the following sources, includes less noisy data compared to the Twitter data, the range used in the collocation analysis will usually be restricted to L2-R2 instead of L3-R3.

- Verbs and other words (reserved, receive, required, provided, one, must) (AF 27, RF 20%)

Table 61: Top 15 words collocation analysis “holy” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full).

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	see	482.56	56	268
2	r	communion	218.53	22	49
3	r	spirit	63.12	8	39
4	l	reserved	45.24	7	66
5	l	receive	30.8	5	54
6	l	approval	28.23	4	28
7	l	minister	25.44	4	39
8	l	permission	25.07	5	95
9	l	sacrifice	22.51	3	17
10	r	required	18.85	4	88
11	r	holy	18.42	4	93
12	l	eucharist	14.68	3	60
13	r	provided	11.03	3	112
14	l	one	9.2	4	318
15	l	must	8.61	4	346

The notion of *holy* is only very rarely related to persons in the HSEC Canon Law; however, this is mainly since “holy persons” in English are usually called “saints.” The word “holy” in the HSEC Canon Law is mostly related to religious and non-religious concepts. Besides the “holy spirit,” it is noteworthy that some of the central rites in the Catholic liturgy, such as the “communion” or the “eucharist,” frequently appear in the neighborhood of the word “holy.” Together with the observation that the “holy see” is also commonly mentioned in the HSEC Canon Law, it becomes clear that the attribute “holy” is connected to some of the most important institutions and rites of the Catholic Church, thereby underlining their extraordinary character.

Sacred

Since the term “holy” only appears a few times in the HSEC Canon Law corpus, the use of “sacred” is also considered in the collocation analysis. Since this book focuses not only on the notion of *holy* but, at least where it is possible, on words related to “holy” as well, only the results from the examination of the use of “holy” in the

Figure 51: Top 15 types collocation analysis “holy” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full) (graph) (c) LancsBox.



HSEC Canon Law corpus are considered in the further analysis and the synoptic table.

This section should shed further light on the use and meaning of “holy” in the HSEC Canon Law in relation to the use of “sacred.” Therefore, the examination of shared collocates between “sacred” and “holy” is also conducted in the following section. The top 15 entries of the HSEC Canon Law collocation analysis of the term “sacred” (6-LL, L2-R2, C3-NC3; function words removed) are found in the following table. The visualization is displayed in the corresponding figure 52. The term “sacred” appears 144 times in the overall HSEC Canon Law (RF 0.13%).

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)”

column in the collocation table (overall words frequency collocation: 120). There is only one shared collocate (“minister”) among the collocates of “holy” and “sacred.”

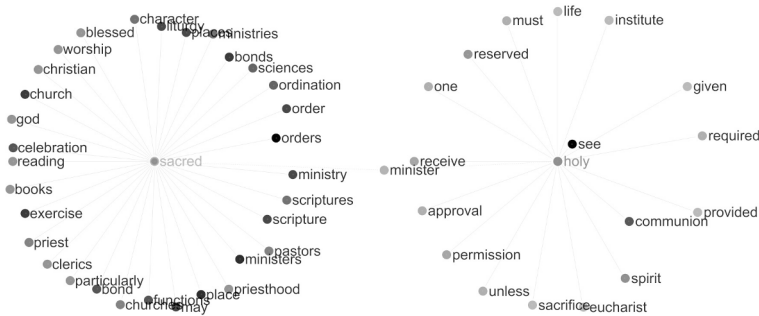
Table 62: Top 15 words collocation analysis “sacred” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full).

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	orders	106.44	14	53
2	r	scripture	95.34	8	8
3	r	bonds	89.1	9	14
4	r	ministers	83.81	10	27
5	r	places	71.87	8	17
6	r	liturgy	62.62	8	27
7	r	sciences	62.43	6	8
8	r	scriptures	54.08	5	6
9	r	ministry	49.95	8	55
10	r	functions	48.48	7	35
11	l	exercise	44.4	9	117
12	r	ministries	44.35	5	11
13	r	bond	43.1	7	50
14	r	place	41.15	10	197
15	r	ordination	41.08	6	31

The words from the collocation analysis can be grouped according to the “do-main” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 120). There is only one shared collocate (“minister”) among the collocates of “holy” and “sacred.”

- Words related to religious “Concepts” that represent institutions and practices (ministers, liturgy, ministry, functions, exercise, ministries, ordination) (AF 53, RF 44%)
- Words related to non-religious “Concepts” and “Objects” (scripture, orders, bonds, sciences, scriptures, bond) (AF 49, RF 41%)
- Words related to “Places” (places, place) (AF 18, RF 15%)

Figure 52: Top 15 types collocation analysis “sacred” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Canon Law (full) and its overlappings with “holy” (graph) (c) Lancs-Box.



Compared to the use of “holy,” the use of “sacred” is in line with the observations made in the excursus at the beginning of this subchapter. There are many similarities between the two terms, which are, for instance, both used to describe religious institutions and concepts. The term “sacred” is, however, more frequently applied to denote the religious character of a place or an object. Furthermore, the word “sacred” is more frequently encountered in the vicinity of concepts and objects that are not religious. Therefore, in the context of the Canon Law, it compares to the ancient use where Latin *sacer* was applied to indicate that a “normal” object now belonged to a deity (see Chapter 2).

5.2.7.4 Synoptic Table – Sentences Selection and Microanalysis

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses as well as their corresponding values and annotation has been created.

The results of the HSEC Canon Law collocation analysis on “sacred” are not considered in the following table. Regarding the frequency lists, only the data from the word frequency list of the smaller subcorpus HSEC Canon Law (holy) are considered.

As explained in Chapter 3, the selection of example sentences for the microanalysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and

discussed in the section below. In case there are several sentences where *holy* appears in the proximity of the selected lemma, the sentence that offers the most potential for the analysis is manually selected. In addition to the English text, the Latin text will be provided as well.

Table 63: Synoptic table of the HSEC Canon Law.

id	lemma	kw	wf_t	cf_t	sum	domain	semantic_field
1	see	1	1;62	1;56	100	concept	see_O2
2	commu- nion	1	1;21	1;22	40	concept	communion_S9
3	spirit	0	0;0	1;8	3	concept	spirit_S9
4	reserve	0	0;0	1;7	3	verb	reserve_A9+
5	receive	0	0;0	1;5	1	verb	receive_A9+
6	approval	0	0;0	1;4	0	concept	approval_S7.4+
7	minister	0	0;0	1;4	0	title	minister_G1.1/S2mf
8	permission	0	0;0	1;5	1	concept	permission_S7.4+
9	sacrifice	0	0;0	1;3	0	concept	sacrifice_S9
10	require	0	0;0	1;4	0	verb	require_X7+
11	holy	0	0;0	1;4	0	concept	holy_S9
12	eucharist	0	0;0	1;3	0	concept	eucharist_S9
13	provide	0	0;0	1;3	0	verb	provide_A9-
14	one	0	0;0	1;4	0	concept	one_N1
15	must	0	0;0	1;4	0	verb	must_S6+
16	bishop	0	1;23	0;0	15	title	bishop_S9/S7.1+/S2mf
17	church	0	1;28	0;0	19	concept	church_S9/H1c
18	diocesan	0	1;22	0;0	15	concept	diocesan_S9/M7
19	obligation	0	1;16	0;0	10	concept	obligation_S6+
20	life	0	1;15	0;0	9	concept	life_L1+
21	faithful	0	1;13	0;0	7	attribute	faithful_S6+
22	institute	0	1;12	0;0	7	concept	institute_P1/S5+c
23	given	0	1;12	0;0	7	attribute	given_A9-
23	episcopal	0	1;12	0;0	7	attribute	episcopal_S9
24	reason	0	1;11	0;0	6	concept	reason_A2.2
25	priest	0	1;11	0;0	6	title	priest_S9/S2mf
26	accordance	0	1;11	0;0	6	concept	accordance_S7.1-
27	special	0	1;10	0;0	5	attribute	special_A4.2+

28	holiday	1	0;0	0;0	4	time	holiday_K1/T1.3
29	feast	1	0;0	0;0	4	time	feast_F1
30	Sunday	1	0;0	0;0	4	time	sunday_T1.3
31	participate	1	0;0	0;0	4	verb	participate_S1.1.3+
32	penitents	1	0;0	0;0	4	concept	penitents_E4.1-
33	catechetical	1	0;0	0;0	4	attribute	catechetical_S9

The domains shown in table 64 appear in the HSEC Canon Law (ordered by their summarized “sum” values):

Table 64: Categories in the HSEC Canon Law in descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	220
Attribute	30
Title	21
Time	12
Verb	8

Concept

With a “sum” value of 100, the most important word in the “concept” domain is the “see.” “See” is thus also the word in the synoptic table with the highest frequency.

(Canon 838, §3) It pertains to Episcopal Conferences to prepare vernacular translations of liturgical books, with appropriate adaptations as allowed by the books themselves and, with the prior review of the Holy See, to publish these translations.

Ad Episcoporum conferentias spectat versiones librorum liturgicorum in linguas vernaculas, convenienter intra limites in ipsis libris liturgicis definitos aptatas, parare, easque edere, praevia recognitione Sanctae Sedis.

Discussion

The combination of “holy” and “see” is used as a *terminus technicus* in the Canon Laws denoting the jurisdiction of the bishop of Rome, the Pope, and hence the head of the Catholic Church. The “Holy See” is used as a translation of the Latin “Sancta Sedes.” The use of “holy” in this combination gives another example of an abstract use case where “holy” no longer serves as an attribute to highlight specific aspects of an

object, concept, or person. Instead, the attribution of “holy” has become part of a fixed expression (in this case, a position in a hierarchy) that stands for itself. The use of “holy” and “see” as a fixed expression is further underlined by the capitalization in both Latin and English.

Attributes

With a “sum” value of 7, the most significant words in the “attribute” domain are “given,” “faithful,” and “episcopal.” As previously mentioned, the “attribute” and “concept” domains are closely related.⁴³ In the HSEC Canon Law, some adjectives such as “diocesan” are assigned to the “concepts” domain, whereas others belonged to the “attributes.”

(Canon 242, §1) In each country there is to be a Charter of Priestly Formation. It is to be drawn up by the Episcopal Conference, taking account of the norms issued by the supreme ecclesiastical authority, and it is to be approved by the Holy See; (...)

In singulis nationibus habeatur institutionis sacerdotalis Ratio, ab Episcoporum conferentia attentis quidem normis a suprema Ecclesiae auctoritate latis, statuenda et a Sancta Sede approbanda (...)

(Canon 387) Mindful that he is bound to give an example of holiness, charity, humility and simplicity of life, the diocesan Bishop is to seek in every way to promote the holiness of Christ’s faithful according to the special vocation of each. Episcopus dioecesanus, cum memor sit se obligatione teneri exemplum sanctitatis praebendi in caritate, humilitate et vitae simplicitate, omni ope promovere studeat sanctitatem christifidelium secundum uniuscuiusque propriam vocationem atque,

Discussion

The importance of attributes and concepts such as “diocesan” or “episcopal” is not only typical for the “holy” subcorpus of the HSEC Canon Law but can also be observed across the Canon Law data. For instance, the word “diocesan” is, with a frequency of 387, also among the top entries in the word frequency list of the HSEC Canon Law. The attribute “episcopal” is also quite common (AF 171). The general importance of these attributes in connection to titles in an institutional Church hierarchy raises the question in how far these words inform us about the broader use of “holy.” The example of Canon 387 demonstrates the extent to which the terminology around “holy” is crucial for the meaning of ecclesiastic titles.

In this example, the term “holiness” (*sanctitas*) is listed together with “humility” (*humilitas*), “charity” (*caritas*), as well as “simplicity” (*simplicitas*) as typical attributes

43 See Chapter 3 on the annotation framework for more details.

of a bishop's conduct of life (by taking the life of Christ as an example). The presence of "holiness" as a part of a longer chain of attributes in this example is comparable to the way *surb* was frequently used in attribute chains in the Armenian texts. Despite the sedimented character of *holy* in fixed expressions such as "the Holy See" or "the Holy Spirit," this sentence illustrates that the use of the terminology of *holy* as a specific attribute is still present.

Title

With a "sum" value of 15, the central word in the "title" domain is the "bishop." The importance of titles, as well as of other institutional aspects, has already been discussed in the section on "attributes." Consequently, a more in-depth analysis of the use of titles such as "bishop" or "priest" will not be conducted at this stage. To reiterate, it is essential to note the general significance of titles and positions in the HSEC Canon Law that is not only observable in subcorpus with sentences where the word "holy" (or "holiness") appears.

Time

With a "sum" value of 4, the most important words in the "time" domain are related to feasts, such as "sunday," "holyday," and "feast." All these words are only part of the synoptic table because they appear as important words in the keyword analysis.

(Canon 534, §1) When he has taken possession of his parish, the parish priest is bound on each Sunday and holyday of obligation in his diocese to apply the Mass for the people entrusted to him. (...)

Parochus, post captam paroeciae possessionem, obligatione tenetur singulis diebus dominicis atque festis in sua dioecesi de praecepto Missam pro populo sibi commisso applicandi;

Discussion

The expression "each Sunday and holyday" translates *singulis diebus dominicis atque festis*. This means that there is no appearance of *sanctus* in the corresponding Latin text. Instead, the noun-adjective combination of *dies festus* is interpreted as a "holy day" (> "holyday"> "holiday") and "holy" thus employed by the translator to underline "something as extraordinary/special."⁴⁴ This example of "holy day" that finally became the "holiday" is another illustrative instance of the transformation of an attribute into a sedimented expression previously noted above. The significance of "holy" as an autonomous attribute is diminished by its merging of "holy" and "day" into "holiday."

44 Because the interruption of everyday life is what defines a feast. Furthermore, the parallelity in the concept of ancient *dies festi* and "holidays" is implied as well.

Furthermore, the link between obligations of bishops and “holy” expressed in this example highlights the interrelation between *holy* and the hierarchical Church structure presented in the Canon Law.

Verb

With a “sum” value of 4, the most significant word in the “verb” domain is the verb “to participate.”

(Canon 663, §2) Each day the members are to make every effort to participate in the Eucharistic sacrifice, receive the most holy Body of Christ and adore the Lord himself present in the Sacrament.

Sodales cotidie pro viribus Sacrificium eucharisticum participant, sanctissimum Corpus Christi recipiant et ipsum Dominum in Sacramento praesentem adorent.

Discussion

The verbs “to participate” and “to receive,” which both occur in the synoptic table, are used in the context of the Eucharist (*Sacrificium eucharisticum*) where people are obliged to participate to receive the “most holy Body of Christ” (*sanctissimum Corpus Christi*). The use case of *holy* in this example is of particular interest since it includes different layers of sedimentation processes.

The first layer is the combination of “holy” and “body” (of Christ). This example could present us with a rare use case of “holy” and a material object in the English corpora. However, the “body of Christ” is written in initial capitals. As previously mentioned with reference to the “Holy See,” the capitalization can be interpreted as an indication of a more abstract or sedimented application of *holy*. This assumption is confirmed by the semantic content of the sentence that is concerned with the Eucharist, and not with the physical body of Christ. The latter, thus, builds the second layer. The “Body of Christ,” in this case, is an expression alluding to the bread in the central liturgical part of every Catholic mass, thereby serving as another instance of the sedimentation of certain words in fixed expressions. However, the term “most holy” (*sanctissimum*) is not part of this fixed expression since it is set to lowercase. Thus, it further qualifies the “Body of Christ.” The question is whether “holy” characterizes the “body” of Christ or the bread in the Eucharist – or both. In theory, both solutions are possible since the combination of “holy” with concepts has been encountered as well as that of “holy” with material objects. A definite answer is not possible in this example. What can be said, however, is that the ascription of “holy” as either an attribute of a human body or of bread oscillates between the material and abstract conceptual spheres. This, in turn, is perfectly in line with the Catholic dogma of transsubstantiation, which assumes that the bread is the body of Christ.

5.2.7.5 Final Discussion

The examination of the HSEC Canon Law data has revealed the following results and insights into the use of *holy*.

- Both word fields “holy” and “sacred” were almost equally present in the data. A collocation analysis of both terms revealed that their use was generally in line with the results presented in the introductory excursus in this subchapter. The term “sacred” was more frequently used to denote objects and non-religious concepts, whereas the term “holy” was almost exclusively applied to describe central religious concepts and titles.
- The use of “holy” was often part of fixed expressions (regularly written with initial capital letters) such as “the Holy See” or “the Holy Spirit.” The fixed expressions usually denoted central religious aspects or institutions. In these cases, the meaning of the term “holy” has merged with the meaning of its counterpart, which, in turn, also impacts the semantic field of *holy*.
- Despite the frequent application of “holy” in fixed expressions such as “the Holy See,” the term “holy” (or “holiness”) is still used in attribute chains to explicitly describe and define certain persons, titles, and institutions. It shows that the word field of “holy” can still potentially function as an attribute to highlight specific aspects alongside its frequent use in sedimented fixed expressions.

5.2.8 HSEC Encyclopaedia Britannica

5.2.8.1 Keyword Analysis

Since the entry on “Sacred” in the HSEC Encyclopaedia Britannica is concerned with the notion in question, this section will not offer a keyword analysis due to the lack of an appropriate reference corpus.

5.2.8.2 Word Frequency List (AF)

The HSEC Encyclopaedia Britannica data consists of 4,177 running words and 1,217 types (stop words included). The 15 most common words (types) in the HSEC Encyclopaedia Britannica (no stop words) are shown in the following table and bar plot (see also the corresponding figure 53).

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The RF_1 is calculated based on the 4,177 running words in the HSEC Encyclopaedia Britannica (including stop words). The RF_2 is calculated based on the overall absolute frequency of the top 15 words (including stop words), which is 345. Note that the word “holy” only occurs eight times. The reason for this is that the HSEC Encyclopaedia Britannica

data mainly focuses on the term “sacred” that is assumed to be a synonym for other terms related to the semantic field of *holy* (such as “holy”).⁴⁵

Table 65: Most frequent words (top 15) in HSEC Encyclopaedia Britannica.

Type	Frequency
sacred	115
religious	28
life	24
power	21
divine	19
human	17
different	16
forms	14
reality	14
religion	14
world	14
social	13
profane	12
one	12
existence	12

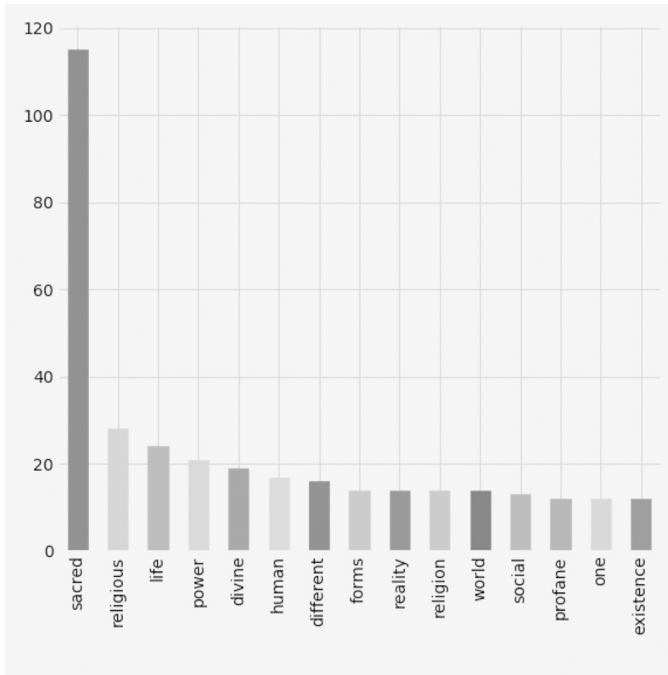
- Words related to general “Concepts” and, to a small extent, “Places” (life, power, human, different, forms, reality, world, social, one, existence) (AF 157, RF₁ 3.76%, RF₂ 45.5%)
- Words related to religious “Concepts” (sacred, religious, divine, religion, profane) (AF 188, RF₁ 4.5%, RF₂ 54.5%)

The overview of the most frequent words in the HSEC Encyclopaedia Britannica shows that the entry on “Sacred” is primarily concerned with non-religious and religious concepts. Among the religious concepts is the duality of *sacred* and *profane*, which is an instance of the *separation-binding* process deemed crucial for our research on the notion of *holy* (see Chapter 2).

Among the non-religious concepts are examples related to very abstract and general ideas concerning the world and the humans that live in it, such as “life,” “human,” “reality,” and “existence.” Interestingly, the notion of power is also found

45 See the introduction of the *Encyclopaedia Britannica* article on “Sacred.”

Figure 53: Most frequent words (top 15) and their absolute frequencies in HSEC Encyclopaedia Britannica.



among the top entries of the word frequency lists, thereby highlighting another critical property of the modern notion of *holy* (see Chapter 2).

The results of this word frequency analysis become even more interesting when searching for the categories which were significant in the other texts (particularly in the Armenian corpora). Persons, titles, or objects are not found among the top entries of the word frequency list of the HSEC Encyclopaedia Britannica entry. This shows that in contemporary entries on “Sacred” (that aim to be representative for the overall notion of *holy*) the focus has shifted from persons and their conduct of life to abstract concepts, such as “power,” “reality,” or “human.”

5.2.8.3 Collocation Analysis

Even though this part of the examination mainly focuses on the word field around “sacred” due to the low frequency of “holy,” the collocation analysis will be conducted for both terms. Both collocation analyses are considered in the final synoptic table.

- Words related to non-religious “Concepts” (idea, ultimate, dependence, destinies, effect, great, illumination) (AF 9, RF 50%)
- Words related to “Persons” (otto, brede, heidegger) (AF 4, RF 22.22%)
- Verbs (appeared, elaborated, exercised, fulfill) (AF 4, 22.22%)
- Words related to religious “Concepts” (heilige) (AF 1, RF 5.55%)

The results of the collocation analysis reveal a strong emphasis on general concepts that mostly have a non-religious or religioid⁴⁶ character such as “ultimate,” “illumination,” or “destiny.” Unlike in the word frequency list, however, the category of “persons” is also present in the collocation analysis. Yet, the persons who appear in this context are not saints but mostly scholars and philosophers who have contributed to the modern understanding of the notion of *holy* (such as Rudolf Otto).

Table 66: Top 15 words collocation analysis “holy” (6-LL, L4-R4, C1-NC1; function words removed) in HSEC Encyclopaedia Britannica.

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	l	idea	19.28	2	3
2	r	ultimate	16.38	2	5
3	l	otto	15.4	2	6
4	r	appeared	11.4	1	1
5	l	brede	11.4	1	1
6	r	dependence	11.4	1	1
7	l	destinies	11.4	1	1
8	l	effect	11.4	1	1
9	l	elaborated	11.4	1	1
10	r	exercised	11.4	1	1
11	r	fulfill	11.4	1	1
12	r	great	11.4	1	1
13	l	heidegger	11.4	1	1
14	l	heilige	11.4	1	1
15	r	illumination	11.4	1	1

Sacred

The top 15 entries of the HSEC Encyclopaedia Britannica collocation analysis of the term “sacred” (6-LL, L2-R2, C2-NC2; function words removed) can be found in

⁴⁶ Note that the term religioid in this context denotes a notion/word with the potential to become religious. However, it does not necessarily need to have its origin in religion.

the following table. The term “sacred” appears 115 times in the overall HSEC Encyclopaedia Britannica. The visualization is displayed in the corresponding figure 55.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 60).

Table 67: Top 15 words collocation analysis “sacred” (6-LL, L2-R2, C2-NC2; function words removed) in HSEC Encyclopaedia Britannica.

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	l	idea	19.28	2	3
2	r	ultimate	16.38	2	5
3	l	otto	15.4	2	6
4	r	appeared	11.4	1	1
5	l	brede	11.4	1	1
6	r	dependence	11.4	1	1
7	l	destinies	11.4	1	1
8	l	effect	11.4	1	1
9	l	elaborated	11.4	1	1
10	r	exercised	11.4	1	1
11	r	fulfill	11.4	1	1
12	r	great	11.4	1	1
13	l	heidegger	11.4	1	1
14	l	heilige	11.4	1	1
15	r	illumination	11.4	1	1

- Words related to non-religious “Concepts” and “Time” (nature, power, notion, ambiguity, basic, dichotomy, manifestation, time, dialectic) (AF 39, RF 65%)
- Verbs (manifested, allows, argued, assumes) (AF 10, RF 16.66%)
- Words related to religious “Concepts” (profane) (AF 6, RF 10%)
- Words related to “Places” (places) (AF 5, AF 8.33%)

The results of the collocation analysis of “sacred” are very much in line with the results of the previous collocation analysis of “holy.” Both analyses have a strong

values and annotation has been created. The results of the HSEC Encyclopaedia Britannica collocation analysis on “sacred” are also considered in the synoptic table together with the results from the collocation analysis of “holy.” This decision has been made due to the overall low frequency of “holy” in the HSEC Encyclopaedia Britannica, which would not yield sufficient data for analysis. Secondly, the term “sacred” is, according to the self-description of the article, used as a synonym for other words such as “holy” that all refer to a common semantic field of *holy*.

As explained in Chapter 3, the selection of example sentences for the micro-analysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *holy* appears in the proximity of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

Table 68: Synoptic table of the HSEC Encyclopaedia Britannica.

id	lemma	wf_t	cf_t	sum	domain	semantic_field
1	sacred	1;115	0;0	100	concept	sacred_S9
2	religion	2;42	0;0	36	concept	religion_S9
3	life	1;24	0;0	20	concept	life_L1+
4	power	1;21	1;9	26	concept	power_S7.1+
5	divine	1;19	0;0	15	concept	divine_S9
6	human	1;17	0;0	14	concept	human_S2
7	different	1;16	0;0	13	concept	different_A6.1-
8	forms	1;14	0;0	11	concept	forms_A4.1
9	reality	1;14	0;0	11	concept	reality_A3+
10	world	1;14	0;0	11	place	world_W1
11	social	1;13	0;0	10	concept	social_S1.1.1
12	profane	1;12	1;6	15	concept	profane_S9
13	one	1;12	0;0	9	concept	one_N1
14	existence	1;12	0;0	9	concept	existence_A3+
15	idea	0;0	1;2	0	concept	idea_X4.1

16	ultimate	0;0	1;2	0	concept	ultimate_T2-
17	otto	0;0	1;2	0	person	otto_Z1
18	appear	0;0	1;1	0	verb	appear_A8
19	brede	0;0	1;1	0	person	brede_Z1
20	dependence	0;0	1;1	0	concept	dependence_A2.2
21	destiny	0;0	1;1	0	concept	destiny_A7+++/S7.1-
22	effect	0;0	1;1	0	concept	effect_A2.2
23	elaborated	0;0	1;1	0	attribute	elaborated_Q2.2
24	exercise	0;0	1;1	0	verb	exercise_K5.1
25	fulfill	0;0	1;1	0	verb	fulfill_X9.2+
26	great	0;0	1;1	0	attribute	great_A5.1+
27	heidegger	0;0	1;1	0	person	heidegger_Z1
28	heilige	0;0	1;1	0	concept	heilige_S9
29	illumination	0;0	1;1	0	concept	illumination_W2
30	nature	0;0	1;6	4	concept	nature_W5
31	manifestation	0;0	2;8	7	concept	manifestation_A10+
32	notion	0;0	1;6	4	concept	notion_X4.1
33	place	0;0	1;5	3	place	place_M2
34	ambiguity	0;0	1;3	1	concept	ambiguity_Q3
35	basic	0;0	1;3	1	concept	basic_A6.2+
36	dichotomy	0;0	1;3	1	concept	dichotomy_A6.1-
37	time	0;0	1;5	3	time	time_T1
38	allow	0;0	1;2	0	verb	allow_S7.4+
39	argue	0;0	1;2	0	verb	argue_Q2.1
40	assume	0;0	1;1	0	verb	assume_X2.1
41	dialectic	0;0	1;1	0	concept	dialectic_Z99

The domains shown in table 69 appear in the HSEC Encyclopaedia Britannica (ordered by their summarized “sum” values):

Table 69: Categories in the HSEC Encyclopaedia Britannica in descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	307
Place	14
Time	3
Attribute	0
Person	0
Verb	0

The “concept” domain is exceptionally dominant in the HSEC Encyclopaedia Britannica entry. As a result, the microanalysis will also mainly concentrate on the “concept” domain. Yet, the “person” and “verb” domains also include unusual words and will, therefore, also be considered. The other “domain” categories are left out as they include only a few relevant words that, in addition, only appear with a very low frequency. The only exception is the term “world” in the “place” domain that will be discussed later in the context of the HSEC Eliade corpus.

Concept

With a “sum” value of 100, the most important word in the “concept” domain is “sacred.” Thus, “sacred” is also the word in the synoptic table with the highest frequency. The high value of “sacred” is a result of its high absolute frequency value in the word frequency list. Since the entire article focuses on this term, a microanalysis does not add much to the overall discussion beyond the points that have already been discussed in the previous macroanalyses. Consequently, this examination will concentrate on three other notions in the “concept” domain, namely “religion,” “power,” and “life.” The microanalysis includes three examples from the “concept” domain because the “concept” domain is exceptionally dominant in the overall HSEC Encyclopaedia Britannica entry.

(RELIGION)

It was during the first quarter of the 20th century that the concept of the sacred became dominant in the comparative study of religions. Since the first quarter of the 20th century many historians of religions have accepted the notion of the sacred and of sacred events, places, people, and acts as being central in religious life if not indeed the essential reality in religious life.

Discussion

The fact that the word “religion” frequently appears in the entry of the HSEC Encyclopaedia Britannica is due to the entire lemma’s emphasis on the importance of the notion of *holy* in the context of religion and the academic field of the study of religion. Thus, other religious and non-religious dimensions of *holy* in contemporary societies and academic discussions, highlighted in Chapter 2, are not further elaborated upon in the HSEC Encyclopaedia Britannica entry. On the contrary, despite the attempt to include different approaches and ideas to define the “sacred” in the nineteenth and particularly twentieth centuries CE, the assertion by Söderblom is adopted that the notion of *holy* and “religion” have the potential to replace each other. A vivid example of this tendency is found among the last sentences in the summary section of the lemma.

Throughout the past hundred years a number of philosophers and social scientists have asserted the disappearance of the sacred and predicted the demise of religion.

The attention given to the academic study of religion in the HSEC Encyclopaedia Britannica entry also accounts for the appearance of persons such as Rudolf Otto, Brede Kristensen, or Martin Heidegger in the synoptic table. In this case, the importance of the “person” domain is not connected to “holy persons” but due to influential scholars who are frequently referred to in the HSEC Encyclopaedia Britannica. The latter are almost always scholars from the study of religion (or related fields such as philosophy) who often have a background in *Religionsphänomenologie* (such as Mircea Eliade and Gerardus van der Leuw) or are closely associated with the works by Rudolf Otto.

(POWER)

Sacred, the power, being, or realm understood by religious persons to be at the core of existence and to have a transformative effect on their lives and destinies. Because the sacred contains notions both of a positive, creative power and a danger that requires stringent prohibitions, the common human reaction is both fear and fascination.

Discussion

The importance of “power” in the “concept” domain of HSEC Encyclopaedia Britannica hints at the influence of different academic traditions and schools of thought on the Encyclopaedia Britannica entry. Most prominent among these are the *Religionsphänomenologie* and the works by Rudolf Otto. This tendency is particularly evident in the first sentences quoted at the beginning of the lemma, where the “sacred” is defined as a “power” that influences the “life and destiny” of “religious persons.”

The prevalence of the attribution of “power” or even “being” to a concept such as *holy* is of particular interest when compared with the other examinations. In the Armenian sources, *surb* was mostly used as an attribute to characterize persons, objects, or concepts. However, the use of *surb* with “spirit” (Arm. *hogi* or *ays*) or as an attribute for persons (“saints”), who were themselves presented as powerful actors, was attested in the Armenian texts as well, thereby granting *surb* an aura of might and power. In parts of the Twitter data and in the HSEC Encyclopaedia Britannica entry, the attribute of *holy* seems to have transformed into an independent agent, either in the form of the “Holy Spirit” or in a more abstract sense as the power of *holy*. The following examinations of HSEC Eliade and HSEC Otto will demonstrate how deeply the idea of *holy* as a powerful independent agency is entangled with academic discourses on *holy*.

(LIFE)

The image and the temple are, in traditional societies, not simply productions by individual artists and architects; they are reflections of the sacred essence of life, and their measurements and forms are specified through sacred communication from the divine sphere.

Or, a river, such as the Nile for the ancient Egyptians and the Ganges for the Hindu, gave witness to the power of life incarnated in geography.

Discussion

The word “life” appears particularly often in the word frequency list. However, it is not part of the top 15 entries in the collocation analysis tables. The word “life” appears only a few times in the direct vicinity of “sacred” or “holy.” One example of

the direct connection appears in the first quote in the form of the “sacred essence of life.” Yet, the word “life” occurs frequently in the context of “power,” as a collocation analysis of “life” with *LancsBox* reveals.⁴⁷ “Power,” in turn, is one of the major concepts used to characterize *holy* in this article. Thus, the importance of the word “life” is indeed connected to *holy* since – according to the HSEC Encyclopaedia Britannica entry – the “power of life” or “the power to give/renew life” is very often assumed to be a key aspect of *holy*.

Person

The most prominent name in the “person” domain is the name “Otto.” The name “Otto” is the last name of Rudolf Otto, an eminent scholar in the study of religion, whose contributions strongly influenced the lemma in the HSEC Encyclopaedia Britannica. Rudolf Otto, as well as the many other persons mentioned in the HSEC Encyclopaedia Britannica who mostly have an academic background, is not called “holy” (or “sacred”) himself but merely related to the notion by having significantly contributed to its contemporary understanding. The influence of their works and the fact that certain persons are quoted (whereas others are not) is central for the lemma and for the understanding of *holy* in the HSEC Encyclopaedia Britannica. It might also be taken as an indication of their significance for academic and non-academic debates in the context of *holy*. Since the influence of scholars with a background in *Religionsphänomenologie* and Rudolf Otto has already been pointed out in the previous part, no other examples will be provided in this section.

Verb

The most significant verbs in the “verb” domain are the verbs “to appear” and “to manifest.”⁴⁸ The verb “to manifest” does not appear in the synoptic table because it was put together with the noun “manifestation” that has a “sum” value of 7 in the “concept” domain.

The sacred appears in myths, sounds, ritual activity, people, and natural objects. Another significant ambiguity is that the sacred manifests itself in concrete forms that are also profane.

Discussion

That the *holy* appears or manifests itself is another strong proof of the already mentioned observation that *holy* in the HSEC Encyclopaedia Britannica entry is assumed to be an independent and powerful entity that may or may not manifest

47 (6-LL; L3-R3, C2-NC2; function words removed).

48 Note that all the verbs in the “verb” domain have a “sum” value of 0. Thus, the verbs that appear to be most interesting have manually been selected.

itself in different situations and forms. Thus, both verbs “to appear” and “to manifest” belong together as they refer to the same potential of the *holy*. The idea that the *holy* appears or manifests itself is central for the school of *Religionsphänomenologie* and particularly present in the works by Mircea Eliade, among others. In Eliade’s influential work, “The Sacred and the Profane,” the sacred’s potential to appear or manifest itself in any given object or situation and is called “hierophany.” Thus, the verb category is another indication of the significant influence of certain schools of thought on the overall entry in the HSEC Encyclopaedia Britannica and thus on the broader meaning of *holy* as well.

5.2.8.5 Final Discussion

The macro- and microanalyses of the HSEC Encyclopaedia Britannica entry have revealed the following results.

- With a few exceptions, the notion of *holy* is primarily applied in the context of non-religious and religious concepts, albeit to a lesser degree. This finding demonstrates that contemporary discussions on *holy* primarily focus on a conceptual and often very abstract understanding of *holy* related to very general terms such as “existence,” “life,” or “power.” Yet, the importance of individual *holy* persons, times, and objects is still present (see the example sentences in the microanalysis). However, these “manifestations” of the *holy* are usually taken as examples or concrete instantiations of more abstract ideas that are in the center of the overall discourse.
- The notion of *holy* appears as an independent and powerful entity that has the potential to change people’s life and destiny. This particular understanding of *holy* as a powerful entity is based on broader academic discourses (for instance, on *taboo* or the works by Rudolf Otto) and the Christian concept of the “holy spirit.”
- Even though no examples were discussed in the microanalysis, the appearance of “profane” and terms such as “ambiguity” or “dichotomy” also indicates the importance of *sacred* and *profane* concepts of *separation-binding* discussed in Chapter 2.
- The importance of several influential tendencies in the study of religion (see also Chapter 2) was perceivable in the mention of scholars in the synoptic table and in the keywords that belong to individual schools of thought such as “manifestation” (Eliade) or the above mentioned “sacred and profane” distinction (Durkheim).

5.2.9 HSEC Eliade

5.2.9.1 Keyword Analysis

There will be no keyword analysis due to the lack of an appropriate reference corpus for the book “The Sacred and the Profane” in the HSEC Eliade.

5.2.9.2 Word Frequency List (AF)

The HSEC Eliade data consists of 46,700 running words and 5,446 types (stop words included). The 15 most common words (“types”) in the HSEC Eliade corpus (no stop words) are shown in the following table and bar plot. The word “holy” only occurs 17 times in the HSEC Eliade (“holiness” appears only once). The reason for this lies in the “The Sacred and the Profane” emphasis on the word field around “sacred.” Yet, since the word “holy” appears and is mostly employed as a synonym for “sacred,” the following collocation analysis and synoptic table include the collocation analysis of both “holy” and “sacred.” But first, the word frequency list of HSEC Eliade and the corresponding bar plot (see the corresponding figure 56) are presented and discussed.

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The RF_1 is calculated based on the 46,700 running words in the HSEC Eliade (including stop words). The RF_2 is calculated based on the total absolute frequency of the top 15 words (without stop words), which is 2,506.

- Words related to “Places” (world, cosmic, cosmos, earth, space) (AF 759, RF_1 1.6%, RF_2 30.3%)
- Words related to non-religious “Concepts” (man, life, human, new, existence) (AF 743, RF_1 1.6%, RF_2 29.6%)
- Words related to religious “Concepts” (religious, sacred, god) (AF 671, RF_1 1.4%, RF_2 26.8%)
- Words related to “Time” (time) (AF 240, RF_1 0.5%, RF_2 9.6%)
- Misc. (hence) (AF 93, RF_1 0.2%, RF_2 3.7%)

Table 70: Most frequent words (top 15) in HSEC Eliade.

Type	Frequency
world	336
religious	281
man	280
sacred	243
time	240
god	147
life	146
cosmic	117
human	113
cosmos	109
new	103
space	103
existence	101
earth	94
hence	93

Unlike in the examinations of the other sources so far, the “space” (“place”) domain is very present in the whole text, at least according to the frequency list. The non-religious concepts are similar to what the analysis of the other sources revealed, particularly the observations in the HSEC Encyclopaedia Britannica entry. This confirms the assumption that the HSEC Encyclopaedia Britannica entry is influenced by works such as Eliade’s “The Sacred and the Profane.”

5.2.9.3 Collocation Analysis

Holy

The top 15 entries of the HSEC Eliade collocation analysis of the term “holy” (6-LL, L2-R2, C1-NC1; function words removed) can be found in the following table. The term “holy” only appears 17 times in the overall HSEC Eliade (RF 0.033%). The visualization is displayed in the corresponding figure 57.

Figure 56: Most frequent words (top 15) and their absolute frequencies in HSEC Eliade.

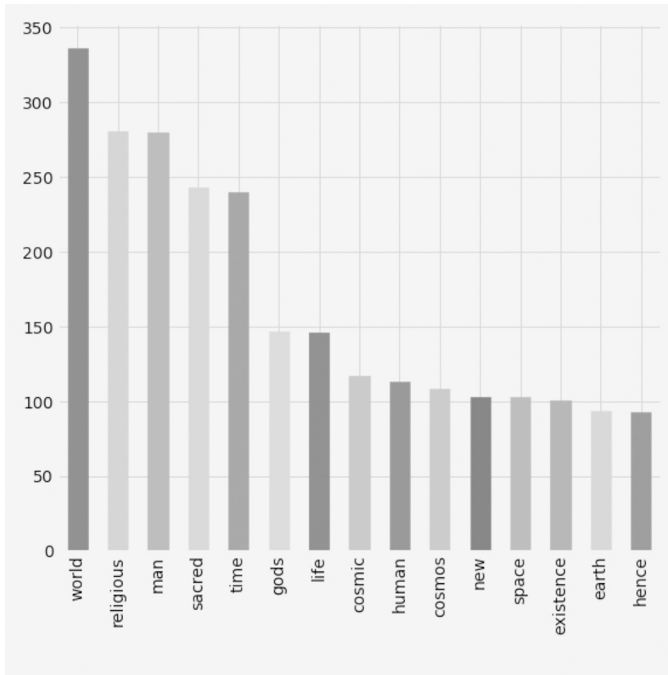
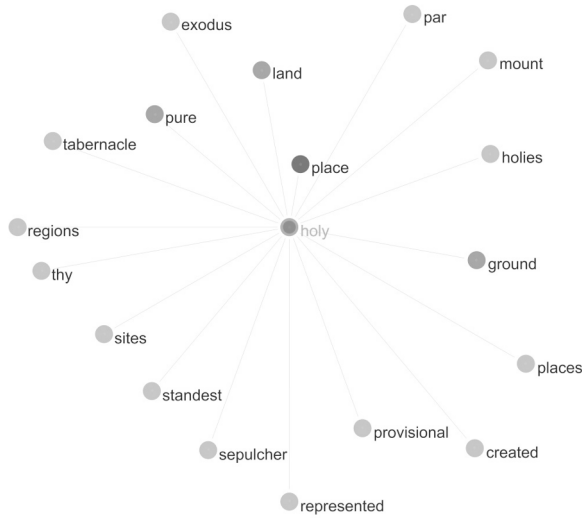


Table 71: Top 15 words collocation analysis “holy” (6-LL, L2-R2, C1-NC1; function words removed) in HSEC Eliade.

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	place	33.74	4	70
2	r	land	22.39	2	9
3	l	pure	20.44	2	14
4	r	ground	18.72	2	21
5	r	holies	15.9	1	1
6	l	provisional	15.9	1	1
7	r	sites	15.9	1	1
8	l	standest	15.9	1	1
9	r	exodus	13.12	1	2
10	r	sepulcher	13.12	1	2
11	r	par (excellence)	12.08	1	3
12	r	tabernacle	11.4	1	4

13	l	thy	11.4	1	4
14	r	mount	10.49	1	6
15	l	regions	9.02	1	12

Figure 57: Top 15 types collocation analysis “holy” (6-LL, L2-R2, C1-NC1; function words removed) in HSEC *Eliade* (graph) (c) LancsBox.



The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 21).

- Words related to “Places” (place, land, ground, sites, mount, regions) (AF 11, RF 52%)
- Words related to religious and religioid “Concepts” (pure, holies, exodus, sepulcher, tabernacle) (AF 6, RF 28.6%)
- Misc. (provisional, standest, par (excellence), thy) (AF 4, RF 19.4%)

The results of the collocation analysis of the term “holy” are somewhat unexpected in view of what has been stated so far regarding the various use cases of “holy” and “sacred.” The most frequent group that appears in the context of “holy” are terms related to spatial concepts. So far, we had noted the following tendency: the term “sacred” is related to spatial concepts, whereas the word “holy” mostly appears in the context of more abstract (religious) concepts. It should, however, be noted that the general focus on spatial concepts in “The Sacred and the Profane” might account for the high frequency of “holy” in this context. As can be observed in the following collocation analysis of “sacred,” spatial terms are also very present in the examination of the use of “sacred” and in the word frequency list. Thus, the occurrence of spatial concepts in the proximity of “holy” might be interpreted as another indication that both terms are often used interchangeably.

Also worth mentioning is the appearance of terms related to purity in the context of “holy.” However, in this case, the notion of purity is related to an aspect which has not been hitherto encountered, namely death (except for martyrs and their burial contexts in the Armenian data). These terms occur neither in the word frequency list nor in the collocation analysis of “sacred.” Thus, they seem to be significant for the meaning of “holy” in “The Sacred and the Profane” by Eliade.

Sacred

The top 15 entries of the HSEC Eliade collocation analysis of the term “sacred” (6-LL, L2-R2, C3-NC3; function words removed) can be found in the following table. The corresponding visualization is displayed in figure 58. The term “sacred” appears 243 times in the overall HSEC Eliade (RF 0.52%).

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 146).

Table 72: Top 15 words collocation analysis “sacred” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Eliade.

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	space	232.16	34	103
2	r	time	161.08	33	240
3	r	pole	57.52	8	20
4	r	calendar	54.21	6	8
5	r	history	53.58	11	77
6	r	lodge	34.47	4	6
7	l	experience	32.64	8	82
8	r	profane	32.25	8	84
9	r	realities	28.69	4	10
10	m	strong	27.74	4	11
11	r	dimension	24.80	4	15
12	l	world	22.48	11	336
13	r	stone	21.82	4	21
14	m	tree	18.80	4	30
15	r	manifests	18.75	3	11

- Words related to “Places” (space, world, lodge, dimension) (AF 53, RF 36.3%)
- Words related to “Time” (time, calendar, history) (AF 50, RF 34.3%)
- Words related to non-religious or religioid “Concepts” (profane, experience, realities, strong) (AF 24, RF 16.4%)
- Words related to “Objects” (stone, tree, pole) (AF 16, RF 11%)
- Verbs (manifest) (AF 3, RF 2.1%)

Interestingly, there are no definite religious concepts in the collocation analysis of “sacred” in HSEC Eliade (with the slight exception of “profane”). The “place” dimension is equally important in the collocation analysis of both “holy” and “sacred.” However, the word “holy” seems to be more closely related to aspects of purity and death. As for the word “sacred,” it is more closely connected to “time” and to “objects.”

Figure 58: Top 15 types collocation analysis “sacred” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Eliade (graph) (c) LancsBox.



5.2.9.4 Synoptic Table – Sentences Selection and Discussion (Microanalysis)

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses and their corresponding values and annotation has been created. The results of the HSEC Eliade collocation analysis on “sacred” are also considered in the synoptic table together with the results of the collocation analysis of “holy.” Firstly, this decision has been made due to the overall low frequency of “holy” in the HSEC Eliade, which would not yield sufficient data for analysis. Secondly, the term “sacred” is often used in a very close relationship with other words such as “holy” in Eliade’s “The Sacred and the Profane.” All these words thus seem to refer to a common semantic field of *holy*.

Table 73: Synoptic table of the HSEC *Eliade*.

id	lemma	wf_t	cf_t	sum	domain	semantic_field
1	world	1;336	1;11	100	place	world_W1
2	man	1;280	0;0	80	person	man_S2.2m
3	religion	1;281	0;0	80	concept	religion_S9
4	sacred	1;243	0;0	69	concept	sacred_S9
5	time	1;240	1;33	78	time	time_T1
6	god	1;147	0;0	42	concept	god_S9
7	life	1;146	0;0	41	concept	life_L1+
8	human	1;113	0;0	32	concept	human_S2
9	new	1;103	0;0	29	concept	new_T3-
10	cosmos	2;226	0;0	65	place	cosmos_W1
11	space	1;103	1;34	39	place	space_N3.6
12	existence	1;101	0;0	28	concept	existence_A3+
13	earth	1;94	0;0	26	place	earth_W3
14	hence	1;93	0;0	26	misc	hence_A2.2
15	place	0;0	1;4	0	place	place_M2
16	land	0;0	1;2	0	place	land_W3
17	thy	0;0	1;1	0	person	thy_Z8
18	mount	0;0	1;1	0	place	mount_T2+
19	pure	0;0	1;2	0	concept	pure_A5.4+
20	ground	0;0	1;2	0	place	ground_M6
21	exodus	0;0	1;1	0	concept	exodus_M1
22	tabernacle	0;0	1;1	0	object	tabernacle_S9/H1
23	holy	0;0	1;1	0	concept	holy_S9
24	provisional	0;0	1;1	0	attribute	provisional_T1.3-
25	sepulcher	0;0	1;1	0	place	sepulcher_H1/L1-
26	site	0;0	1;1	0	place	site_M7
27	standest	0;0	1;1	0	verb	standest_l2.2/H5
28	par (excellence)	0;0	1;1	0	concept	excellence_A5.1+++
29	region	0;0	1;1	0	place	region_M7
30	pole	0;0	1;8	2	object	pole_O2
31	calendar	0;0	1;6	1	time	calendar_O2/T1.3
32	history	0;0	1;11	2	time	history_T1.1.1

33	lodge	0;0	1;4	0	place	lodge_H4
34	experience	0;0	1;8	2	concept	experience_A2.1+/A3+
35	profane	0;0	1;8	2	concept	profane_S9
36	reality	0;0	1;4	0	concept	reality_A3+
37	strong	0;0	1;4	0	attribute	strong_S1.2.5+
38	dimension	0;0	1;4	0	concept	dimension_A4.1
39	stone	0;0	1;4	0	object	stone_O1.1
40	tree	0;0	1;4	0	object	tree_L3
41	manifest	0;0	1;3	0	verb	manifest_A10+

As explained in Chapter 3, the selection of example sentences for the micro-analysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *holy* appears in the proximity of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

The following domains appear in the HSEC Eliade (ordered by their summarized “sum” values):

Table 74: Categories in the HSEC Eliade in descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	325
Place	230
Time	81
Person	80
Misc	26
Object	2
Attribute	0
Verb	0

Concept

With a “sum” value of 80, the most significant word in the “concept” domain is the term “religion” (mostly in the form of “religious”).

In other words, for those who have a religious experience all nature is capable of revealing itself as cosmic sacrality. (Eliade and Willard R. Trask 1959, 12)

Our primary concern is to present the specific dimensions of religious experience, to bring out the differences between it and profane experience of the world. (Eliade and Willard R. Trask 1959, 17)

Discussion

The word “religious”⁴⁹ is among the top entries in the word frequency list. There is another religious concept among the top entries in the word frequency list, namely “god.” This is not surprising since the whole work “The Sacred and the Profane” by Eliade is primarily concerned with the topic of religion and the subcategories of “religious man,” “religious experience,” and “religious behaviour.” This assumption is also reflected in a collocation analysis of “religious” with *LancsBox*.⁵⁰

This collocation analysis yields that the words “man,” “experience,” and “behaviour” are among the top entries. They thus underline their frequent appearance in the vicinity of “religious.” Since the primary aim of Eliade’s book is to examine the “Sacred and the Profane,” one may surmise that these two words are understood as notions related to “sacred” and *holy*. This is best exemplified by the last sentence where the “religious experience” is contrasted with a “profane experience of the world,” thereby underlining the partial identification of both *holy* (“sacred”) with religion.

Besides the religious concepts, there are other non-religious concepts with a religioid character, such as “life,” “human,” and “new,” that were also encountered in other sources such as the HSEC Encyclopaedia Britannica.

Place

With a “sum” value of 100, the most important word in the “place” domain is the “world.” The term “world” is thus also the most important word in the overall synoptic table.

This is as much as to say that religious man can live only in a sacred world, because it is only in such a world that he participates in being, that he has a real existence. (Eliade and Willard R. Trask 1959, 64)

Discussion

Even though the word “world” is the most significant term in the “place” domain, there are other frequently used words such as “cosmos” (“sum” value of 65) and

49 That is used as an adjective, but due to its close relationship with “religion,” it is subsumed under the “concept” domain.

50 (6-LL; L2-R2; C3-NC3; no function words).

“space” (“sum” value of 39). These further point to the idea that the notion of *holy* (“sacred”) in Eliade’s “Sacred and the Profane” is not exclusively related to specific spaces such as monasteries (see the Armenian sources), but also to the entire world and cosmos. The *holy* in Eliade’s “Sacred and the Profane” therefore has a much more abstract as well as a universal dimension since it comprises the entire cosmos and world. This hypothesis is very much in line with the general standpoint of “The Sacred and the Profane,” which is primarily concerned with a universal idea of the “sacred” (and the “profane”) beyond historical and cultural boundaries.

Time

With a “sum” value of 78, the central word in the “time” domain is “time.”

The experience of sacred time will make it possible for religious man periodically to experience the cosmos as it was in principio, that is, at the mythical moment of Creation. (Eliade and Willard R. Trask 1959, 65)

Discussion

The word “time” appeared in many of the other HSEC sources as well. Since the word “time” is also used in many expressions, such as “it is time to (...),” it was usually not regarded as relevant for the understanding of *holy* and was left out of the microanalysis. This situation is different in the case of Eliade’s “The Sacred and the Profane.” The topic of “sacred time” holds a crucial position in the entire work. Among others, the title of the second chapter, “Sacred Time and Myths,” allows grasping its importance.

The attribution of “sacred” to “time” is comparable to the attribution of “sacred” to “world” or “cosmos” in the “place” domain. It suggests that from a spatial and temporal perspective, the entire world can be regarded as “sacred” (*holy*). This observation further underlines the previously noted abstract and comprehensive character of *holy* in “The Sacred and the Profane.”

Person

With a “sum” value of 80, the most significant word in the “person” domain is the “man.”

The above example helps us to understand the perspective adopted by the man of the archaic societies; for him, the whole of life is capable of being sanctified. (Eliade and Willard R. Trask 1959, 167)

Discussion

A look at the collocation analysis and the word frequency table in this chapter shows that “man” is rarely directly connected to terms such as “sacred” and “holy.” Instead,

the term “man” is first and foremost present in the synoptic table due to its high frequency in the overall work of “The Sacred and the Profane.”

A separate collocation analysis of “man” reveals that it is most commonly used together with attributes such as “(non-)religious,” “modern,” “primitive,” and “profane.”⁵¹ Even though “man” is rarely directly connected to “sacred,” an indirect connection demonstrates the pertinence of this term for the use of words related to “sacred” (*holy*) in “The Sacred and the Profane.” This connection is reflected in the title of the fourth chapter, “Human Existence and Sanctified Life,” that illustrates the central role of “man” in a sacred world and in sacred time. It is the religious/primitive man who finds his or her entire life sanctified or embedded in a sacred world (consisting of sacred space and time).

This indirect connection between “man” and *holy* constitutes an essential contrast to the use of terms such as *surb* or “holy” in the “person” domain in many of the other sources. Instead of using the terminology around *holy* to define specific groups of persons by ascribing them attributes such as “holy,” “chaste,” or “virtuous,” or an established title (the “saints”), the terminology is here applied as a signifier to characterize the entire world, life, and time in which religious/primitive men are living. Thus, “sacred” still has a differentiating function, but it now separates two world views, namely that of religious/primitive man from that of a modern/profane man.

Also, it has been pointed out in the previous part on “concepts” that the term “religious” is often identified as “sacred,” although the use contexts of both terms might differ from one another. The partial similarity is perceivable in the distinction between a “religious man” and a “profane man,” where the latter is defined as being part of a sacred cosmos and time.

Object

With a “sum” value of 2, the most important word in the “object” domain is the “pole.” Other words that appear in this domain are often connected to natural objects, such as “tree” or “stone.”

But he felt the need to live at the Center always — like the Achilpa, who, as we saw, always carried the sacred pole, the axis mundi, with them, so that they should never be far from the Center and should remain in communication with the supraterrrestrial world. (Eliade and Willard R. Trask 1959, 43f.)

51 (6-LL; L2-R2; C2-CN2; function words removed).

Discussion

The “pole” is a symbolic object⁵² that symbolizes the center of the (sacred) universe (*axis mundi*) and thus holds an essential function for religious/primitive men. It also illustrates the variety of objects that can represent the “sacred” in Eliade’s “The Sacred and the Profane.” The aforementioned close connection between the world view of religious/primitive man and the “sacred” in the form of various objects and concepts is also expressed via natural objects, such as stones and trees. These objects underline the archaic and premodern state of the “sacred” as opposed to the “profaneness” of modernity.

Miscellaneous, Verbs, and Attributes

The words in the remaining three domains are not discussed with a microanalysis because they are not relevant for the understanding of *holy*.

There is one exception, though, namely the verb “to manifest” in the “verb” domain. The importance of “to manifest” or “manifestation” is related to the so-called “hierophanies” in “The Sacred and the Profane.” These signify that the *holy* manifests itself in different forms in the history of religion and in the life of religious man. Examples of objects that are particularly relevant in the context of the “hierophanies” have been discussed in the paragraph dedicated to the “object” domain. The importance of words related to “(to) manifest” in the HSEC Encyclopaedia Britannica bespeaks the influence of the views Eliade develops in “The Sacred and the Profane.” As discussed in the microanalysis of HSEC Encyclopaedia Britannica, the fact that the *holy* (“sacred”) manifests itself ascribes an agency to this notion. It thus goes far beyond the meaning and application of *holy* in the sense of an attribute highlighting specific aspects of persons, concepts, or objects.

5.2.9.5 Final Discussion

The analysis of the HSEC Eliade has revealed the following results.

- The terminology around “sacred” and “holy” frequently appeared in the context of abstract and comprehensive words related to time and space, such as “world,” “cosmos,” and “time.” The “time” and “space” categories primarily serve to separate the “sacred world” and “sacred time” of the primitive/religious man from that of the modern/profane man. The generic character of the terminology that was underlined by concepts such as “life” fits with the universal interest of Eliade’s thoughts that goes beyond single objects, persons, or concepts.

52 For instance, in the form of a tent pole.

- Unlike many of the other sources, the synoptic table of the HSEC Eliade also includes natural objects. These objects mostly serve as examples or instantiations of the aforementioned abstract spheres of “space” and “time” and show possible objects in which the *holy* (“sacred”) can reveal itself (“hierophany”).
- The influence of “The Sacred and the Profane” on later works in the study of religion such as the entry in the HSEC Encyclopaedia Britannica was visible in the parallels between words in both synoptic tables and the corresponding microanalyses (for instance, the similar importance of verbs such as “to manifest”).
- It is remarkable that even though the term “man” plays an important role, the idea of “holy men” in the sense that it was used in the Armenian sources was only of minor importance in the HSEC Eliade. Instead, the term “man” was related to the religious/primitive and the profane/modern man and their relation to the concept of the *holy* (“sacred”) that partially coincided with “religion.”
- Words related to purity did not show up among the top entries of the collocation analysis of “sacred” but only in the context of the collocation analysis of “holy.” Yet, due to the overall low frequency of “holy,” this result needs to be treated with caution.

5.2.10 HSEC Otto

5.2.10.1 Keyword Analysis

There will be no keyword analysis due to the lack of an appropriate reference corpus for the book “The Idea of the Holy” in the HSEC Otto.

5.2.10.2 Word Frequency List (AF)

The HSEC Otto data consists of the cleaned text of “The Idea of the Holy” and holds 84,824 running words and 9,182 types (stop words included).⁵³ The 15 most common words (“types”) in the HSEC Otto (no stop words) are shown in the following table and bar plot (see the corresponding figure 59).

53 Note that the digitized text of Otto’s “The Idea of the Holy” still includes a lot of noisy data, among others, due to wrongly transcribed quotations, footnotes, words written in non-Latin alphabets, etc.

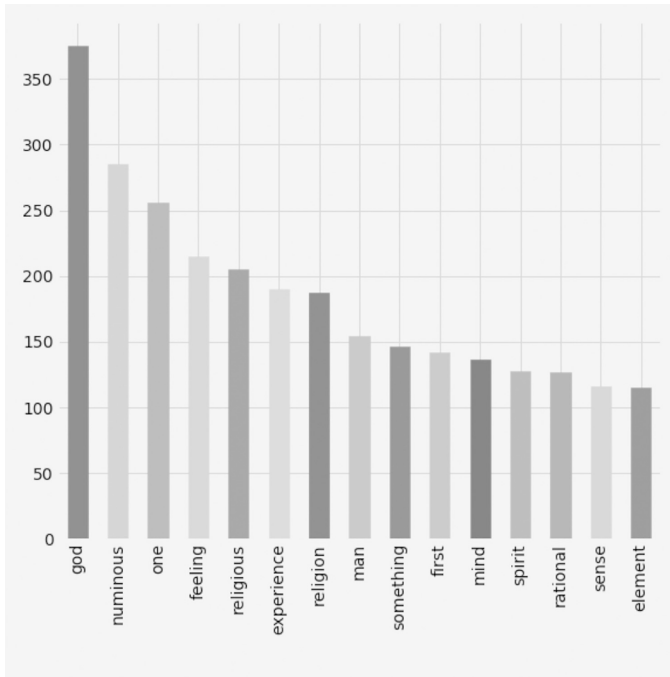
Table 75: Most frequent words (top 15) in HSEC Otto.

Type	Frequency
god	375
numinous	285
one	256
feeling	215
religious	205
experience	190
religion	187
man	154
something	146
first	142
mind	137
spirit	128
rational	127
sense	116
element	115

The words in the table can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The RF_1 is calculated based on the 84,824 running words in the HSEC Otto (including stop words). The RF_2 is calculated based on the overall absolute frequency of the top 15 words (without stop words), which is 2,778.

- Words related to religious “Concepts” (god, numinous, religious, religion, spirit) (AF 1,180, RF_1 1.4%, RF_2 42.5%)
- Words related to various non-religious “Concepts” (one, man, something, first, rational, element) (AF 940, RF_1 1.1%, RF_2 33.8%)
- Words related to non-religious “Concepts” mostly representing experiences and sensations (feeling, experience, mind, sense) (AF 658, RF_1 0.8%, RF_2 23.7%)

Figure 59: Most frequent words (top 15) and their absolute frequencies in HSEC Otto.



A closer look at the word frequency list reveals that the word “numinous” is among the top entries with an absolute frequency of 285. Rudolf Otto defines the term “numinous” as follows:⁵⁴

Any one who uses it today does undoubtedly always feel “the morally good” to be implied in “holy;” and accordingly in our inquiry into that element which is separate and peculiar to the idea of the holy it will be useful, at least for the temporary purpose of the investigation, to invent a special term to stand for “the holy” minus its moral factor or “moment;” and, as we can now add, minus its “rational” aspect altogether. (Otto and Harvey 1923, 6)

This “special term” is called the “numinous” and plays a central role in “The Idea of the Holy.” According to Otto, the numinous is meant to represent an original layer

54 A word derived from the Latin term *numen*, see Otto and Harvey (1923, 7). For the meaning and importance of these terms in Otto’s “The Idea of the Holy,” please see the corresponding chapters in the book, such as Otto and Harvey (1923, 4ff.) and Alles (2010, 206ff.).

of the meaning of *holy* that is still part of its contemporary meaning, namely the idea of *holy* without the moral factor.⁵⁵ Due to the high frequency of this term, its overall importance for “The Idea of the Holy,” and its close relationship to the word “holy,” the numinous is also considered in the collocation analysis (together with “holy”), the final synoptic table, and the microanalysis.

5.2.10.3 Collocation Analysis

Holy

The top 15 entries of the HSEC Otto collocation analysis on the terms “holy” (and “holiness”) (6-LL; L2-R2; C3-NC3; function words removed) can be found in the following table. The visualization is displayed in the corresponding figure 60. The term “holy” appears 124 times in the HSEC Otto.⁵⁶ Unlike most of the other sources, the term “holiness” is also considered in this collocation analysis as it has a relatively high frequency (AF 38) as well.

The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 80).

- Words related to non-religious “Concepts” (category, one, meaning, word, experience, something, awe) (AF 33, RF 41.25%)
- Words related to religious “Concepts” (holy, sacred, spirit, sanctity) (AF 32, RF 40%)
- Verbs (manifest, made, becomes) (AF 12, RF 15%)
- Words related to “Persons” (men) (AF 3, RF 3.75%)

55 That can further be split into the *mysterium fascinans* and *tremendum*. For a concise overview of key ideas in Otto’s “The Idea of the Holy,” see Alles (2010, 206ff.).

56 The word “holy” is not part of the word frequency list as the term “holy” has usually been manually removed from word frequency lists. Thus, removing “holy” from the word frequency list in this section was mainly done due to coherency reasons.

Table 76: Top 15 words collocation analysis “holy” and “holiness” (6-LL; L2-R2; C3-NC3; function words removed) in HSEC Otto.

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	m	holy	89.45	14	124
2	r	sacred	81.44	8	14
3	l	category	35.45	5	31
4	r	spirit	34.08	7	128
5	r	manifest	31.15	4	18
6	r	one	30.25	8	256
7	r	sanctity	26.17	3	9
8	r	made	21.69	4	55
9	m	becomes	20.99	4	60
10	l	meaning	19.20	4	75
11	l	word	17.76	4	90
12	l	experience	17.19	5	190
13	r	men	15.21	3	49
14	l	something	14.03	4	146
15	r	awe	13.53	3	65

The collocation analysis of “holy/holiness” reveals a close relationship between “sacred” and “holy/holiness” in the HSEC Otto. Due to the relatively low frequency of “sacred” (AF 14), the term “sacred” will not be examined in a separate collocation analysis. The word “holy” mainly appears in the collocation analysis because the term “holiness” was also among the search terms. The terms “holy” and “holiness” are primarily connected to religious and non-religious concepts (over 80%).

Numinous

The top 15 entries of the HSEC Otto collocation analysis on the term “numinous” (6-LL, L2-R2, C3-NC3; function words removed) can be found in the following table. The visualization is displayed in the corresponding figure 61. The term “numinous” appears 285 times in the overall HSEC Otto.

Figure 60: Top 15 types collocation analysis “holy” and “holiness” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Otto (graph) (c) LancsBox.



Table 77: Top 15 words collocation analysis “numinous” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Otto.

id	pos	coll	stat	freq_(coll)	freq_(corpus)
1	r	consciousness	140.27	21	96
2	r	feeling	118.69	23	215
3	r	character	45.18	9	88
4	r	experience	42.79	11	190
5	r	emotion	35.84	6	37
6	m	something	30.19	8	146
7	r	object	29.58	6	61
8	r	awe	28.81	6	65
9	r	element	27.83	7	115
10	r	impression	25.44	5	47

11	r	atmosphere	23.64	3	8
12	l	non-rational	22.88	6	107
13	l	primal	20.21	3	13
14	l	purely	15.99	3	25
15	r	moment	12.96	3	41

Figure 61: Top 15 types collocation analysis “numinous” (6-LL, L2-R2, C3-NC3; function words removed) in HSEC Otto (graph) (c) LancsBox.



The words from the collocation analysis can be grouped according to the “domain” scheme developed in Chapter 3. Unlike what can be observed in the synoptic tables, words can be part of multiple categories in the following overview. The absolute frequencies below are calculated on the basis of the values in the “freq_(coll)” column in the collocation table (overall words frequency collocation: 120).

- Words related to non-religious “Concepts” with a focus on emotions and cognition (consciousness, feeling, experience, emotion, awe, impression, atmosphere, non-rational) (AF 81, RF 67.5%)
- Words related to other primarily non-religious “Concepts” (character, something, object, element, primal, purely, moment) (AF 39, RF 32.5%)

Interestingly, there are no clearly religious words in the collocation analysis of “numinous.” Instead, the collocation analysis reveals many words related to concepts of which the majority is related to emotions (such as “emotion,” “feeling,” or “atmosphere”) and cognition (such as “experience” or “consciousness”).

5.2.10.4 Synoptic Table – Sentences Selection and Discussion (Microanalysis)

In accordance with the calculations described in Chapter 3, a synoptic table which integrates all the words from the previous analyses as well as their corresponding values and annotation has been created.

The results of the HSEC Otto collocation analysis of “numinous” are also considered in the synoptic table, together with the results of the collocation analysis of “holy” and “holiness.” This decision was made because the “numinous” is, according to Otto’s definition, “the holy minus its moral factor” (Otto and Harvey 1923, 6) and frequently appears in “The Idea of the Holy.”

As explained in Chapter 3, the selection of example sentences for the microanalysis relies upon the “domains” and “sum” categories of the synoptic table. For each “domain,” the lemma with the highest “sum” value is usually selected. However, there might be some exceptions. Subsequently to the selection of the lemma, a corresponding sentence from the text where this lemma appears is chosen and discussed in the section below. In case there are several sentences where *holy* appears in the proximity of the selected lemma, the sentence that offers the most potential for the analysis is manually selected.

Table 78: Synoptic table of the HSEC Otto.

id	lemma	wf_t	cf_t	sum	domain	semantic_field
1	god	1;375	0;0	95	concept	god_S9
2	numinous	1;285	0;0	72	concept	numinous_S9
3	one	1;256	1;8	67	concept	one_N1
4	feeling	1;256	1;23	71	concept	feeling_X2.1
5	religion	2;392	0;0	100	concept	religion_S9
6	experience	1;190	2;16	52	concept	experience_A2.1+/A3+
7	man	1;154	1;3	39	person	man_S2.2m
8	something	1;146	2;12	40	concept	something_Z8
9	first	1;142	0;0	35	concept	first_N4
10	mind	1;137	0;0	34	concept	mind_E6+
11	spirit	1;128	1;7	34	concept	spirit_S9
12	rational	1;127	0;0	31	concept	rational_S1.2.6+
13	sense	1;116	0;0	28	concept	sense_A4.1
14	element	1;115	1;7	30	concept	element_A4.1
15	holy	0;0	1;14	2	concept	holy_S9
16	sacred	0;0	1;8	1	concept	sacred_S9
17	category	0;0	1;5	0	concept	category_A4.1
18	manifest	0;0	1;4	0	verb	manifest_A10+
19	sanctity	0;0	1;3	0	concept	sanctity_A11.1+
20	make	0;0	1;4	0	verb	make_A1.1.1
21	become	0;0	1;4	0	verb	become_A2.1+
22	meaning	0;0	1;4	0	concept	meaning_Q1.1
23	word	0;0	1;4	0	concept	word_Q3
24	awe	0;0	2;9	1	concept	awe_E1
25	consciousness	0;0	1;21	4	concept	consciousness_B1
26	character	0;0	1;9	1	concept	concept_X4.1
27	emotion	0;0	1;6	0	concept	emotion_E1
28	object	0;0	1;6	0	concept	object_A6.1-
29	impression	0;0	1;5	0	concept	impression_X2.1
30	atmosphere	0;0	1;3	0	concept	atmosphere_W3
31	non-rational	0;0	1;6	0	concept	non-rational_S1
32	primal	0;0	1;3	0	concept	primal_Z99
33	purely	0;0	1;3	0	concept	purely_A14
34	moment	0;0	1;3	0	time	moment_T1.2

The following domains appear in the HSEC Otto (ordered by their summarized “sum” values):

Table 79: Categories in the HSEC Otto in descending order according to their “sum” values in the synoptic table.

Domain	Sum
Concept	698
Person	39
Time	0
Verb	0

Due to the dominance of the “concept” domain and the very few entries in the other domains, the following examination will take into account several entries from the “concept” domain. This provides a more in-depth overview of important use cases of *holy* in the context of various concepts.

Concept

The most important word in the “concept” domain with a “sum” value of 100 is the word “religious.” The word “religious” in the synoptic table includes the appearances of “religion” and that of “religious” in the word frequency analysis. The term “religious” is thus also the most prominent word field in the synoptic table. However, neither the term “religious” nor the term “religion” appear in the collocation analysis, meaning that the terms are of importance in the overall work but only seldom appear in the direct vicinity of “holy,” “holiness,” and “numinous.”

However, as the main topic of “The Idea of the Holy” is the meaning of *holy*, the words in the word frequency list are also informative for the overall understanding of *holy*. A collocation analysis of “religious” with *LancsBox* furthermore shows that the top entries of this query are all part of the synoptic table as well.⁵⁷

The top three words in this collocation analysis of “religious” are “experience,” “consciousness,” and “feeling.” All these words are not only commonly related to “religious” but to “holy” and “numinous” as well. They consequently serve as intermediaries between these two word fields. This also demonstrates the very close relationship between the terminology of *holy* (such as “numinous”), religion, and cognitive as well as emotional concepts of “feeling,” “experience,” and “consciousness.” This tripartite relationship can indeed be seen as a good summary of central ideas in “The Idea of the Holy,” as can be observed in the following two quotes.⁵⁸

57 (6-LL; L2-R2; C3-CN3; no function words).

58 Concerning the importance of emotions and cognitive concepts in Otto’s work, see the influence of Kant and Fries on Rudolf Otto in Alles (2010, 202f.).

The reader is invited to direct his mind to a moment of deeply-felt religious experience, as little as possible qualified by other forms of consciousness. (Otto and Harvey 1923, 8)

This is already to be observed on the lowest and earliest level of the religion of primitive man, where the numinous consciousness is but an inchoate stirring of the feelings. (Otto and Harvey 1923, 26)

Discussion

The two example sentences underline the importance of words such as “experience” or “consciousness” in Otto’s “The Idea of the Holy.” Furthermore, they demonstrate that Otto’s approach to evaluating the meaning of *holy* is closely linked to broader questions of religion(s) and religiosity of primitive men in particular. This interest in a (fictional) *Urzustand* of primitive religiosity was also present in Eliade’s “The Sacred and the Profane.” Contrary to an intuitive assumption that might arise in the context of Otto’s interest in primitive religiosity, he was an opponent of the influential school of *Religionsevolutionismus* (Alles 2010, 201ff.).

Otto’s central hypothesis in “The Idea of the Holy” is that religion, as well as the core meaning of *holy* (in the form of the “numinous”), is closely related to certain types of irrational experiences such as fear or awe. According to Otto, these experiences have always been part of the phenomenon of religion, which is a category *sui generis*, and are not confined to an earlier stage of primitive religion. This opinion is also mirrored in the synoptic table and in the collocation analyses where words such as “awe,” “emotion,” and “non-rational” appear.

In the following, more sentences will be discussed to further illustrate the relationship between *holy* (“numinous”) and cognitive as well as emotional concepts such as “consciousness,” “feeling,” and “awe.”

(CONSCIOUSNESS)

A collocation analysis with *LancsBox* of “consciousness” shows that the term frequently appears in the context of “numinous” and “religious.”⁵⁹ This, again, underlines Otto’s interest in both the meaning of *holy* (“numinous”) itself and the *holy*’s importance for religion. A sentence including the combination of “numinous consciousness” has already been examined. Therefore, the following part only shows and discusses an example of “religious consciousness.”

The intimate interpretation of the non-rational with the rational elements of the religious consciousness, like the interweaving of warp and woof in a fabric, may be elucidated by taking another familiar case, in which a universal human feeling, that of personal affection, is similarly interpenetrated by a likewise thoroughly

59 (6-LL; L2-R2; C3-CN3; no function words).

non-rational and separate element, namely, the sex instinct. (Otto and Harvey 1923, 47)

Discussion

This sentence is another example of an instantiation of Otto's overarching theory of the interweaving of non-rational with rational elements in religion and the *holy*. The "numinous consciousness" thus represents the irrational part of the consciousness, whereas the "religious consciousness" represents the rational part. Usually, both aspects are found together in a religion, just like the category of *holy* is the irrational category of the numinous with the rational moral part.

(FEELING)

A collocation analysis with *LancsBox* of "feeling" shows that the term frequently appears in the context of "numinous" and "religious" as well.⁶⁰

Musical feeling is rather (like numinous feeling) something "wholly other," which, while it affords analogies and here and there will run parallel to the ordinary emotions of life, cannot be made to coincide with them by a detailed point-to-point correspondence. (Otto and Harvey 1923, 50)

In every highly-developed religion the appreciation of moral obligation and duty, ranking as a claim of the deity upon man, has been developed side by side with the religious feeling itself. (Otto and Harvey 1923, 53)

Discussion

In contrast to the use of "religious" and "numinous" in the case of "consciousness," there does not seem to be a huge difference in the usage of "religious" and "numinous" in the context of "feeling." Both words are equally used as attributes to distinguish these "wholly other" feelings from rational or moral ideas.

(AWE)

A collocation analysis with *LancsBox* of "awe" shows that the term appears very often in the context of "numinous" and "religious" as well.⁶¹ Other words that appear in the context of "awe" are "dread," "primal," and "mere."

It comprises, first, a manifestation of the numinous awe, viz. the feeling that the "profane" creature cannot forthwith approach the numen, but has need of a covering or shield against the orgê [Greek term meaning anger, wrath] of the numen. (Otto and Harvey 1923, 56)

60 (6-LL; L2-R2; C3-CN3; no function words).

61 (6-LL; L2-R2; C3-CN3; no function words).

But the object of religious awe or reverence - the tremendum and augustum, cannot be fully determined conceptually; (...) (Otto and Harvey 1923, 61)

Discussion

Similarly to the examination of “feeling,” the combination of “religious awe” is also used almost synonymously with “numinous awe.” Both describe a specific subtype of feelings and experiences that constitute the non-rational part of religion and *holy*.

Person

The most important term in the “person” domain is the word “man” with a “sum” value of 39. A collocation analysis with *LancsBox* of “man” shows that this term is frequently used with attributes such as “primitive” and “natural.”⁶² Similarly to Eliade’s “The Sacred and the Profane,” this indicates that Otto’s “The Idea of the Holy” is also strongly interested in reconstructing religious forms (such as *holy*) and religion in general based on an analysis of primitive man in archaic cultures. This approach was very influential in the study of religion during the nineteenth and early twentieth centuries CE and is commonly attributed to the school of *Religionsevolutionismus*. Yet, as explained above, Rudolf Otto is usually not associated with this school. For Otto, phenomena from the earlier stages of humankind are still present in contemporary religion. Since the use of “man” in “The Idea of the Holy” resembles the use of “man” in Eliade’s “The Sacred and the Profane,” there will not be a discussion of concrete examples in this section.

Verb

The last domain discussed in this section is the “verb” domain and the word “(to) manifest” in particular. The word field around “(to) manifest” was also present in the synoptic tables of HSEC Encyclopaedia Britannica and HSEC Eliade and thus shows that these three works need to be seen in relation to one another. The HSEC Encyclopaedia Britannica draws heavily from HSEC Eliade and from HSEC Otto, thereby demonstrating the great influence of these works on the academic notion and use of *holy* in contemporary discourses inside and outside of academia.

In the case of Luther we find next after this element of “Wrath” the numinous manifesting itself in the set of ideas which we may fairly call those of Job. (Otto and Harvey 1923, 104)

62 (6-LL; L2-R2; C3-CN3; no function words).

Discussion

The idea that “the holy” or “the numinous” manifest themselves is present in both Eliade’s “The Sacred and the Profane” as well as in Otto’s “The Idea of the Holy.”

In the latter, however, the numinous manifests itself primarily in feelings, experiences, and notions instead of objects or “in the world.” Yet, these feelings are driven by concrete phenomena or objects. A telling example in this context is chapter XV in “The Idea of the Holy” on “Its Earliest Manifestations” (Otto and Harvey 1923, 121ff.) where Rudolf Otto traces the manifestation of “the numinous” in notions and religious concepts such as “clean-unclean” or “spirit.”

5.2.10.5 Final Discussion

- The most obvious difference between the use of *holy* (“numinous”) in Otto’s “The Idea of the Holy” compared to the other sources is its very close connection to feelings, experiences, and consciousness.
- Furthermore, the *holy* (“numinous”) is regarded as a religious category *a priori* that cannot be explained or derived from other categories such as morality. Thus, the *holy* (“numinous”) constitutes a universal (religious) state of mind through different numinous experiences and their corresponding feelings such as awe, fear, or the sublime. This universalism of *holy* was also found in Eliade’s “The Sacred and the Profane,” although in a different form. Yet, both Otto and Eliade treat the concept of *holy* (or “numinous” or “sacred”) as a universal phenomenon that goes far beyond the restrictions of an attribute denoting certain (distinguishing) aspects of persons and objects or a position in an ecclesiastic hierarchy. Instead, the *holy* is a category able to explain and define entire fields of human behavior and religion as such. However, Eliade restricts certain forms of the *holy* to an earlier primitive man, whereas Otto assumes that aspects of the “numinous” are still present in the religious life of contemporary religious persons.

6. Comparison

This chapter summarizes and compares the results of the previous examinations. It aims to provide responses to the research questions which guide this book: What does the notion of *surb* mean in Ancient Armenian texts from the fifth century CE? Are *surb*, *holy*, and the comparative notion of *holy* in the study of religion interrelated?

The first part concerns the respective significations and applications of *surb* in the context of the Ancient Armenian texts in the AAFTC and AASC and of *holy* in the context of the HSEC data (Chapter 6.1). The second part (Chapter 6.2) engages with the comparison of the semantic field of *surb* and the English semantic field of *holy* and tries to unravel their different strings of relationship. In a subsequent step, both notions will be examined against the intermediate *tertium comparationis* of *holy* as reflected in the study of religion (see Chapter 2).

Chapter 6.1 starts with the examination of the two individual semantic fields of *surb* and *holy* with the help of semantic graphs that include different layers of abstraction. There will be two versions of each of the semantic graphs. The first version is based on the “domain” annotation in the merged version of the synoptic tables (Chapter 5). The second variant considers the “semantic field” annotation from the same merged synoptic tables (see also Chapter 3).

The following semantic graphs and filters are utilized in Chapter 6.1.

- The first type of graphs are condensed graphs that display the general relevance of each “domain” and “semantic field” category for *holy* and *surb*. The significance of each “domain” or “semantic field” is based on the summarized “sum” values from the synoptic tables in Chapter 5. To receive the overall “sum” values for each “domain” or “semantic field,” the “sum” values from the synoptic tables in Chapter 5 are summarized and normalized¹ for each subcorpus. These normalized “sum” values are then added up to obtain an overall value for the different “domains” or “semantic fields” in the HSEC and AAFTC/AASC. The

1 Again, with the help of min-max normalization.

global values are once more normalized with the help of a min-max normalization to facilitate the comparison in Chapter 6.2. Finally, the significance of each “domain” and “semantic field” for the AAFT/AASC and HSEC data is visualized in the form of a graph.

- Secondly, extended graphs are created that display every “domain” or “semantic field” included in the HSEC and AAFT/AASC together with all the words that appear in each category.
- Depending on the corpus, different filters are then applied to the graphs mentioned above. One filter only displays those “domains” or “semantic fields” (and their corresponding words) that appear in all (or the majority) of the subcorpora. This permits looking for common patterns of the use of *holy* and *surb* throughout the diverse subcorpora of both the HSEC and the AAFT/AASC. Another filter only displays those “domains” or “semantic fields” (and their corresponding words) that appear in only one of the subcorpora. This enables achieving a better idea of use cases of *holy* or *surb* that only appear in specific contexts.
- The individual graphs will eventually be examined together to acquire an overview of the meaning and use of the notions of *holy* and *surb* in their respective contexts. In addition to the macroanalyses from Chapter 5 and the comparison in this chapter, this part will also consider the results of the microanalyses in Chapter 5.

Chapter 6.2 includes the comparison of the “domain” and “semantic field” graphs from Chapter 6.1 between the AAFT/AASC and HSEC. This comparison analyzes potential differences and overlappings between the Ancient Armenian use and meaning of *surb* and the contemporary notion of *holy*. This section also includes a comparison of *surb* and *holy* to the comparative notion of *holy* in the study of religion. This last step should first evaluate how far the empirical material is reflected in the more abstract discussions in the study of religion. Secondly, it should highlight potential blind spots in the comparative notion of *holy* that might need to be adjusted in order to better reflect the empirical material.

6.1 Examination of the Individual Semantic Fields

6.1.1 Overview of the “Domain” Section

In the following section, several semantic graphs are displayed and analyzed. All graphs are based on the condensed “sum” values of the synoptic tables from Chapter 5 and focus on the “domain” section.

The abbreviation “swai” in the column names of the tables stands for “number of subcorpora in which the word appears” whereas “noa” stands for “number of appearances.”

6.1.1.1 Domain Overview (No Words)

The first graphs and tables provide an overview of the importance of every single “domain” that appears in the HSEC and AAFTC/AASC data. The corresponding two figures (AAFTC/AASC and HSEC) show the “domains” based on the normalized “sum” values from the synoptic tables in Chapter 5 (see also the corresponding tables below). The first table for each corpus displays the “domain” values that correspond to the graphs, the second table displays in how many of the subcorpora each of the “domains” appears.

AAFTC/AASC

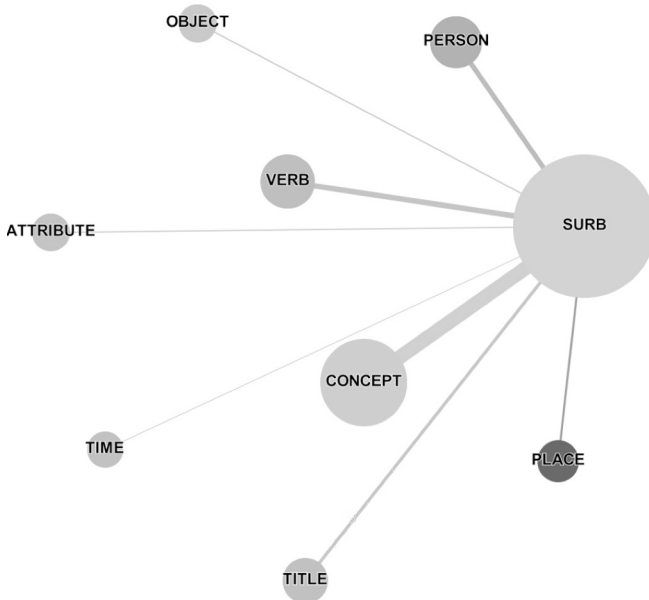
Table 80: Table of the “domain” section in the AAFTC and AASC (min-max normalized).

domain	value
concept	100
verb	36
person	32
title	19
place	12
object	5
attribute	4
time	1

Table 81: Table of the “domains” with the number of corpora in which the domains appear (AAFTC/AASC).

domain	noa
concept	4
verb	4
title	4
person	4
place	3
attribute	3
object	3
time	1

Figure 62: Visualization of the “domain” graph of the Armenian texts in the AAFTC and AASC; Fruchterman Reingold Graph rendered with (c) Gephi.



HSEC

Table 82: Overall table of the “domain” section of the HSEC (min-max normalized).

domain	value
concept	100
object	19
place	14
person	6
time	5
verb	1
title	1
attribute	1

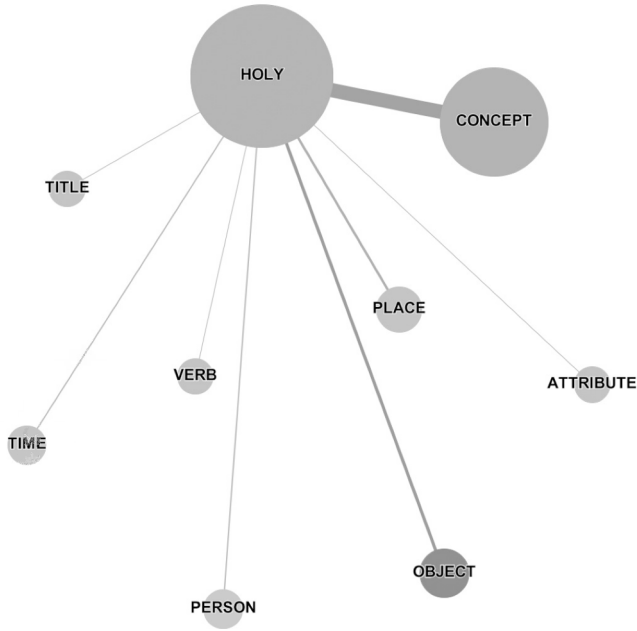
Table 83: Table of the “domains” with the number of corpora in which the domains appear (HSEC).

domain	noa
concept	5
verb	5
time	5
person	4
attribute	3
place	2
title	2
object	2

6.1.1.2 Domain Overview (All Words)

The following graphs give an overview of the “domain” distribution and, in addition, all the words assigned to the respective domains. Even though they provide an overview of the most important “domains” and their corresponding words, these graphs include too many words and become too confusing to provide any further insights. Therefore, the visualized graphs in the following figures 64 and 65 (AAFTC/AASC and HSEC), which display all words, should rather be understood as rough impressions of the general distribution of words. However, the corresponding tables only list the top 15 words (according to their “sum” value in the merged tables) and are thus more suitable for further examination.

Figure 63: Visualization of the domain graph of the texts in the HSEC; Fruchterman Reingold Graph rendered with (c) Gephi.



AAFTC/AASC

Table 84a: Table including the top 15 words plus their corresponding “domains” in the AAFTC and AASC data.

word	domain	value	swai
all (ամենայն)	concept	223	4
spirit (հոգի)	concept	220	3
god (աստուած)	concept	179	4
to make (առնել)	verb	155	4
name (անուն)	concept	131	3
church (եկեղեցի)	concept	112	3
bishop (եպիսկոպոս)	title	102	2
Nersēs (ներսէս)	person	100	1
Armenia(n) (հայ)	place	94	2
many (բազում)	concept	89	2

HSEC

Table 85: Table including the top 15 words plus their corresponding domains in the HSEC data.

word	domain	value	swai
religion	concept	216	3
sacred	concept	170	3
god	concept	146	3
man	person	119	2
world	place	111	2
see	concept	100	1
shit	object	100	1
time	time	81	2
one	concept	80	4
numinous	concept	72	1
feeling	concept	71	1
life	concept	70	3
cosmos	place	65	1
experience	concept	54	2

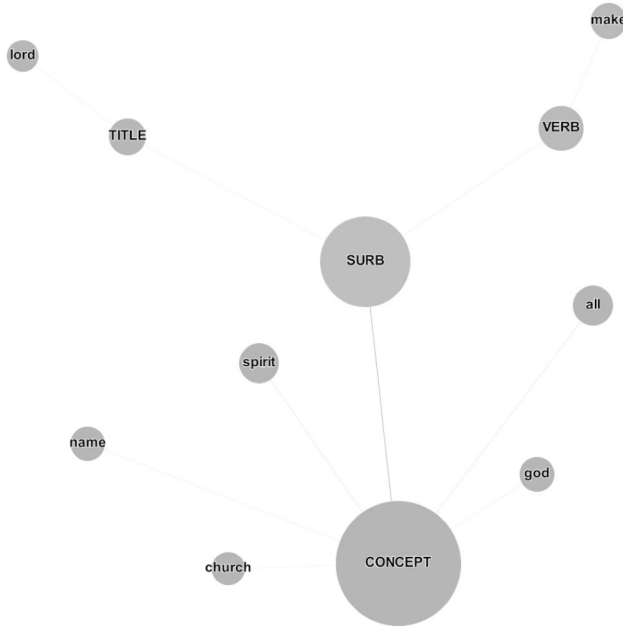
6.1.1.3 Domain Overview (Only Domains and Words that Appear in All Subcorpora)

In the case of the HSEC with five subcorpora, only the examination of words that appear in three or more subcorpora yielded meaningful results. Concerning the AAFTC/AASC, words that appear in at least three of four subcorpora have been selected. Since these graphs only include very few words, there are no additional tables in this section. Words that appear in the AAFTC/AASC in three or more subcorpora are displayed in figure 66. The visualization of words that appear in the HSEC in three or more subcorpora is shown in figure 67.

6.1.1.4 Overall Domain Overview (Only Domains and Words that Appear in Only One Subcorpus)

This part of the examination focuses on words that appear in only one subcorpus of the AAFTC/AASC and the HSEC (see the corresponding figures 68 and 69). The tables in this section only include the top 15 entries.

Figure 66: Visualization of the “domain” graph of the Armenian texts in the AAFTC/AASC, including all words and “domains” that appear in all subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi.



AAFTC/AASC

Table 86a: Table including the top 15 words plus their corresponding “domains” in the AAFTC/AASC data that appear in only one subcorpus.

word	domain	value	swai
Nersēs (ներսէս)	person	100	1
to intercede (բարեխօսել)	verb	71	1
Gregory (գրիգոր)	person	56	1
to call (կոչել)	verb	53	1
city (քաղաք)	place	50	1
to pass (անցել)	verb	50	1
pious (բարեպաշտ)	concept	46	1
region (կողմն)	place	43	1
Sahak (սահակ)	person	43	1
shrine (խորան)	place	43	1

Figure 67: Visualization of the “domain” graph of the English texts in the HSEC, including all words and “domains” that appear in all subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi.

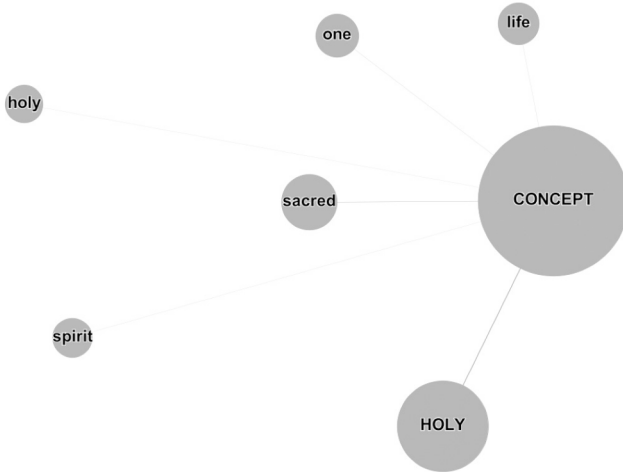
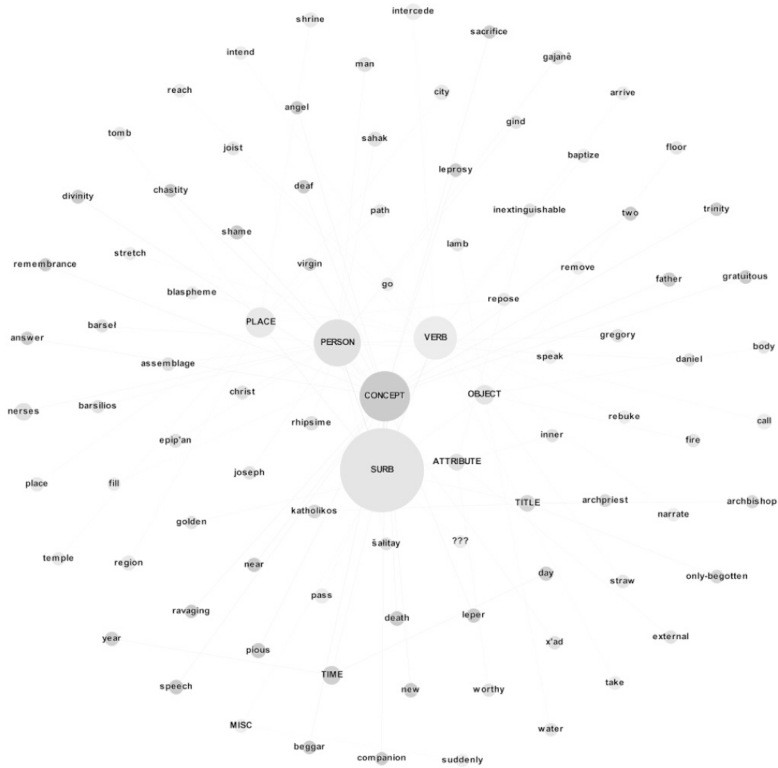


Table 86b: Table including the top 15 words plus their corresponding “domains” in the AAFTC/AASC data that appear in only one subcorpus.

word	domain	value	swai
man (այր)	person	40	1
death (վախճան)	concept	37	1
day (օր)	time	31	1
speech (բարբառ)	concept	31	1

Figure 68: Visualization of the “domain” graph of the Armenian texts in the AAFTC/AASC, including all words that appear in only one subcorpus; Fruchtermann Reingold Graph rendered with (c) Gephi.



HSEC

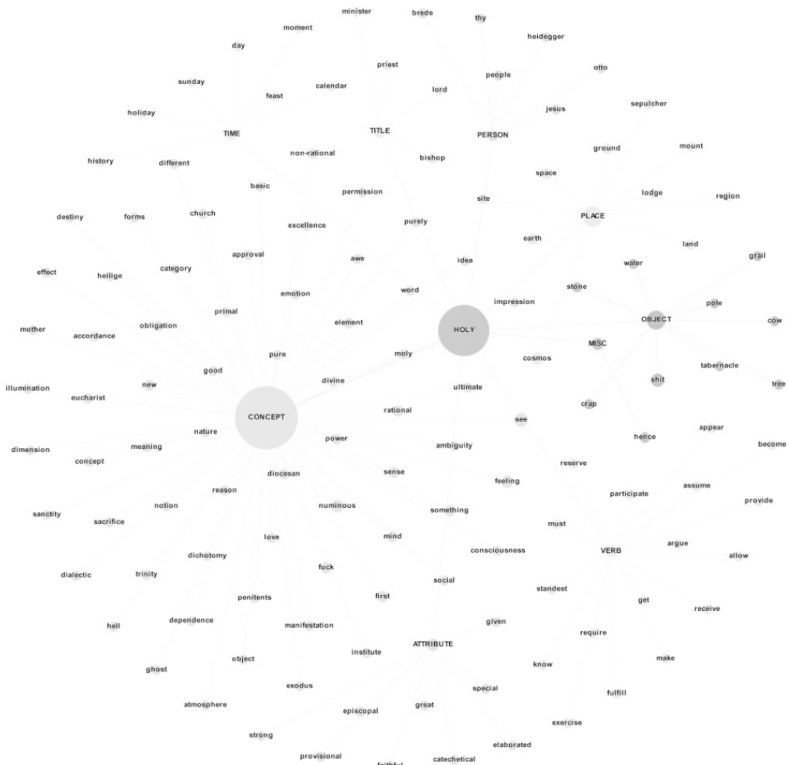
Table 87a: Table including the top 15 words plus their corresponding “domains” in the HSEC data that appear in only one subcorpus.

word	domain	value	swai
shit	object	100	1
see	concept	100	1
numinous	concept	72	1
feeling	concept	71	1
cosmos	place	65	1
something	concept	40	1
space	place	39	1

Table 87b: Table including the top 15 words plus their corresponding “domains” in the HSEC data that appear in only one subcorpus.

fuck	concept	36	1
first	concept	35	1
mind	concept	34	1
rational	concept	31	1
element	concept	30	1
new	concept	29	1
sense	concept	28	1

Figure 69: Visualization of the “domain” graph of the HSEC, including all words that appear in only one subcorpus; Fruchtermann Reingold Graph rendered with (c) Gephi.



6.1.2 Overview of the “Semantic Field” Section

The different versions of the “semantic field” graphs and tables follow the structure of the “domain” analysis in the previous section. Unlike in the study of the “domains,” the tables in this section usually only show the top 15 “semantic fields.” Only the abbreviations of the full “semantic field” titles are displayed in the visualized graphs. The full titles can be found in the corresponding sections of the documentation on the USAS website.² Only the first layer of the USAS tagging is applied that includes the most abstract layer from the complex semantic field hierarchy used in the USAS tagging system. Otherwise, the results would be too diverse to recognize any patterns.

6.1.2.1 Semantic Field Overview (No Words)

The first graphs and tables give an overview of the importance of every single “semantic field” in the HSEC and AAFTC/AASC data. Figure 70 (AAFTC/AASC) and figure 71 (HSEC) illustrate the importance of the individual “semantic fields” based on their merged “sum” values from the synoptic tables in Chapter 5 (see also the corresponding tables below). The first tables of each corpus display the “semantic field” values corresponding to the graphs. The following tables display in how many of the subcorpora each of the “semantic fields” appears.

2 Documentation USAS Semantic Tagger: <http://ucrel.lancs.ac.uk/usas/semtags.txt> (last accessed: 11/08/2021).

AAFTC/AASC

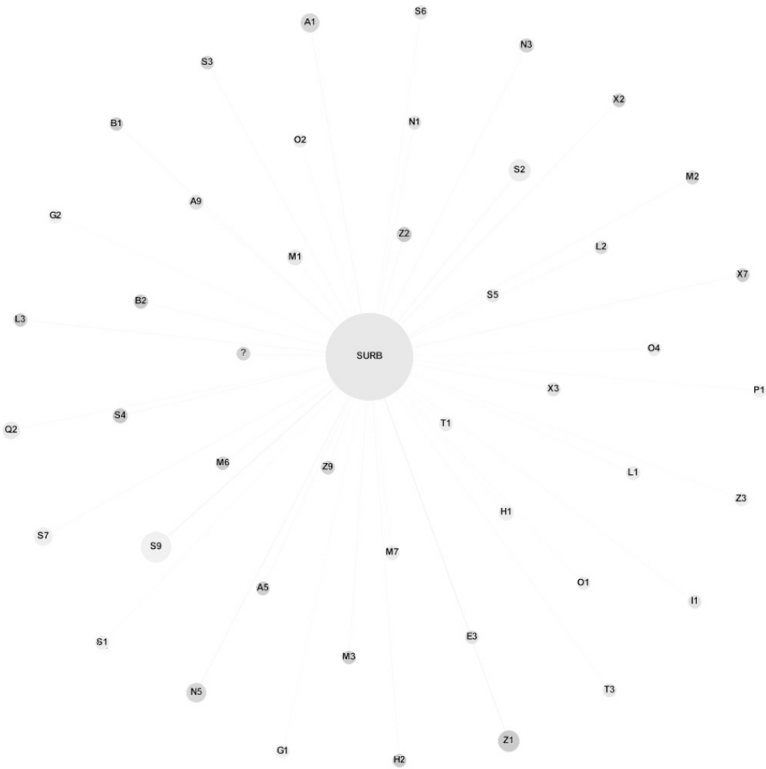
Table 88: Table with top 15 entries of the “semantic field” section in the AAFTC/AASC data (min-max normalized).

semantic_field_abbr	semantic_field_long	value
S9	Religion and the supernatural	100
S2	People	50
Z1	Personal names	47
N5	Quantities	37
A1	General and abstract terms	32
S7	Power relationship	31
Q2	Speech acts	24
M1	Moving, coming and going	14
H1	Architecture and kinds of houses and buildings	12
S4	Kin	9
Z2	Geographical names	9
M7	Places	7
A9	Getting and giving; possession	7
B2	Health and disease	6
S6	Obligation and necessity	6

Table 89: Table with top 15 entries of the “semantic field” section and the numbers of subcorpora they appear in.

semantic_field_abbrev	semantic_field_long	noa
S9	Religion and the supernatural	4
Q2	Speech acts	4
Z1	Personal names	4
N5	Quantities	4
A1	General and abstract terms	4
S7	Power relationship	4
S2	People	4
S4	Kin	3
M1	Moving, coming and going	3
H1	Architecture and kinds of houses and buildings	3
B1	Anatomy and physiology	3
N3	Measurement	3
S3	Relationship	2
X3	Sensory	2
Z2	Geographical names	2

Figure 70: Visualization of the “semantic field” graph (no words) of the Armenian texts in the AAFTC/AASC; Fruchterman Reingold Graph rendered with (c) Gephi.



HSEC

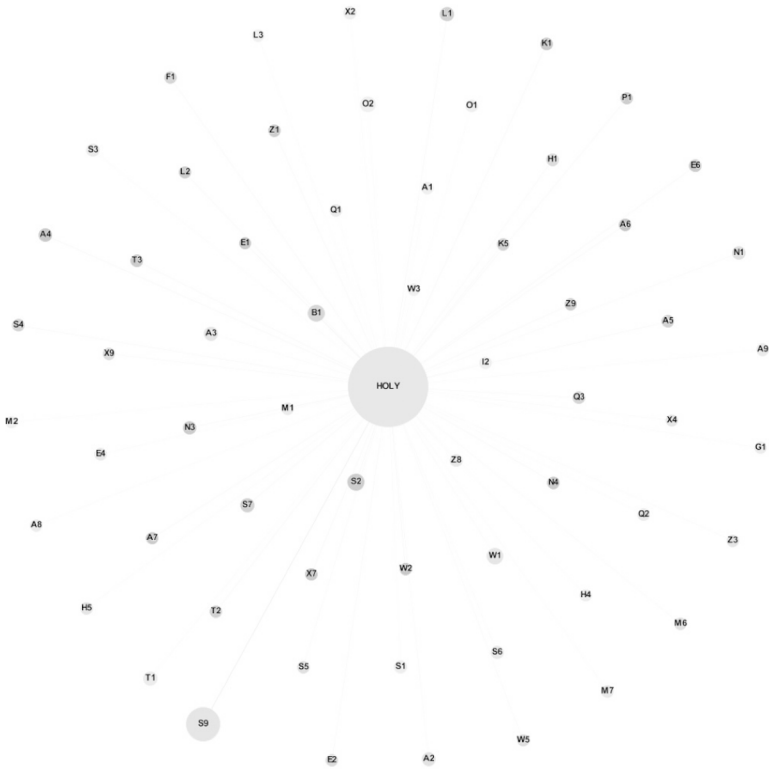
Table 90: Table with top 15 entries of the “semantic field” section of the HSEC (min-max normalized).

semantic_field_abbr	semantic_field_long	value
S9	Religion and the supernatural	100
S2	People	22
B1	Anatomy and physiology	22
W1	The universe	20
O2	General objects	16
T1	Time	9
A3	Being	9
S7	Power relationship	9
L1	Life and living things	9
A2	Affect	8
A4	Classification	6
N1	Numbers	6
X2	Mental actions and processes	5
S3	Relationship	5

Table 91: Table with top 15 entries of the “semantic field” section and the numbers of sub-corpora they appear in. Note that the Z9 semantic field “not defined” was dropped from the table.

semantic_field_abbrev	semantic_field_long	noa
S9	Religion and the supernatural	5
T1	Time	5
S2	People	5
A4	Classification	4
A2	Affect	4
S1	Social actions, states and processes	4
N1	Numbers	4
X2	Mental actions and processes	3
A3	Being	3
A5	Evaluation	3
L1	Life and living things	3
A1	General and abstract terms	3
S7	Power relationship	3
O2	General objects	3
Q3	Language, speech and grammar	2

Figure 71: Visualization of the “semantic field” graph (no words) of the English texts in the HSEC; Fruchterman Reingold Graph rendered with (c) Gephi.



6.1.2.2 Semantic Field Overview (All Words)

AAFTC/AASC

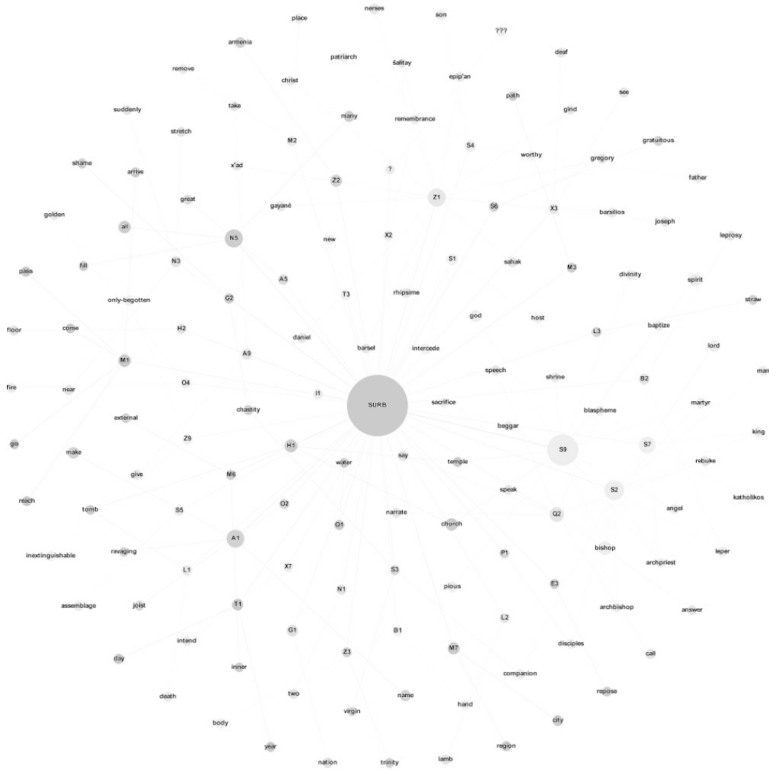
Table 92: Table including the top 15 words plus their corresponding “semantic fields” in the AAFTC/AASC data.

word	semantic_field	value	swai
all (ամենայն)	Quantities	223	4
spirit (հոգի)	Religion and the supernatural	220	3
god (աստուած)	Religion and the supernatural	179	4
to make (առնել)	General and abstract terms	155	4
name (անուն)	General and abstract terms	131	3
church (եկեղեցի)	Religion and the supernatural (a.o.)	112	3
bishop (եպիսկոպոս)	Religion and the supernatural (a.o.)	102	2
Nersēs (ներսէս)	People	100	1
Armenia(n) (հայ)	Geographical names	94	2
many (բազում)	Quantities	89	2
Lord (տէր)	Power relationship	86	3
to intercede (բարեխօսել)	Social actions, states and processes	71	1
nation_(ազգ)	Government, Politics and elections	66	2
to say (ասել)	Speech acts	59	2

The following graphs give an overview of the “semantic field” distribution in addition to all the words assigned to the respective semantic fields. Even though they provide an overview of the most important “semantic fields” and their corresponding words, these graphs include too many words and become too confusing to provide any further insights. Therefore, the visualized graphs in figure 72 (AAFTC/AASC) and figure 73 (HSEC), which display all words, should rather be understood as rough impressions of the general distribution of words. However, the corresponding tables only list the top 15 words (according to their “sum” value in the merged tables) and are thus more suitable for further examination.

Some of the words might be related to several “semantic fields.” Even though multiple connections are shown in the visualized graph, the tables only include the first “semantic field” attributed to each word. Furthermore, some of the tagging had to be done manually. For instance, the noun “name” was initially tagged as belonging to the “Speech acts” semantic field. Since “name” is mostly used as a noun in both the Armenian and English texts, it was manually added to the “Abstract terms” category.

Figure 72: Visualization of the “semantic field” graph of the Armenian texts in the AAFTC/AASC including all words and “semantic fields;” Fruchtermann Reingold Graph rendered with (c) Gephi.



HSEC

Table 93a: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data.

word	semantic_field	value	swai
religion	Religion and the supernatural	216	3
sacred	Religion and the supernatural	170	3
god	Religion and the supernatural	146	3
man	People	119	2
world	The universe	111	2
see	General objects	100	1
shit	Anatomy and physiology	100	1

Table 93b: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data.

word	semantic_field	value	swai
time	Time	81	2
one	Numbers	80	4
numinous	Religion and the supernatural	72	1
feeling	Mental actions and processes	71	1
life	Life and living things	70	3
cosmos	The universe	65	1
experience	Affect	54	2

6.1.2.3 Semantic Field Overview (Only Semantic Fields and Words that Appear in All Subcorpora)

In the case of the HSEC with five subcorpora, only the examination of words that appear in three or more subcorpora yielded meaningful results. Concerning the AAFTC/AASC, words that appear in at least three of four subcorpora have been selected. Since these graphs only include very few words, there are no additional tables in this section. The words that appear in the AAFTC/AASC in three or more texts are displayed in figure 74. The visualization of words that appear in the HSEC in three or more texts is shown in figure 75.

Figure 74: Visualization of the “semantic field” graph of the Armenian texts in the AAFTC/AASC including all words and “semantic fields” that appear in all subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi.

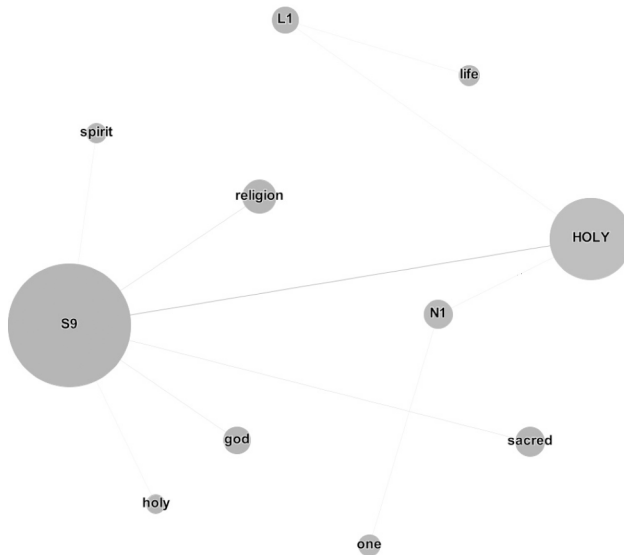


AAFTC/AASC

Table 94: Table including the top 15 words plus their corresponding “semantic fields” in the AAFTC/AASC data.

word	semantic_field	value	swai
Nersēs (նէրսէս)	Personal names	100	1
to intercede (բարեխօսել)	Social actions, states and processes	71	1
Gregory (գրիգոր)	Personal names	56	1
to call (կոչել)	Speech acts	53	1
city (քաղաք)	Places	50	1
to pass (անցել)	Moving, coming and going	50	1
pious (բարեպաշտ)	Religion and the supernatural	46	1
region (կողմն)	Places	43	1
Sahak (սահակ)	Personal names	43	1
shrine (խորան)	Places	43	1
man (այր)	People	40	1
death (վախճան)	Life and living things	37	1
day (օր)	Time	31	1
speech (բարբառ)	Speech acts	31	1

Figure 75: Visualization of the “semantic field” graph of the English texts in the HSEC including all words and “semantic fields” that appear in all three subcorpora; Fruchtermann Reingold Graph rendered with (c) Gephi.



HSEC

Table 95a: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data.

word	semantic_field	value	swai
shit	Anatomy and physiology	100	1
see	Speech acts	100	1
numinous	Religion and the supernatural	72	1
feeling	Mental actions and processes	71	1
cosmos	The universe	65	1
something	Pronouns	40	1
space	Measurement	39	1

Figure 76: Visualization of the “semantic field” graph of the Armenian texts in the AAFTC/AASC including all words and “semantic fields” that appear in only one subcorpus; Fruchtermann Reingold Graph rendered with (c) Gephi.

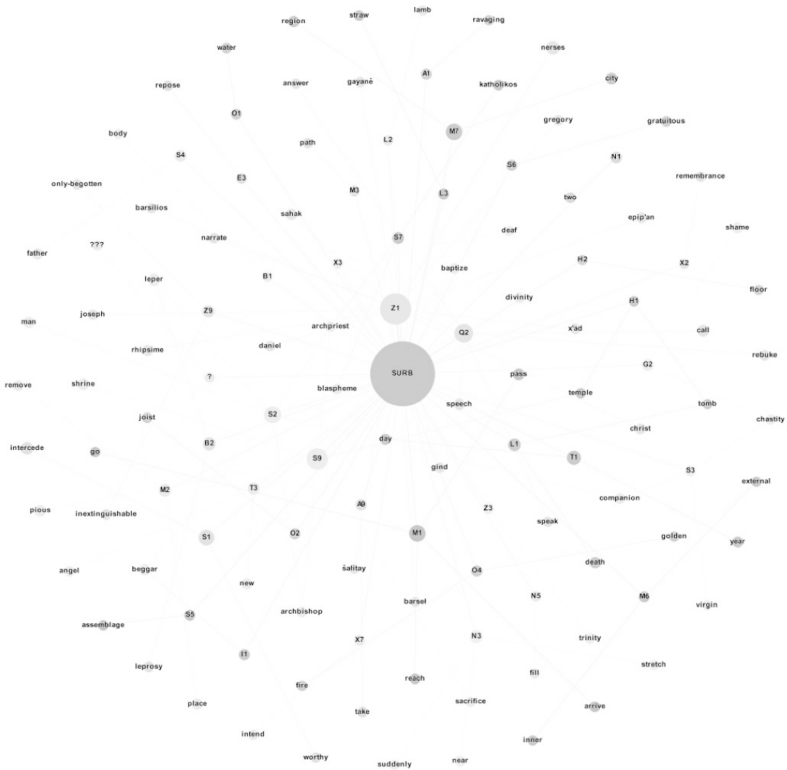


Table 95b: Table including the top 15 words plus their corresponding “semantic fields” in the HSEC data.

word	semantic_field	value	swai
fuck	Anatomy and physiology	36	1
first	Linear order	35	1
mind	Worry, concern, confident	34	1
rational	Social actions, states and processes	31	1
element	Classification	30	1
new	Time	29	1
sense	Classification	28	1

- The top “domain” in the context of *surb* is the “concept” domain. It is followed by the “verb,” “person,” and “title” domains. All four domains appear in every subcorpus of the AAFTC/AASC, meaning that they are commonly found in the context of *surb*. However, the “concept” domain is very unspecific and includes many different words, such as “god,” “all,” and “nation.” Furthermore, it has very blurry boundaries with other domains, such as “attributes.” The “semantic field” annotation offers a more detailed insight into the subgroups within the very broad “concept” domain. According to the “semantic field” annotation, the most important semantic field in the AAFTC/AASC is “religion and the supernatural” (S9), containing words such as “god,” “spirit,” and “church.” Other semantic fields are related to “personal names” (Z1) and “people” (S2), thereby illustrating the diversity of the “person” domain in the AAFTC/AASC that contains both individual persons and groups of persons. Taken together, the “people” (S2) and “personal names” (Z1) semantic fields are almost as important as the “religion” (S9) semantic field. Other important semantic fields are “quantities” (N5), “abstract terms” (A1), and “power relationship” (S7). All these semantic fields appear in all four subcorpora of the AAFTC/AASC and thus represent widespread associative contexts of *surb*. The “quantities” (N5) semantic field includes words such as “many” and “all,” which are of lesser relevance for the meaning and use of *surb*, among others because they regularly appear among the lockwords in the analyses of keywords (see Chapter 5). The “power relationship” (S7) semantic field includes many titles from the corresponding “titles” domain, such as “bishop” or “archpriest.” In contrast to the “domain” section, the “semantic field” annotation system permits multiple attributions. It thus highlights the overlappings between the “persons” semantic field and the semantic field of “power relationship” in the case of titles, such as “bishop” or “king.” In sum, the “domain” section reveals the importance of concepts, persons, and titles in the AAFTC/AASC data. The “semantic field” section adds more details to these observations by demonstrating that many of the words in these domains are related to religion, quantities, abstract terms, and power relationships. Furthermore, the second annotation theme reveals that the focus on persons in the AAFTC/AASC lies both on individual persons and on abstract groups of persons.
- The previous observations are also mirrored in the words that appear in all subcorpora. Among these words are religious concepts, such as “god,” “church,” or “spirit,” quantities such as “all,” and abstract terms such as “name”³ or titles such as “lord.” These words mostly represent very general religious con-

3 Note that the word “name” was initially wrongly classified by the USAS semantic tagger. It has manually been assigned to the “Abstract terms” (A1) semantic field.

cepts that were heavily influenced by Christian traditions; most prominently the “spirit” and the “church.”

- The overviews of the inflected forms and derivatives of *surb* in Chapter 5 reveal that the terminology was mostly used as an adjective or nominalized adjective. The form of *surb*-ness was also common but less frequently applied. The use of *surb* as a verb is found rather seldom. Yet, it often hints at the important application and understanding of *surb* in the sense of purity that will be discussed below.
- The table and graph with words that only appear in one of the subcorpora include many names that are particularly relevant in one of the subcorpora, among them Sahak (Koriwn), Nersēs (“Epic Histories”), and Gregory (Agat’angelos). Furthermore, they underline the importance of specific places in the subcorpora, particularly in the AAFTC/AASC Koriwn and the AAFTC/AASC Agat’angelos. The “place” domain and the corresponding “semantic fields” such as “Geographical names” (Z2) or “Architecture” (H1) also appear among the top entries of the most frequent “semantic fields” and “domains” in Chapter 6.1. The “place” domain comprises diverse words some of which are referring to concrete local places such as “shrine” or “city” and others to more abstract terms such as “Armenia” or “region.” Related to the two latter terms are other words from the macroanalysis, namely the “nation” and the “king.” The ensemble of “nation,” “Armenia,” “region,” and “king” demonstrates that the terminology around *surb* is not only closely related to religious concepts or places but to political ideas of territory, power, and identity as well, at least in the AAFTC/AASC Koriwn and the AAFTC/AASC Agat’angelos.
- The “verb” domain can also be further subdivided with the help of the “semantic field” annotation. The verbs that occur in the macroanalysis emphasize the importance of persons in the AAFTC/AASC previously noted. One group contains verbs of social actions such as “to make” or “to give” (S1). Another group comprises words related to the semantic field of “Speech acts” (Q2), such as “to say” or “to narrate.” Another interesting verb that appears in only one of the texts, although with a high “sum” value, is the verb “to intercede” that is usually found in the context of saints and martyrs. Other verbs such as “to baptize” are directly related to the semantic fields of religion and purity. Lastly, there is an important subgroup of verbs related to movements, including verbs such as “to go,” “to pass,” or “to arrive.” Together with the “place” and “person” domains and the corresponding “semantic fields,” this group indicates the central role of traveling missionaries in the Armenian accounts and the context of *surb* as well.

6.1.3.2 Microanalysis

The microanalysis of the AAFTC/AASC subcorpora in Chapter 5 underlines central aspects observed in the macroanalysis, among them the significant role of persons, religion, and abstract concepts in the context of *surb*. Yet, it also sheds further light on the understanding of the use and meaning of *surb*.

- One important use case of *surb* which mostly appeared in the microanalysis is the embedding of *surb* in attribute chains that include other religious and non-religious attributes such as “blessed,” “worthy,” “virtuous,” “chaste,” or “great.” In these attribute chains, which are mostly used to characterize persons, the term *surb* does not necessarily hold a more prominent position compared to the other attributes. Instead, *surb* is only used as one among many other characteristics. The importance of attributes is also mirrored in the macro data, where words such as “great” or “worthy” also appear. Yet, due to the very blurry boundaries between the “concept” and “attribute” domains as well as the different attributions in the “semantic field” section, their joint use is less evident from a macro perspective. It becomes particularly evident when considering specific sentences.
- Closely related to the use of *surb* in attribute chains is the occurrence of terms that represent the counterpart of *surb*. In this context, the word *pilc* (պիլծ) was of particular importance as it was commonly applied to represent the opposite of *surb*, for instance, in the Gospels and the “History of the Armenians” by Agat’ angelos. Similar to the application of *surb* in positive attribute chains, the terminology related to *pilc* (պիլծ) frequently appears in negative attribute chains to characterize the enemies of the *surb* persons in the form of attributes such as *pilc* (պիլծ) and “impious” (անսարէն).⁴ Yet, the term *pilc* did not appear in the macroanalyses.
- The relationship between *surb* and *pilc* was particularly evident in the Gospels. In this text, *surb* was usually applied to translate Greek terms such as *kat’ arós* and *hágios* whereas *pilc* was used to translate words such as *akat’ arós*. This inter-relationship between the Armenian and the more differentiated Greek terminologies hints at a central aspect of the application and understanding of *surb* in the Armenian data that was barely perceivable in the macroanalysis, namely the relationship between the notion of *surb* and the semantic field of purity. To some degree, this connection is observable in all of the Armenian corpora, although it was most evident in the comparison of the Greek and Armenian versions of the Gospels. The notion of *surb* covered a wide range of different forms of purity, ranging from mere physical purity in the sense of “clean clothes” over

4 See §143 in the “History of the Armenians.”

sexual and ethical purity to religious purity and healing. The overlappings between *surb* and healing as well as sickness are also present in the macro data in the form of terms such as “leprosy” or “leper.”

6.1.3.3 Summary

According to the macro- and microanalyses, the notion of *surb* frequently occurred in the proximity of abstract and religious concepts. Among these concepts were general Christian ideas or institutions, such as the “spirit,” the “church,” or “titles.” The close relationship between *surb* and “spirit” and its high frequency in the AAFTC/AASC particularly underlines the importance of trans-regional concepts in early Christianity on both Armenian literature as well as the use and meaning of notions such as *surb*. A case in point is the *surb hogi* who was not an Armenian invention but based on the concept of the “holy spirit” taken from other Christian texts and from the Bible in particular. Thus, the application of *surb* is often used as an Armenian instantiation and translation of already existing non-Armenian concepts (such as *hágios*), which consequently had a strong influence on the overall understanding of *surb*.

Furthermore, *surb* was commonly used as one among many other attributes to describe persons and groups of persons. Unlike the other attributes used in these contexts, the attribute *surb* was exclusively granted to groups of religious persons. The term *surb* was frequently related to places as well. Yet, its attribution to objects was only occasionally found.

An essential aspect of the application and understanding of *surb* was exclusively apparent in the microanalyses, namely its connection to a wide range of phenomena related to purity. In this context, the counter notion of *surb (pilc)* regularly appeared as well. That this relationship between purity and *surb* was mainly visible in the microanalyses and not in the macroanalysis strengthens the assumption that this layer of meaning was already overshadowed by a more general understanding of *surb* as a religious notion and attribute. The latter more unspecific religious use case of *surb* in the Armenian texts indicates a general closeness of persons or institutions to god, as it is most prominently observable in the “holy spirit” as part of the Trinity. Furthermore, the increasing abstraction of *surb* can rightfully be regarded as a first step towards so-called sedimentation processes. In this context, sedimentation means that *surb* no longer indicates a specific characteristic (such as “pure” or “close to god”); rather, the terminology is used in fixed expressions (such as “holy see,” “holy spirit,” or “holy katholikos”) to evoke a very general associative context, namely “religion.”

6.1.4 Overview of the Semantic Field of *Holy*

This section examines the semantic graphs in combination with the results of the microanalyses from Chapter 5 to evaluate the meaning of *holy* in contemporary English.

6.1.4.1 Macroanalysis

The macroanalyses of the HSEC and *English Web corpus 2015 (enTenTen15)* in Chapter 5 and the semantic graphs in Chapter 6 yielded the following results of the application and understanding of *holy*.

- The most important domain in the HSEC is the “concept” domain. It is followed by the “object,” “place,” and “person” domains. Only the “concept” domain appears in all five subcorpora, whereas the “person” domain is found in four and the “object” and “place” domains in two of the five subcorpora. This observation indicates that the use of *holy* in the context of “concepts” and “persons” is relatively widespread. In contrast, the appearance of *holy* in the context of “places” and “objects” represents special use cases. These observations are very much in line with the previous analysis of the semantic field of *surb*. Yet, a closer look at the more differentiated “semantic field” annotation reveals that the “person” domain mainly consists of “People” (S2) and only to a minimal extent of “individual persons” (Z1). The few individual persons found in the HSEC macroanalysis stem primarily from the HSEC Encyclopaedia Britannica subcorpus and represent important scholars such as Rudolf Otto or Martin Heidegger. The words in the “People” (S2) category are mostly very abstract and include terms such as “man,” “people,” and “human.” Regarding the “concept” domain, the “semantic field” annotation indicates that there are mainly concepts related to “Religion and the supernatural” (S9), “Being” (A3), “Power relationships” (S7), “Affects” (A2), and “Mental actions” (X2). Yet, among these different semantic fields, only “Religion and the supernatural” (S9) appears in all five subcorpora. Conversely, the other categories occur in four (“Affects” [A2]) or fewer subcorpora (“Being” [A3] and “Mental actions” [X2]). Thus, besides a strong focus on religion, the lexical fields around *holy* also seem to have a significant connection to cognition and feelings.
- Even though there is a huge gap between the “objects” and “place” domains on the one hand and the “concept” domain on the other hand, there are several “semantic fields” in the tables and graphs that indicate the importance of “objects” and “places” for the understanding of *holy* in specific use cases. Regarding the “object” domain, the semantic fields of “Anatomy and physiology” (B1) and “General objects” (O2) reflect its inner diversity. The “Anatomy and physiology” (B1)

semantic field comprises words that belong to both the “object” domain (“shit”) and the “concept” domain (“consciousness,” “fuck”). The fact that most of these words only appear in one single corpus, namely the HSEC Twitter Data, illustrates their importance in specific contexts. The same observation holds for the “General objects” (O2) semantic field, including words such as “see,” “grail,” and “pole.” Of these words, the “see” is particularly relevant in the HSEC Canon Law data, the “grail” is commonly used in the HSEC Twitter Data, and the “pole” is part of the HSEC Eliade. Therefore, the objects related to *surb* are usually found in specific contexts but are less representative of a more widespread meaning. In sum, the examples from the “object” domain and the corresponding “semantic fields” demonstrate that the use of *holy* in the context of objects is rather an exception and that the objects vary greatly. However, when these exceptions occur, they can be relevant for the overall use of *holy* in these specific contexts and appear with a very high frequency. An example is the expression “holy shit” in everyday language or the *terminus technicus* “holy see” in the Canon Law. The application of *holy* in both cases is restricted to specific circumstances but appears very frequently therein.

- The “place” domain is represented by several “semantic fields,” most prominently “The universe” (W1) (with words such as “world” and “cosmos”), “Measurement” (N3) (“space”), and “Places” (M7). Most of the words in the “place” domain only appear in one or two subcorpora (HSEC Eliade and HSEC Encyclopaedia Britannica). The majority of the words in the “place” domain and the corresponding “semantic fields” either represent very abstract and broad spatial concepts, such as “cosmos,” “world” and “earth,” or natural spaces, such as “land,” “ground” or “mount.” Almost none of the words in this category have an immediate religious meaning; however, some of them, such as “sepulcher,” are religioid.
- Regarding the words, “domains,” and “semantic fields” that appear in (almost) all of the subcorpora, the “Religion and the supernatural” (S9) semantic field with general religious words such as “god” or “spirit” is very present. The appearance of “one” is less relevant since this word is frequently used in several English expressions and cannot be restricted to one dominant use case with particular relevance for the application and understanding of *holy*. Interestingly, the word “life” also appears in a majority of the subcorpora and is thus also counted among the top words that are typical for widespread use of *holy*.
- In the table and semantic graph of the words and categories that appear in only one of the subcorpora, many words from the already mentioned “place” domain appear as well as words related to emotions and feelings. These words, such as “feeling,” “awe,” or “emotion,” belong to different “semantic fields,” such as “Mental processes” (M2) and “Emotional actions” (E1). Their particular relevance for the overall corpus mostly derives from HSEC Otto. Generally, many

words and categories that have already been discussed are found in the semantic graph and table of words that appear in only one subcorpus, thereby indicating the very diverse application of *holy* that differs significantly between the subcorpora. “Time” is another “domain” and “semantic field” that appears in the table and semantic graph with words that appear in one subcorpus only. Many words related to “time” such as “calendar” or “day” occur in only one subcorpus. The term “time” only appears in two subcorpora (HSEC Eliade and HSEC Encyclopaedia Britannica). Its application in the HSEC Eliade resonates with the discussion of the words in the “place” and “person” domains, indicating that *holy* in the HSEC is often applied in the context of comprehensive and abstract words.

6.1.4.2 Microanalysis

It should be kept in mind that the terminology of *holy* differed significantly between the subcorpora in the HSEC. For instance, the word “sacred” was frequently used in the HSEC Eliade, whereas the word “holy” was dominant in the Twitter data. Furthermore, other words such as “numinous” in the HSEC Rudolf Otto could be applied in place of *holy* as well. It has been demonstrated throughout this book that they are all related to a joint semantic field of *holy*. The microanalyses of the HSEC mostly confirm the results of the examination of the macro data.

First and foremost, the microanalyses of the different subcorpora reflect the great variety of use cases of *holy*. It also illustrates that their application greatly varied between the different subcorpora. Yet, the microanalyses also highlight several aspects which have not been mentioned so far.

- The word “holy” was often part of fixed religious and non-religious expressions, such as “Holy See,” “Holy Communion,” “Holy Spirit,” or “holy shit.” Many of these fixed expressions commonly start with a capital letter. The frequent use in fixed expressions can be taken as an indication of sedimentation processes. Yet, in some cases, the term “holy” is still used as an independent adjective, thereby demonstrating that the static use in fixed expressions and the application as a dynamic attribute can coexist.
- Even though words such as “profane” did also appear in the macroanalysis, the presence and importance of the *sacred* and *profane* distinction was especially perceivable in the microanalysis.
- The *holy* regularly appeared as a self-acting, powerful, and often mighty entity or even agent. This aspect was visible in the prayers in the HSEC Twitter Data, in the idea of an autonomous “sacred” entity that could reveal itself in the history of religion in HSEC Eliade (“hierophany”), and lastly, in the HSEC Otto in the form of the numinous. The idea of the *holy* as an acting entity might be a

result of the overarching and continuing influence of the Christian concept of the “holy spirit.”

- The everyday use of “holy” in the context of words such as “fuck,” “shit,” or “crap” in the HSEC Twitter Data revealed the perduring relationship between *holy* and purity. Yet, this connection between *holy* and purity was almost exclusively present in the non-religious everyday use of “holy” and nearly non-existent in the other subcorpora of the HSEC.

Summary

According to the macro- and microanalyses, the notion of *holy* was commonly applied in the context of abstract words. These abstract words could either be related to cognitive concepts (“feelings,” “emotions”), persons (“human,” “man”), places and other spatial concepts (“cosmos,” “world”), or time (“time”). Concrete names or objects seldom appeared in the HSEC analyses, and if so, they were mostly used as examples of more abstract ideas.

The term “holy” was frequently used in fixed religious and non-religious expressions, thereby showing that it had often undergone a process of sedimentation. Neither the word “sacred” nor the word “holy” was regularly applied to function as a distinct adjective in order to attribute specific characteristics to an object, person, or concept.

Instead, the notion of *holy* often occurred as a self-sufficient actor or power that could do certain things; for instance, directly changing the destiny of a person in prayer or revealing itself in the history of man or in particular “numinous” situations.

The overall importance of religion was visible throughout all five subcorpora, even though some of the words from the semantic field annotation “religion” (S9) such as “grail” are mostly used in a non-religious sense. Overall, the majority of use cases of *holy* in the subcorpora of the HSEC can be described as representing an unspecific, autonomous, efficacious, and potentially harmful power that is related to religion in its broadest understanding as a sphere of complex transcendence (Luckmann). An interesting subcase of the application and understanding of *holy* was observable in the HSEC Twitter Data. Here, the term “holy” was frequently used together with words from a semantic field of impurity to evoke an ambiguous state. This relationship between *holy* and purity was almost absent in the other English sources.

6.2 The Comparison

The comparative Chapter 6.2 is divided into two sections. The first section examines the relationship between the semantic field of *surb* from the AAFTC/AASC analyses

with the semantic field of *holy* from the HSEC analyses in the previous section (Chapter 6.1). The second section investigates potential overlappings between these two semantic fields and the *tertium comparationis* of *holy* in the study of religion as outlined in Chapter 2.

6.2.1 Comparison I: *Holy* and *Surb*

The comparison between the semantic field of *surb* in the AAFTC/AASC and the semantic field of *holy* in the HSEC is based on the synoptic tables and microanalyses from Chapter 5 and on the results of Chapter 6.1. The focus in the first part of this subchapter lies on the macroanalyses, although the results of the microanalyses are also briefly considered. The second part of Chapter 6.2 considers both the macro- and the microanalyses.

The graphs and tables below provide the following information about potential overlappings between the semantic fields of *surb* and *holy*.

1. The first graph (see the corresponding figure 78) and table display the intersection of domains from the “domain” annotation scheme between the HSEC and the AAFTC/AASC. The size of the nodes indicates the “weighted degree.” The “weighted degree” is calculated based on the “weight” values of the edges of each node, which are the normalized “sum” values of the “domains” (see the graphs and corresponding tables), and the number of edges connected to the node. The thickness of the edges reflects the normalized “sum” value of the connected “domain” in either the HSEC or AAFTC/AASC. Due to these calculations, a large node does not necessarily indicate that a “domain” or “semantic field” is equally important in both corpora. It might also suggest that a specific “domain” is only relevant in one of the two corpora. Thus, the interpretation of the semantic graphs needs to consider both the size of the node and the thickness of both connected edges. The underlying “sum” values are displayed in the corresponding tables of each semantic graph.
2. The second semantic graph (see the corresponding figure 79) and table are almost equal to the “domain” graph described above. However, instead of “domains,” they represent the intersections of “semantic fields” between the HSEC and the AAFTC/AASC.
3. The third and last graph (see the corresponding figure 80) and table provide an overview of the intersection of words that appear in the synoptic tables of the HSEC and the AAFTC/AASC. The words from the AAFTC/AASC are written in their original Armenian alphabet. Yet, since they are connected to their corresponding terms in the HSEC, the English word in the graph also serves as a translation of the Armenian term. The corresponding “semantic fields” to

which these words belong are also displayed in the graph. The color of the nodes indicates different “domains” from the “domain” annotation scheme.

6.2.1.1 Intersection between AAFTC/AASC and HSEC (Domains)

Figure 78: Intersection of “domains” between the HSEC and the AAFTC/AASC. (c) Gephi.

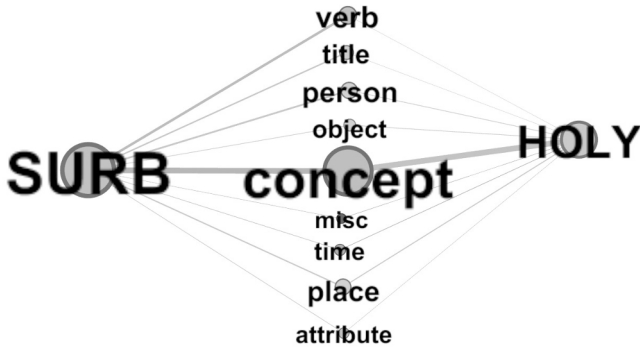


Table 96: Comparison of “domain” values between the HSEC and AAFTC/AASC according to their normalized “sum” values.

domain	sum_en	sum_arm
attribute	1.147343	4.078858
concept	100.0	100.0
misc	0.905797	0
object	6.340580	5.982325
person	6.884058	27.872196
place	14.070048	18.354861
time	5.253623	3.535010
title	0.724638	21.549966
verb	0	41.740313

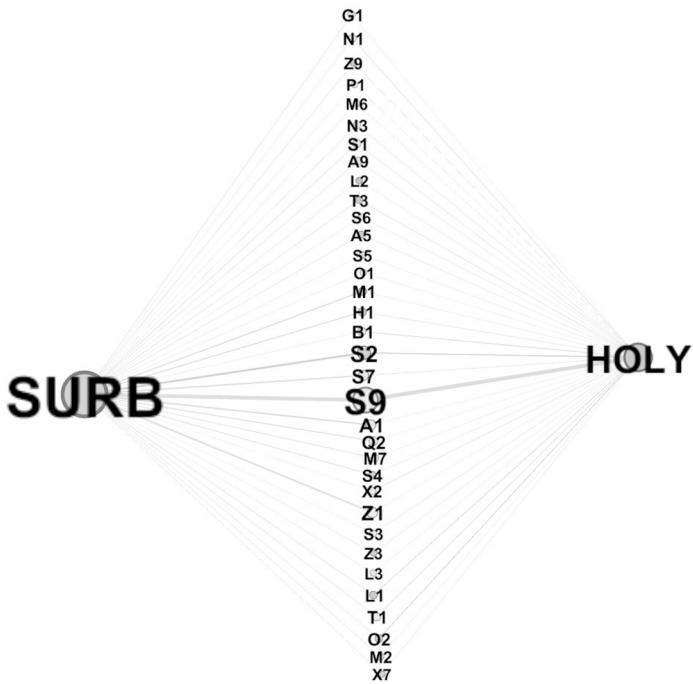
6.2.1.2 Intersection between AAFTC/AASC and HSEC (Semantic Field)

Table 97: Comparison of “semantic field” values between the HSEC and AAFTC/AASC according to their normalized “sum” values.

sf_short	sf_long	sum_en	sum_arm
A1	General and abstract terms	0.8	36.9
A5	Evaluation	1.8	3.2
A9	Getting and giving; possession	1.5	5.8
B1	Anatomy and physiology	16.6	3.0
G1	Government, Politics and elections	0.1	8.0
H1	Architecture and kinds of houses and buildings	2.3	14.7
L1	Life and living things	8.3	4.6
L2	Living creatures generally	0.1	0.7
L3	Plants	0.1	1.6
M1	Moving, coming and going	0.1	14.7
M2	Putting, taking, pulling, pushing, transporting &c.	0.4	4.9
M6	Location and direction	0.1	1.4
M7	Places	1.8	11.4
N1	Numbers	9.5	2.5
N3	Measurement	4.6	4.9
O1	Substances and materials generally	0.1	1.6
O2	Substances and materials generally: Solid	12.2	1.6
P1	Education in general	0.8	0.3
Q2	Speech acts	0.1	22.5
S1	Social actions, states and processes	5.3	10.0
S2	People	22.6	48.4
S3	Relationship	4.3	0.9
S4	Kin	0.1	9.3
S5	Groups and affiliation	0.8	1.7
S6	Obligation and necessity	2.0	1.6
S7	Power relationship	5.9	32.5
S9	Religion and the supernatural	100	100
T1	Time	11.2	7.2
T3	Time: Old, new and young; age	3.4	3.3

X2	Mental actions and processes	8.6	0.1
X7	Wanting; planning; choosing	0.1	0.9
Z1	Personal names	0.5	42.8
Z3	Other proper names	0.1	1.4
Z9	not defined	0.1	3.3

Figure 79: Intersection of “semantic fields” between the HSEC and the AAFTC/AASC. (c) Gephi.



6.2.1.3 Intersection between AAFTC/AASC and HSEC (Words)

Table 98: Intersection of words and their corresponding “domains” and “semantic fields” between the HSEC and the AAFTC/AASC.

engl	arm	sum_en	sum_arm	domain	sf
bishop	Եպիսկոպոս	6.9	45.739910	title	bishop_S9/S7.1+/S2mf
church	Եկեղեցի	8.796296	50.224215	concept	church_S9/H1c
day	օր	0.925926	13.901345	time	day_T1.3
god	աստուած	67.592593	80.269058	concept	god_S9
great	մեծ	0.1	12.107623	attribute	great_A5.1+
lord	տէր	0.925926	38.565022	title	Lord_S7.1+/S2.2m
make	առնել	0.1	69.506726	verb	make_A1.1.1
man	այր	55.092593	17.937220	person	man_S2.2m
man	մարդ	55.092593	13.004484	person	man_S2.2m
place	դիր	1.388889	0.448430	place	place_M2
place	տեղի	1.388889	11.210762	place	place_M2
region	կողմն	0.1	19.282511	place	region_M7
sacrifice	պատարագ	0.1	3.139013	concept	sacrifice_S9
spirit	հոգի	25.000000	98.654709	concept	spirit_S9
trinity	Երրորդութիւն	0.1	5.381166	concept	trinity_Z3c
water	ջուր	0.1	6.278027	object	water_O1.2

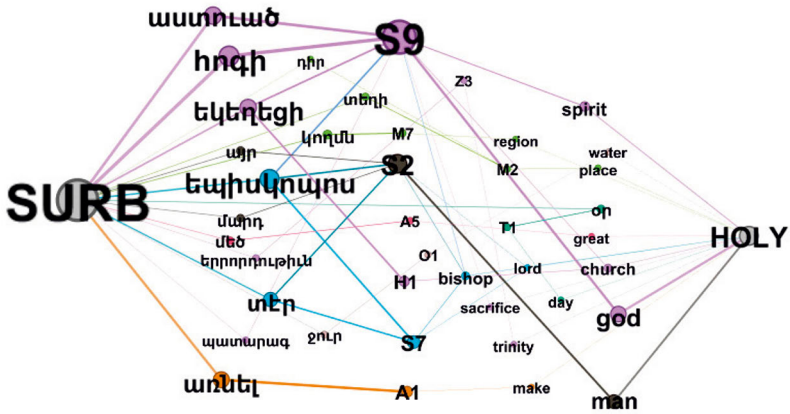
6.2.1.4 Examination

Intersection of Words

Only 14 of 147 individual words that appear in the HSEC synoptic tables are also part of the synoptic tables of the AAFTC/AASC (approximately 10%). From the 111 individual words in the AAFTC/AASC synoptic tables, only 16 are part of an intersection with the HSEC synoptic tables (approximately 15%). All domains in the “domain” annotation scheme occur in this intersection of words. However, words related to the “concept” domain (AF 5, RF 31%) and the “place” domain (AF 3, RF 19%) are dominant.

Regarding the “semantic field” annotation of the words in the intersection, words related to “religion and the supernatural” (S9) are conspicuous (AF 5, RF

Figure 80: Intersection of words and their “semantic fields” between the HSEC and the AAFTC/AASC. Different colors indicate different “domains.” (c) Gephi.



31%).⁵ Most of these terms are very general and usually found among the top entries in at least one of the corpora. Among the religious words in the intersection are “god” (աստուած), “spirit” (հոգի), “sacrifice” (պատարագ), and “church” (եկեղեցի). In the “person” domain, the word “man” (Arm. այր and մարդ) appears in both corpora. Concerning the “place” domain, the very general spatial concepts “place” (տեղի) and “region” (կողմն) are listed.

Many of the words in the intersection are not equally important in both corpora. For instance, the word “place” (տեղի) is approximately eleven times more important in the Armenian corpus than in the English corpus. The same holds for many other words in the intersection. Most of them frequently appear in either the English or the Armenian corpus. The only exceptions are “god” (աստուած) and, to some extent, “spirit” (հոգի) and “man” (այր and մարդ), which are important in both corpora.

Overall, an examination of the intersection of words illustrates that both corpora only share a few and mostly very general words. Most prominently and of equal significance in both corpora are religious “concepts,” such as “god” and “spirit.” The absence of words that hint to a more specific use case of *holy* or *surb*,

5 Note that other religioid or usually religiously used words such as “lord” (S7) or “trinity” (Z3) are not assigned to the “religion and the supernatural” (S9) category by the USAS tagger. Thus, the overall amount of religious words in this intersection is even higher than 29%.

with the slight exception of “water” (ջուր), indicates the very diverse use of the terminology between the Armenian and the English corpora.

Intersection of Domains

The intersection of domains from the “domain” annotation scheme reveals that the “concept” category is central in both corpora. Yet, since the “concept” domain is very broad, there are no immediate conclusions that can be drawn from this observation. The “person,” “title,” and “verb” domains are much more relevant in the AAFTC/AASC compared to the HSEC. The “place,” “time,” and “object” domains are equally (ir-)relevant in both corpora. The “person,” “title,” and “verb” domains are interrelated and can be deemed as a strong indication of the overall relevance of religious persons and their deeds in the AAFTC/AASC. The focus on persons and their acts, however, seems to be rather unimportant in the HSEC.

Intersection of Semantic Fields

A closer look at the graph and table displaying the intersection of “semantic fields” offers a more differentiated comparative perspective on the data. The semantic field “Religion and the supernatural” (S9) is the essential “semantic field” in both the HSEC and the AAFTC/AASC. Semantic fields related to “time” (T1) are almost equally present in both corpora. Several categories from the “semantic field” annotation are much more present in the Armenian data than in the HSEC. Among these categories are “Architecture” (H1), “Places” (M7), “People” (S2), “Power relationship” (S7), and “Personal names” (Z1). The dominance of the “Personal names” (Z1) category and several categories that are related to verbs (such as Q2) in the AAFTC/AASC underlines the overall importance of individual persons and their deeds in the Armenian data. On the other hand, “Substances” (O2), “Life and living things” (L1), and “Mental Actions and Processes” (X2) are more present in the English corpus. Furthermore, it is interesting to notice the “semantic fields” that do not appear in the merged table but which are highly relevant in only one of the respective corpora. Among these categories is the “Body and anatomy” (B1) semantic field that is important in the HSEC tables but does not appear in the Armenian data.

6.2.1.5 Summary

An inquiry of the intersections between the “domains,” “semantic fields,” and words of the HSEC and AAFTC/AASC revealed significant differences between the use of *surb* and *holy* in both corpora. In particular, the intersection of words only yields very general terms that mostly stem from the semantic field of religion. This shows that most of the more specific use cases are not shared between the corpora.

Overall, these general religious notions indicate that the semantic field of religion functions as the major link between the semantic fields of *surb* and *holy*.

In the AAFTC/AASC, the terminology around *surb* is frequently used in the context of religious “concepts” and individual religious “persons.” In the HSEC, terms related to *holy* are also used in the context of religious “concepts,” but they are less often associated with individual “persons.” Instead, *holy* is applied in the context of very general and abstract terms that belong to “objects” (“grail”), non-religious “concepts” (“life,” “nature”), “places” (“cosmos,” “earth”), or feelings such as “awe.”

An interesting link between both terminologies that is primarily visible in the microanalyses is purity. The notion of *surb* was associated with diverse forms of purity, and the relationship between both semantic fields was observable in most parts of the Armenian data. Yet, the relationship between the semantic field of *holy* and purity was evident in the HSEC Twitter Data only to a lesser degree.

The term *surb* seemed to be more frequently used to function as a specific religious marker that made a tangible difference akin to other attributes that appeared in its vicinity, such as “virtuous,” “chaste,” or “worthy.” The lexical fields in the context of *holy* seemed to be more sedimented in fixed expressions where they only evoke a generic connotation of religion. Yet, the use of fixed expressions was also perceivable in the Armenian data, showing that the process of sedimentation started relatively early. That the use of *holy* or *surb* in fixed expressions always coexisted with a more dynamic application in the form of an attribute was observable in the Armenian and in the English data. Yet, the use of *surb* as a dynamic attribute was dominant in the Armenian data, whereas the English use of *holy* was primarily perceivable in the use of fixed expressions. This result illustrates the chronological dimension of sedimentation processes.

Finally, the question remains in how far the two semantic fields of *surb* and *holy* are related to each other. According to the results of the analysis of the merged synoptic tables with the intersections of “domains,” “semantic fields,” and words, there are more differences between the use of *surb* and *holy* than similarities. Thus, by only assessing the values and overall numbers from the statistical analysis, it is tempting to declare that both semantic fields do not have a lot in common, except for some very general overlappings, such as “god,” religion, etc. Besides, the diverging use cases and meanings of *surb* and *holy* were also underlined in the microanalyses in Chapters 5 and 6.

Yet, the assumption that both semantic fields only have little in common is too simplistic. The first aspect that challenges this assumption pertains to the structure of both corpora. As has already been mentioned in Chapter 4, the HSEC and the AAFTC/AASC include very different text genres. The diversity of texts in the corpora might consequently have led to the result in very few overlappings between the use of *surb* and *holy* in the different (sub-)corpora. In turn, this dichotomy can also be interpreted as a hint at the distinct use of *surb* and *holy* between different text genres rather than as the evidence of the general incompatibility of both notions.

The second point that challenges the assumption that both terminologies are not related to each other concerns the overall methodology of this book. Even though this book also includes extensive microanalyses, the global focus of the examination and comparison lies in the text statistics. Even though quantitative analyses have revealed compelling results, relying on this method alone can be misleading. The microanalyses in Chapter 5 have demonstrated that relying on quantitative statistical results can sometimes lead to partially wrong assumptions, for instance, regarding the importance of purity. Thus, the quantitative results of the comparisons need to be reviewed from a qualitative perspective as well.

From a micro perspective, the few notions that appear in the merged synoptic tables above function as stable links between *surb* and *holy*. Telling examples are the “holy spirit,” “god,” and the notion of the “church.” All these notions are present in the majority of texts in the HSEC and the AAFTC/AASC. Furthermore, they are important from both a quantitative and qualitative perspective. They consequently build very few but also central and stable intertextual and diachronic links between the two semantic fields.

In sum, the combination of statistical macroanalyses and qualitative microanalyses has pointed out that the concrete use cases and thus meanings of *surb* and *holy* were often very different between the corpora. Among others, this was observable in:

- the dichotomy of attribute chains (AAFTC/AASC) versus the use of *holy* in fixed expressions (HSEC),
- the attribution to (mostly individual) persons in the Armenian data versus the preferred attribution to very abstract concepts in the English data, as well as
- a strong interrelation between purity and *surb* in the AAFTC/AASC that was mostly missing in the HSEC (except for the Twitter data).

Yet, from a qualitative perspective, the few but important linking words, such as “god,” “church,” and “spirit,” as well as the intersections of “domains” (concept) and “semantic fields” (religion) between both corpora also revealed the existence of few but central links between both notions.

6.2.2 Comparison II: *Holy* in the Study of Religion

In this last section of the comparative chapter, the semantic fields of *surb* and *holy* are compared to the semantic field of *holy* in the study of religion. The semantic field of *holy* in the study of religion was described in Chapter 2 as including a great variety of different associative contexts, among others “holy as power” and “holy as an aesthetic category.”

In this section, each domain from the semantic field of *holy* in the study of religion (displayed in figure 81) is compared to the results of the previous examinations of the HSEC and the AAFTC/AASC data (Chapters 5 and 6.1). This comparison should determine whether the associative contexts of the notion of *holy* in the study of religion reflect the actual use cases of the terminology in the corpora of this book.

6.2.2.1 Holy as an Aesthetic Category (Feelings)

Surb in the Armenian AAFTC/AASC

The idea that *holy* is “used as an adjective to characterize external objects that cause impressions, feelings, and emotions of a particular kind (via an aesthetic input)” (see Chapter 2) or that *holy* itself is a particular feeling triggered by aesthetic inputs is not present in the Armenian data in this book. There are neither any words in the synoptic tables nor results in the microanalyses of the Armenian data that hint at an aesthetic function of *surb*.

Holy in the English HSEC

The relationship between aesthetics and *holy* is mainly perceivable in the HSEC Britannica, the HSEC Eliade, and the HSEC Otto.

In the HSEC Britannica, words such as “appear,” “nature,” “illumination,” and “manifestation” hint at an aesthetic connotation of the notion of *holy*. In the HSEC Eliade, words such as “experience,” “reality,” “stone,” “tree,” and “manifest” mark external objects in which “the sacred” can manifest itself. These objects can then be experienced by humans via an aesthetic input.

The majority of words related to *holy* in the context of aesthetics and feelings are found in the HSEC Otto. In the corresponding synoptic table of this subcorpus, terms such as “feeling,” “experience,” “sense,” “manifest,” “awe,” “consciousness,” “emotion,” “impression,” and “atmosphere” indicate a close relationship between *holy* and feelings as well as certain ways of aesthetic perceptions (“impression”). That the HSEC Otto is of particular importance in this context is not surprising. Rudolf Otto’s work “The Idea of the Holy” with its focus on the “numinous” defined as *holy* without its moral and rational trait (Otto and Harvey 1923, 6) was important for the integration of feelings and aesthetics in the modern academic notion of *holy*.⁶

The HSEC Eliade is also relevant for the interrelation between *holy* and aesthetics since it coined the idea of the *holy* and its diverse manifestations in (natural)

6 For the importance of “nature” or natural objects/events in the context of *holy*, see Otto and Harvey (1923, 148ff.) and Schleiermacher (1980 [1806]) as well as the literary quotes in the corresponding section in Chapter 2.

objects and places (“hierophany”). This implies that these appearances of *holy* exert an aesthetic influence, too.

Notably, the connection between *holy* and feelings and aesthetics only plays a role in those sources that stem from an academic background. The religious data such as the HSEC Canon Law or the HSEC Twitter Data (religious) barely evoke this associative context. Consequently, the idea that *holy* is related to feelings and aesthetic impacts seems to be a scholarly discussion that is not reflected in the majority of religious texts that mention *holy*.

6.2.2.2 Holy as a Synonym for Religion

Surb in the Armenian AATC/AASC

Much like the observations in the previous section, the idea that *holy* (*surb*) functions as a synonym for religion is not mirrored in the Armenian data. Instead, the terminology related to *surb* was usually applied to underline specific religious characteristics. Thus, it was almost exclusively associated with religion, but it was neither considered as a synonym for religion nor expressed as being fundamental for the understanding or functioning of religion.

Holy in the English HSEC

The idea that *holy* is a synonym for or an essential part of religion was particularly observable in the HSEC Eliade, the HSEC Britannica, and the HSEC Otto. Akin to the previous discussion of “Holy as an Aesthetic Category (Feelings),” the idea that *holy* could replace religion was not observable in the religious sources, such as the HSEC Twitter Data (religious) or the HSEC Canon Law.

The practically synonymous use of “religion” and “holy/sacred” in the HSEC Eliade has been examined in the microanalysis of the “concept” domain in the respective section of Chapter 5. The close connection between both notions was demonstrated via the interrelation between “religious man” and “profane man” in Eliade’s “The Sacred and the Profane,” thereby illustrating that the *sacred/profane* dichotomy is represented by *religious/profane* in Eliade’s work. Furthermore, the synoptic table revealed that many abstract concepts such as “cosmos,” “world,” “space,” “life,” and “time” appeared in the overall account of “The Sacred and the Profane,” thereby underlining the encompassing character of *holy* in Eliade’s influential book.

The relationship between religion and *holy* is equally close in “The Idea of the Holy” by Rudolf Otto. Among others, this is observable in the outstanding “sum” value of “religion” in the HSEC Otto synoptic table. As discussed in the corresponding part of the microanalysis, Otto’s examination of *holy* (and the “numinous” in particular) goes hand in hand with his approach to establish an understanding of religion as a category *a priori* (Otto 2004 [1917], 202ff.). This concurrence of reli-

gion and *holy* or better the fundamental character of the *holy* for religion is visible throughout the entire work, among others in quotes such as the following ones:

Plainly, then, Religion is only the offspring of history in so far as history on the one hand develops our disposition for knowing the holy, and on the other is itself repeatedly the manifestation of the holy. (Otto and Harvey 1923, 181)

It will be our endeavour to suggest this unnamed Something [the “numinous”] to the reader as far as we may, so that he may himself feel it. There is no religion in which it does not live as the real inner most core, and without it no religion would be worthy of the name. (Otto and Harvey 1923, 6)

6.2.2.3 Holy and Sacralization Processes

Surb in the Armenian AAFTC/AASC

The difficulty of tracing sacralization processes and religioid applications of religious notions in the Armenian material has already been discussed in the section on “sacralization processes” in Chapter 2. This problem is largely due to the few sources from the fifth century CE that almost exclusively include texts written from a Christian and thus religious perspective. Consequently, there are no historical Armenian texts that could be compared to the HSEC Twitter Data or other contemporary language samples that naturally have a higher potential of also including religioid use cases of notions such as *holy*.

Yet, it is remarkable that the notion of *surb* in the Armenian sources is almost exclusively used in a religious context. For instance, even though kings and other political or military leaders often played an important role in the Armenian sources such as the “Epic Histories” or the “History of the Armenians,” they were never called *surb*. Unlike other more ambivalent notions and attributes such as “worthy” or “great,” the attribute of *surb* was only granted to persons that held a high position in the church hierarchy or were regarded as martyrs or otherwise important religious persons.

Consequently, sacralization processes such as those defined in Chapter 2 were not observable in the context of *surb* in the Armenian sources. With the exception of very few examples of possibly older use cases of *surb* in the sense of (physical) “cleansing,” the terminology was almost continuously used in a religious context. This observation supports the assumption that sacralization processes are rather modern phenomena. Yet, the idea that *surb* could potentially have also been used in a religioid or non-religious sense cannot be entirely abandoned. A point in case is the use of religious terms or names of deities in expressions such as *meherculē(s)* in other ancient languages such as Latin to convey one’s surprise or assertion. To some extent, these formulations are comparable to the contemporary non-religious use of *holy* in expressions such as “holy crap.” The Latin example illustrates that a

non-religious use of religious notions did also exist in ancient languages and that it might thus also be possible in the Armenian case. However, we do not possess any examples in the sources. Furthermore, the question remains in how far such expressions can be regarded as religioids since they neither represent a “kollektive Verständigung über das Heilige” (see below) nor include a realistic potential to switch back to a religious meaning.

Holy in the English HSEC

The application of *holy* as a *Diskurstopper* or expression to mark something as extremely valuable (see Chapter 2) seldom appeared in the English data. Yet, expressions such as “holy grail” in the HSEC Twitter Data (non-religious) partially served this function.

In both the non-religious and religious data, *holy* was commonly used in fixed expressions such as the already mentioned “holy grail,” “holy see,” “holiday,” “holy spirit,” “holy crap,” “holy land,” etc. Depending on the context, many of these fixed expressions had a religioid meaning. Telling examples are the “holy days > holidays” and the “holy land.” Both expressions could be used as *termini technici* denoting an institution or a territory, even in religious texts such as the HSEC Canon Law. Yet, they also kept a strong potential to switch back to a more religious meaning or to constantly oscillate between a religious and a non-religious sphere.

A particular case of a non-religious use of *holy* was the omnipresent phenomenon of the application of *holy* in expressions of surprise or disgust such as “holy shit” in the HSEC Twitter Data. Even though these instances represent examples of the non-religious use of *holy* in fixed expressions, they do not fit into the categories of “sacralization processes” or “religioids” since they neither represent a “(...) kollektive Verständigung über das Heilige [im] nicht-religiösen Bereich” (Schlette and Krech 2017, 437) nor do they hold a potential to oscillate between a religious and a non-religious meaning.

6.2.2.4 Holy and Profane

Due to the importance of Durkheim’s works and their ongoing reception in the study of religion, the *holy* and *profane* distinction still plays a significant role in the detection of a semantic field of *holy* in both contemporary and historical cultures (see Lutzky and others in the section on “Historical Notions of Holy” in Chapter 2). Thus, the presence or absence of such a dichotomy can rightfully be considered as a strong indication of the existence or absence of a general semantic field of *holy*.

Surb in the Armenian AAFTC/AASC

The *holy* and *profane* dichotomy existed in most of the texts in the AAFTC/AASC. In the context of the Gospels and the other Armenian sources, the term *surb* (սուրբ)

is often contrasted with the term *pilc* (պիղծ) that functions as its antonym. Just as *surb* has a close relationship to purity (see below), the notion of *pilc* has a strong connotation of impurity. It is often found in the proximity of *ays* in the Gospels to denote the “impure spirit” (*ays pilc*) in opposition to the “holy/pure spirit” (*surb hogi*).

In the AAFTC/AASC Gospels, both terms appear as translations in the context of the Greek dichotomy of *kat' arós* (Arm. *surb*) and *akat' arós* (Arm. *pilc*). In the other Armenian sources such as AAFTC/AASC Agat' angelos, the dichotomy of *surb* and *pilc* occurs as well. The following quote that was also discussed in the course of the microanalysis of the AAFTC/AASC Agat' angelos provides a revealing example.

Իսկ երանելին պարկեշտասէրն Գայիանէ, սրբասնելովն Հռիփսիմեաւ հանդերձ եւ այլ ընկերաքն իւրեանց, յիշեալ զուխտն սրբութեան, զաւրինաւոր կրանիցն զգաստութեան սրբութիւնն յոր մտեալ էին՝ ողբային յանձինս իւրեանց վասն նկարացոյց հրամանին, առ ի նկատել զպատկերս նոցա պիղծ եւ անաւրէն թագաւորին. (§143)

But the blessed and chaste Gaiane, with the saintly Rhipsime and their other companions, remembered the covenant of holiness, the religious rule of chastity into which they had entered, and lamented amongst themselves over the impure and impious emperor's command to have their portraits painted.

In this passage, the “saintly” (սրբասնելովն, *srbasnelovn*), “chaste” (պարկեշտասէրն, *parkeštasêrn*), and “blessed” (երանելին, *yeranelin*) saints Hřip'simê and Gayianê are opposed to the “impure” (պիղծ, *pilc*) and “impious” (անաւրէն, *anorên*) emperor Diocletian. This example illustrates the existence of an Armenian dichotomy that can justly be related to a *holy* and *profane* distinction in its contemporary understanding, although with a stronger focus on (im-)purity. The focus on impurity, however, is not a problem since the precise meaning of the *holy* and *profane* distinction can differ between cultures and religions; yet, it needs to keep the function of *separation-binding* (see Chapter 2) to be counted as an instantiation of the *holy* and *profane* distinction. The sentence from Agat' angelos also demonstrates that the dichotomy of *holy* and *profane* is not restricted to the use of two precise notions. It can also manifest itself in the form of attribute chains from which it then emerges. In this example, all the positive attributes granted to the two female saints stand for *holy* whereas the several negative characteristics of the emperor portray him as *profane*.

The terminology of *surb* and *pilc* is mostly used in the context of persons where it serves as a concrete attribute that ascribes specific qualities to the persons. Yet, the constant appearance in the proximity of “spirit” (Arm. *hogi*) also illustrates that *surb*, as well as *pilc*, were already part of fixed expressions (see also the contemporary use of *holy* in the HSEC). However, there is no indication of an application of

the Armenian *surb/pilic* dichotomy that characterizes holistic and abstract concepts such as the *Holy* versus the *Profane* (see also the discussion of “Holy as a Synonym for Religion” in this section).

Holy in the English HSEC

The term “profane” only appears in the synoptic tables of the HSEC Britannica and the HSEC Eliade. The occurrence of “profane” in the latter corpus is not surprising since the dichotomy of the *sacred* and the *profane* is already found in the title of Eliade’s “The Sacred and the Profane.” The appearance of “sacred” and “profane” in the HSEC Britannica can be explained by the overall importance of this distinction for the understanding of the contemporary notion of *holy* in the study of religion.

The terms “sacred” and “profane” in the HSEC are mostly applied to differentiate very abstract concepts and dimensions. As discussed in the corresponding examinations of the HSEC Eliade and the HSEC Britannica in Chapter 5, the terms “sacred” and “religious” are often used interchangeably. In this context, the counter notion “profane” is regularly applied to describe a “profane dimension/worldview” that is seen as opposed to a “sacred dimension/worldview.” Thus, the dichotomy of *sacred* and *profane* is present in the English texts as well. Unlike in the Armenian sources, however, it is rarely employed to underline concrete opposing attributes of persons, concepts, or objects (such as “pure” and “impure”), but rather as a marker to differentiate between two entirely different states of existence (religious vs. non-religious).

An interesting exception is the use of *holy* in expressions such as “holy crap” in the HSEC Twitter Data. In these phrases, the dichotomy between *sacred* and *profane* is present as well, but in the form of different terms and with a focus on a purity/impurity distinction. On the one hand, these examples illustrate that a more concrete understanding of the dichotomy between *holy* and *profane* can coexist with an abstract and holistic use of the terminology. On the other hand, it also underlines that the *holy* and *profane* distinction does not need to be instantiated in the form of two concrete terms, such as “sacred” and “profane.” Instead, the *holy* and *profane* distinction can emerge from the use of different words that belong to opposed (societal) spheres (in this case purity and impurity) and which do not need to be explicitly labeled as “sacred” and “profane” or “pure” and “impure.”

6.2.2.5 Holy as Power

Surb in the Armenian AAFTC/AASC

Even though the direct connection between *surb* and power is only hardly visible in the synoptic tables of the Armenian sources, it is indirectly reflected in the high frequency of the “holy spirit” and persons that are called *surb* (“saints”). The “holy spirit” has already been characterized as a compelling representation of the power

of *holy* as it is usually the agency of the “holy spirit” that intervenes in human life and that is evoked in prayers. The power of the *surb* persons, such as Agat’angelos or Hřip’ simê in the “History of the Armenians,” is faintly perceptible in the synoptic tables, or maybe to a certain extent in words such as “to intercede.”

Nevertheless, the way these persons are described in the respective accounts illustrates their might and their extraordinary capabilities (see the microanalyses of the AAFTC/AASC data). For instance, the deeds that the *surb* persons fulfill range from the violent destruction of former pagan sites and the expulsion of evil spirits to the successful resistance of *surb* Hřip’ simê against the violent approaches by mighty emperors such as Trdat or Diocletian. Thus, even though the notion of *surb* is not directly associated with *power*, its close connection to concepts and persons that act with great power illuminates the relationship between *holy* and power throughout all the sources in the Armenian corpus.

Holy in the English HSEC

That the idea of *holy* is closely related to power is visible in different texts of the HSEC, among them HSEC Encyclopaedia Britannica and HSEC Otto. Similarly to the use in the Armenian sources, the power of *holy* is visible in its connection with “spirit,” a powerful actor whose ascribed influence is particularly evident in the prayers in the HSEC Twitter Data (religious). Unlike in the Armenian sources, however, the attribution of power to *holy* does not mainly derive from its attribution to influential persons. Instead, the might of *holy* is more explicitly expressed in the form of a direct connection between power and *holy*. Telling examples in this context are the definitions in the HSEC Britannica where *holy* is explicitly defined as a “power” (see the microanalysis of the HSEC Encyclopaedia Britannica in Chapter 5). Another example is Rudolf Otto’s “The Idea of the Holy,” where the author states:

It is not that the awe of holiness is itself simply fear in face of what is absolutely overpowering, before which there is no alternative to blind, awe-struck obedience. *Tu solus sanctus* is rather a paeon of praise, which, so far from being merely a faltering confession of the divine supremacy, recognizes and extols a value, precious beyond all conceiving. The object of such praise is not simply absolute Might, making its claims and compelling their fulfilment, but a might that has at the same time the supremest right to make the highest claim to service, and receives praise because it is in an absolute sense worthy to be praised. (Otto and Harvey 1923, 54)

6.2.2.6 Holy as Perfection

Surb in the Armenian AAFTC/AASC

The terminology of *surb* in the Armenian sources is not applied to denote or reflect a state of abstract moral or ethical perfection as in some of the modern sources. Yet, the already noted use of *surb* in the context of attribute chains related to ex-

traordinary persons (“saints”) links the two spheres of *surb* and perfection, at least indirectly.

The frequent occurrence of (individual) “saints” in the synoptic tables and words such as “chaste,” “virgin,” “virtuous,” “great,” and “golden” bespeak the existing connection between *surb* and perfection. Most of these words stem from the synoptic tables of the HSEC Agat’angelos and the HSEC Epic Histories. They are mainly used in attribute chains to characterize extraordinary religious persons (“saints”). Many of these attributes hint at an exceptional moral or ethical behavior (“chastity,” “virtuosity”) of the respective persons, thus underlining their aspiration to moral and ethical perfection.

Another word that evokes the association of perfection in the Armenian material is the frequent appearance of “god” or “divinity” in the direct vicinity of *surb*. As has been outlined in the examinations of the Armenian sources, the term “god” is frequently used as another attribute in the attribute chains such as in “(...), a *surb* and virtuous man of God, (...)” Since the Christian god is commonly regarded as being perfect, the close connection between “god” and *surb* also transfers the perfection of god to the latter notion. This observation is very much in line with the Christian notion of *holy* as elucidated in Chapter 2.

Holy in the English HSEC

The words in the synoptic tables of the English sources are generally not linked to a semantic field of perfection, even to a lesser degree than in the Armenian sources. However, this does not mean that the idea of *holy* as perfection is not mentioned or discussed in the English sources. As has already been pointed out, Rudolf Otto was significantly influenced by Immanuel Kant, among others, and therefore drew upon Kant’s idea of *holy* as moral perfection in his works as well (Otto and Harvey 1923, 5). Yet, in the sources that build the *Holy Sacred English Corpus*, the idea of *holy* as an abstract moral or ethical perfection does not seem to play a significant role. This observation is supported by the lack of corresponding words in the synoptic tables and by the absence of this topic in the microanalyses (with very few exceptions).

6.2.2.7 Holy and Persons

Surb in the Armenian AAFTC/AASC

Next to the “concepts” domain, the “persons” domain is the most important category in the AAFTC/AASC. Except for the Gospels, persons appear to a large extent in all of the Armenian subcorpora. The importance of individual persons and groups of persons for the overall understanding of *surb* has already been pointed out. Furthermore, it has previously been stated that the focus on *surb* and “persons” is very much in line with the broader developments in the context of *holy* in early Chris-

tianity.⁷ Consequently, the above-mentioned findings and assessments will not be repeated.

Holy in the English HSEC

The observation that individual “persons” are relatively unimportant in the HSEC data has already been thematized, among others in the microanalyses of the English corpora. One reason why this attribution is only rarely perceivable in the English data is the fact that the English terminology related to *holy* has a specific term for holy persons, namely “saint.” Yet, this term as well appeared only rarely in the sources, thereby underlining the overall loss of significance of “holy persons” for the use and meaning of *holy* in contemporary English.

6.2.2.8 Holy as a Distinct Religious State and Category of Research (Phenomenology)

Surb in the Armenian AATC/AASC

The assumptions that the examination of *surb* would demand a “switching on” of the religious phenomenon “in one’s own life” or an otherwise particular phenomenological standpoint of the scholar to “grasp its essence” (Michaels and Bergunder 2019) cannot be regarded as valid for the examination of the AATC/AASC. Instead, the examination of *surb* yielded meaningful results by applying methods from the field of corpus linguistics in the macroanalysis and a close reading of smaller text samples in the microanalysis.

Overall, the use of *surb* in the Armenian sources does not convey the impression that it is “wholly other” and not approachable by the human intellect. Quite the contrary, *surb* is very often applied as one among (many) other attributes to describe and characterize persons, concepts, and objects, thereby implying that it needs and wants to be understood by the reader. Yet, what *surb* in each context specifically meant is another and more complicated question. I have tried to highlight some important meanings throughout this book, above all in the corresponding section in Chapter 6.1.

Holy in the English HSEC

In the case of the English data in the HSEC, one needs to distinguish between the religious texts represented in the HSEC Twitter Data (religious) and the HSEC Canon Law on the one hand and the academic sources represented in the HSEC Eliade, HSEC Otto, and HSEC Encyclopaedia Britannica on the other hand. That the idea of *holy* as a distinct religious state was echoed in the HSEC Otto and the

7 Among others, see Chapter 2 and the microanalyses of the “person” domain in the Armenian sources.

HSEC Eliade and, to some extent, in the HSEC Encyclopaedia Britannica is not surprising. All these sources expound *holy* and construe it as a distinct religious state and category of research in the contemporary study of religion.

It is remarkable, however, that the assumed “wholly otherness” of *holy* is not observable in the contemporary religious sources in the HSEC. Rather, the everyday use of *holy* and the institutionalized religious application of *holy* in the Canon Law consist of fixed expressions or the use of *holy* in other contexts that are far from being “wholly other” or “not graspable from a rational point of view.”

This observation underlines the importance of a question posed in Chapter 2, namely whether the overall idea of *holy* as the “wholly other” can be regarded as an empirically observable religious phenomenon. The results of this book compel us to consider the entire idea of *holy* as the “wholly other” as a construct of an academic, in parts crypto-religious, *Gedankenspiel* that finds little justification in the majority of religious sources.

6.2.2.9 Holy as Purity

Surb in the Armenian AATC/AASC

Even though words related to purity are among the entries in the synoptic tables,⁸ the relevance of purity for the understanding of *surb* was particularly manifest in the microanalysis. As has already been explained in the examinations of the Armenian data and in the respective section in Chapter 6.1, *surb* was related to physical, spiritual, sexual, and ethical purity as well as healing, thus covering a broad spectrum of the complex semantic field of purity. The observation that the Armenian words related to *surb* were applied to translate Greek terms usually associated with purity further underlines the importance of purity for the understanding of *surb* or even the convergence of both semantic fields in the Armenian terminology.

Holy in the English HSEC

Interestingly, a direct connection between purity and *holy* almost exclusively appeared in the non-religious use of the terminology in the HSEC Twitter Data (non-religious). Yet, in particular the application of *holy* in fixed expressions such as “holy see” or “holy spirit” might still include aspects of purity as well. However, similarly to the observations in the different lemmata from academic publications on *holy*, the context of purity in these contemporary cases is only one among many other associative contexts of the complex notion of *holy*. One goal of this book is to draw more attention to the often neglected relationship between purity and *holy* in concurrent applications of *holy*, particularly since it seems to be overshadowed by other associative contexts of *holy*, such as “perfection,” “power,” or “religion.”

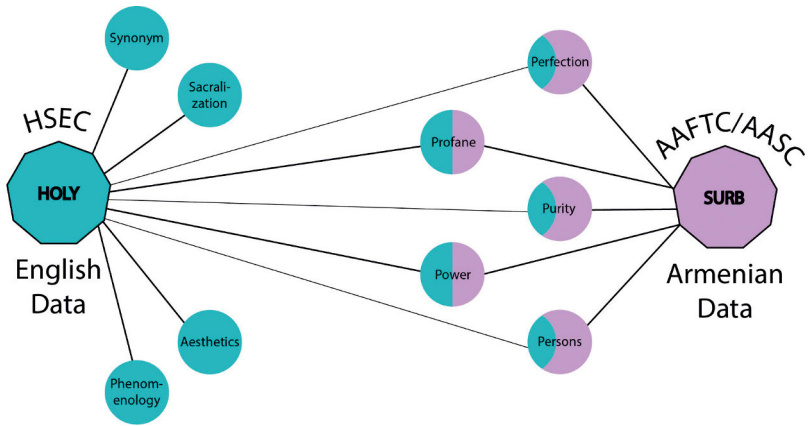
8 Among others in terms such as “chastity” or “water.”

6.2.2.10 Summary

The visualization in figure 81 sketches the relationship between the notions of *surb* and *holy* and the semantic field of *holy* in the study of religion. Unlike most of the other graphs in Chapter 6.2.1, the visualized relations are not based on numbers or results of statistical analyses. Instead, they represent a manual attempt to sketch the results of the previous evaluations in this section.

The idea that the semantic field of *surb* is different from the semantic field of *holy* (see the summary of Chapter 6.2.1) was also substantiated by the study of the notion of *holy* in the study of religion as a *tertium comparationis* between *surb* and *holy*. Accordingly, the disparities between both semantic fields are also reflected in the graph in figure 81.

Figure 81: Visualization of the connections between *surb* (AAFTC/AASC) and *holy* (HSEC) when compared to the notion of *holy* in the study of religion as a *tertium comparationis* (see Chapter 2).



On the one hand, associative contexts such as “Holy as a Synonym for Religion,” “Sacralization Processes,” “Aesthetics,” and “Phenomenology (Feelings)” were almost absent in the application and signification of *surb*. On the other hand, topics such as “Holy and Perfection,” “Holy and Purity,” and “Holy and Persons” are much more present in the Armenian data than in the application of *holy* in the English corpus. The only two domains of equal importance in both corpora are “Holy and Profane” and “Holy and Power.”

Since many associative contexts of the comparative concept of *holy* are only present in either of these notions, it is tempting to conclude that Armenian *surb* and English *holy* only have little in common. Similar to the investigations in Chapter 6.2.1, however, a merely quantitative approach does not do justice to the evalua-

tion of the conjunctions between the three semantic fields. Even though it is true that the notions of *surb* and *holy* respectively mostly refer to different aspects of the comparative notion of *holy*, the fact that the essential associative contexts of “Holy and Profane” and “Holy as Power” are shared between all three notions is a strong indication that they are interrelated. In particular, the crucial significance of the “Holy and Profane” (see Chapter 2 “Holy and Profane” and the section on the historical notions) and its presence in both corpora outweigh the several differences. The “Holy and Profane” distinction has a similar standing than the “holy spirit” in Chapter 6.2.1. It builds a link that binds all three semantic fields together, despite all their differences regarding more specific use cases or associative contexts.

Thus, according to the analysis in this section, the three semantic fields have only a few but sufficiently strong overlappings to regard them as interrelated. Yet, the observation that the majority of associative context from the comparative notion of *holy* were relevant for only either of the two notions also highlights the underlying differences between *surb* and *holy*.

The fact that the comparative notion of *holy* was able to bring the overlappings between *surb* and *holy* to the fore while also pointing out their differences demonstrates the notion's inherent potential for comparative research in the study of religion. One reason why the current comparative notion of *holy* captures both the differences and the common qualities of diverse notions such as *holy* and *surb* needs to be envisaged within the ongoing discourse in academia about this *tertium comparationis* (see Chapter 2). Even though there is room for further adjustments (see below), the diversity of the approaches, which include historical-philological (Burkert), sociological (Durkheim), as well as philosophical approaches (Kant and Otto), have rendered the comparative category of *holy* flexible enough to enable a differentiated discussion of historical and contemporary examples.

Nevertheless, the results of this book invites to revisit two issues regarding the comparative notion of *holy*. The first issue concerns the association of *holy* with purity. Even though this relationship is often considered in the contemporary discourses on *holy*, its overall importance lags behind other topics, such as “Holy and Feelings,” “Holy and Persons,” or “Holy as a Synonym for Religion.” However, the investigation of the Armenian material has demonstrated that *holy* (or *surb*) is often more closely related to purity than it might be anticipated against the background of the current understanding of *holy*. In the early Armenian context, the connection between *surb* and purity manifests itself primarily in the ecclesiastical understanding of *surb* as a designation for something that is redeemed from impurity by the *holy* God Jesus Christ and is therefore exemplarily Christian. Particularly in those cases where a connection between purity and *holy* is less evident than in the Armenian material, relying on the current comparative notion of *holy* as outlined in Chapter 2 might become problematic. In such a case, a scholar might either overlook the potential connection between a notion with a strong focus on purity and

the semantic field of *holy* or the relevance of purity for a notion that is closely related to *holy*.

The second issue concerns associative contexts such as “Holy as a Distinct State of Mind” or “Holy as a Synonym for Religion.” At least in the corpora used in this book, these subfields of the comparative notion of *holy* were barely visible in the religious use cases of both *surb* and *holy*. Instead, they were emphasized in the non-religious texts concerned with *holy* (Otto, Eliade, *Encyclopaedia Britannica* entry). This observation raises the question of how far such abstract and often encompassing conceptualizations of *holy* can legitimately be understood as central for the understanding of the terminology in religion. Maybe they should instead be seen as academic and partially semi-religiously driven philosophical attempts that constitute a separate line of thought in the overall history of the notion of *holy*.

This line of thought, however, is only barely traceable in the contemporary and historical religious sources examined in this book. It remains to be asked whether the encompassing understanding of *holy* is perhaps only absent in certain text genres but more visible in others. One might think of historical and contemporary mystical texts or theological writings, in which the conceptual space around *holy* might be very different from that in the texts examined here. The investigation of this hypothesis remains a desideratum, which could not be accomplished in the context of this book for several mostly pragmatic but partially also legal reasons (see Chapter 3 the discussion about the practical bias of the corpus composition).

7. Conclusion

This conclusion starts with an outline of the results of the individual examinations of *surb* and *holy* in Chapter 6.1 In a second step, I will discuss the interrelations between these two semantic fields and their overlappings with the comparative notion of *holy* in the study of religion. The third section evaluates what the study in this book has contributed to the following discussions:

- The understanding of the notion of *surb* in early Armenian texts.
- The broader discussions on early Christian notions of *holy*.
- The systematic research of comparative notions in the study of religion.

The final part of the conclusion provides a critical review of the methodological approach in this book.

The following sections will not repeat every single detail from the previous summaries in Chapters 5 and 6. Instead, the focus lies on bringing the various results together and to assess them in context of each other.

7.1 Results of the Individual Examinations

The first research focus of this book was concerned with the meaning of the notion of *surb* in early Armenian texts from the fifth century CE. In the course of this analysis, the meaning and application of the English notion of *holy* were also examined to set up the two comparands for Chapter 6. The investigations of both semantic fields included macro- and the microanalyses of the AAFTC/AASC and the HSEC data, which yielded a differentiated picture of both semantic fields. The detailed overviews of the semantic fields of *surb* and *holy* can be found in the final section of Chapter 6.1.

The first important result of the examination of *surb* was the observation that its general use was very much in line with the application of other early Christian terminologies commonly related to *holy*, such as Greek *hágios* and Latin *sanctus* (see Chapter 2). This outcome highlighted the embedding of the early Armenian termi-

nology within the larger late antique Christian cosmos. The closeness between *surb* and other early Christian terminologies of *holy* was also perceivable in its application in widespread Christian expressions, such as “holy spirit” (Arm. *surb hogi*) or “holy church” (Arm. *surb yekelec`i*).

Furthermore, *surb* was frequently employed as an adjective with respect to persons. In this context, it served as one among several other attributes in so-called attribute chains. Unlike most of the other attributes, however, *surb* was almost exclusively used in the context of central religious persons. The use of *surb* as a religious attribute for persons and the observation that the status of *surb* derived from god can be regarded as being consistent with other early Christian traditions (Angenendt 2007; Brown 2015; Gemeinhardt and Heyden 2012).

Despite these similarities, the close relationship between the Armenian term *surb*¹ and the semantic field of purity also revealed important differences between the notion of *surb* and other early Christian terminologies of *holy*, for instance, in Ancient Greek. One potential explanation for the use of *surb* for both *holy* (in the sense of Grk. *hágios*) and “pure” (in the sense of Grk. *kat`arós*) was seen in the Persian and more precisely Zoroastrian influence on the Armenian terminology (see Chapter 2). Furthermore, the close link between *holy* and pure in the Armenian terminology urges to re-evaluate the relationship between *holy* and pure in both historical and contemporary notions of *holy*.

The examination of the use and meaning of *holy* in contemporary English texts, which was based on the analysis of the HSEC and the *English Web corpus 2015* (*en-TenTen15*), yielded rather different results compared to the analysis of *surb*. The notion of *holy* in the English text corpora was related to individual persons only to a negligible extent. Instead, it was mostly applied in fixed expressions with either religious or non-religious meaning (such as “Holy Spirit” or “holy shit”). Especially in the non-religious subcorpora, the notion of *holy* was also frequently utilized as an attribute or synonym for abstract concepts such as “man,” “religion,” “world,” “time,” or “place.” Yet, the widespread but rather distinct application of *holy* in certain expressions of the everyday language also highlighted aspects of purity in the contemporary understanding of the terminology, although exclusively in a non-religious sense. Overall, the process of a sedimentation of the terminology in fixed expressions, which was partially already perceivable in the Armenian material (for example, the *surb hogi*), seems to have increased in the contemporary religious understanding and use of *holy*, whereas the Armenian use of *surb* more frequently illustrated its potential to serve as a distinct and precise attribute.

1 Particularly when *surb* is used as a verb, such as *srbel*.

7.2 Overlappings between *Surb*, *Holy*, and the Comparative Notion of *Holy*

The comparison of the semantic fields of *surb* and *holy* in Chapter 6.2.1 resulted in a differentiated picture. The detailed analyses and summaries of this comparison are provided in Chapters 6.1.3 and 6.1.4. According to the results of the comparison of the “domains” and “semantic fields” annotations, as well as the examination of the overlappings of words between the synoptic tables of the AAFTC/AASC and the HSEC, both semantic fields had only a few but important overlappings.

Regarding the “domain” and “semantic field” annotations, words belonging to either the “concept,” “person,” “verb,” or “title” domains were of crucial importance in the semantic field of *surb*. Primarily the ensemble of the “verb,” “title,” and “person” domains indicated the general significance of *surb* in the context of extraordinary religious persons who often held a prominent position in the Armenian Church hierarchy. The “concept” domain in the AAFTC/AASC consisted mainly of words related to “Religion and the Supernatural” (S9) or non-religious “Abstract Terms” (A1), such as “name.”

The “concept” domain was also important for the semantic field of *holy*, followed by the “object,” and “place” domains. A closer look at the “semantic field” annotation² revealed that the broad “concept” domain in the HSEC data particularly needs to be further differentiated. It mainly includes “Religion and the Supernatural” (S9), “Mental Actions and Processes” (X2), and “Linguistic Actions, States & Processes: Communication” (Q1) with abstract terms such as “meaning.”

In the case of *surb* and that of *holy*, the “concept” domain and the “semantic field” of “religion and the supernatural” (S9) were the most important layers of both semantic fields. These “domains” and “semantic fields” also form the primary links between *surb* and *holy*. This impression was further confirmed by the investigation of the merged word table that only includes very few and mostly religious words, such as “god,” “spirit,” and “church.” Thus, the fundamental and almost exclusive part of the intersection between the semantic fields of *surb* and *holy* is represented only by a few general religious terms. Nearly all the other “domains” and “semantic fields” that appear in the synoptic tables either considerably differ in relevance between the English and the Armenian corpora or are not even part of the intersection.

The overall result of the comparison thus revealed only a few but strong links between the terminologies, among them the “holy spirit” and “god.” Yet, the majority of the more specific use cases differed significantly between the corpora.

2 Again, note that the USAS Annotation System often did not work as expected. Some of the false attributions have manually been corrected. Others, however, might still be part of the data.

This tendency was strengthened by the results of the comparison between *holy* and *surb* and the comparative notion of *holy* in the study of religion. Similar to the previous results, this comparison yielded many differences and only two subfields from the comparative notion of *holy* that were equally important in both corpora, namely “Holy and Profane” and “Holy and Power.” Akin to the central position of the “holy spirit” and the semantic field of religion in the comparison of *surb* and *holy*, the “Holy and Profane” and the “Holy and Power” subfields function as strong links that tie the semantic fields of *surb* and *holy* together.

Above all, the *holy* and *profane* distinction is of central importance for the overall constitution of the comparative concept of *holy* (see Chapter 2).¹ Its presence in both semantic fields of *holy* and *surb* can justly be interpreted as a central argument for subsuming *holy* and *surb* under a more abstract semantic field of *holy*. Even though it is less fundamental than the “Holy and Profane” distinction, the idea that *holy* is a powerful concept or sometimes even a mighty agent³ also forms an important link between both semantic fields by underlining their common Christian background.

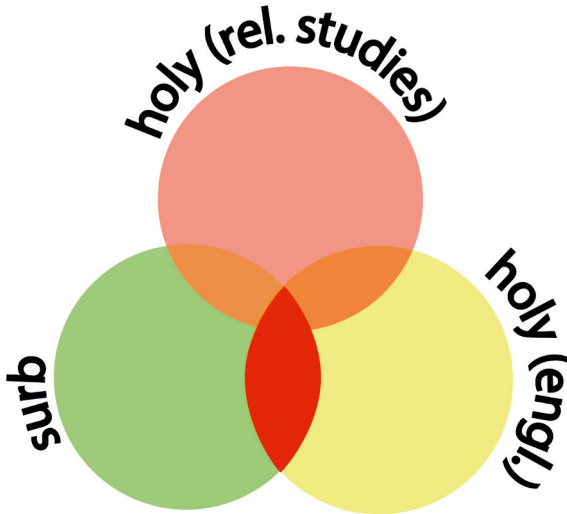
Despite the long list of differences between the use and meaning of *holy* and *surb*,⁴ the final evaluation of the individual analyses resulted in the conclusion that the semantic fields of *holy* and *surb* can rightfully be seen as interrelated. This was supported by the observation that both semantic fields also correspond to (different) subfields of the comparative notion of *holy* that was employed as a *tertium comparationis*. Yet, the interrelations between *surb* and *holy* do not comprise their specific use cases, which differ significantly between the historical and contemporary terminologies. Instead, they mostly consist of few but strong links that tie both fields together.

In conclusion, the links between *surb* and *holy* only consist of very few overlappings, at least from a quantitative perspective. Yet, from a qualitative perspective, these few interrelations are strong enough to justify the integration of both notions under a more abstract and broader notion of *holy*. This result is reflected in the visualization in figure 82. The red space in the visualization represents the relatively few overlappings between the semantic fields and their (approximate) quantitative relation to the overall semantic fields.

3 See Chapter 2 as well as the discussions regarding the “holy spirit” in the respective parts of Chapters 5 and 6.

4 To name but a few: Attribute chains (AAFTC/AASC) versus fixed expressions (HSEC), (individual) persons (AAFTC/AASC) versus abstract concepts (HSEC), mental states (feelings) appeared only in the HSEC, whereas the Armenian data showed a strong focus on purity, etc.

Figure 82: Visualization of the few but central overlappings between the three semantic fields from a quantitative perspective.



7.3 Contributions to the History of Religion and the Comparative Notion of *Holy*

The contribution of this book to the history of early Armenian Christianity lies in its extensive and thorough examination of the notion of *surb* in early Armenian texts. Other scholars have already singled out specific aspects of *surb* discussed in this book.⁵ Yet, the latter provides a detailed analysis and account of the different applications and meanings of *surb*. Furthermore, the transparent and strictly empirical methodology has not only confirmed existing hypotheses, for instance, regarding the close connection between *surb* and purity, but it also produced a more precise and differentiated picture of the application and meaning of *surb*.

The most important results of this book for the understanding of the early Armenian notion of *surb* and early Armenian Christianity in general are here recapitulated:

5 Even though briefly and mostly from a linguistic Indo-European perspective, see Stempel (1988); Morani (2001). Particularly Stempel (1988) has pointed to the close relationship between *surb* and purity, but mostly from a very general linguistic perspective that is especially interested in etymology and without any detailed examination of the Armenian texts.

- This book contains a thorough review of the close relationship between *surb* and various sub-genres of the semantic field of purity. These interrelations could range from mere physical purity to sexual, spiritual, ethical, and religious purity. The omnipresence of the connotation of pure in both *surb* and its antonym *pilc* was particularly observable in the comparison between the Greek and the Armenian versions of the Gospels. The aspect of purity was also present in the other Armenian sources, for example, in the episode about Gayianê and Hřip' simê in the AAFTC/AASC Agat' angelos and expressions such as the "holy kiss" in the AAFTC/AASC Koriwn. Yet, the question remains in how far other use cases of *surb*, which are intuitively assumed to be closely connected to a broader notion of *holy* (such as in Grk. *hágios*), should also be re-evaluated in light of their potential connection to purity.
- The central position of purity in the context of *surb* also raised the question of how far other ancient terminologies of *holy* were related to purity. At least the Greek and Latin texts already possessed a differentiated vocabulary for both *holy* and purity, which does not impede potential semantic overlappings between both spheres. Yet, the example of the Zoroastrian and Persian vocabulary in Chapter 2 with its strong focus on the notion of purity and the assumed absence of a notion of *holy* seemed to be comparable to the Armenian situation. Considering the close relationship between Armenia and the Persian realm (see also Chapter 2), one may legitimately ask in how far the terminology around *surb* was also influenced by notions of purity and *holy* from the Persian-Zoroastrian realm. Yet, since neither Stempel (1988) nor Hübschmann (1897, 492) ascertain any traceable links between *surb* and Persian terminologies, this assumption can only be valid when minimizing the importance of loan words between Persian and Armenian. However, I do not see any reason why a semantic coincidence between two concepts should only be expressed by terms that share the same etymological root or emerge as direct loan words from other languages. If this was the case, the idea of intercultural and inter-religious notions such as *holy* would need to be entirely abandoned. As Stempel convincingly states, most of the Indo-European terminologies related to *holy* are neither based on loan words nor do they share common etymological roots (Stempel 1988, 241).
- The last result of the analysis of the Armenian corpora stresses the embedding of early Armenian Christianity within the wider late antique Christian cosmos. The examples discussed and analyzed in this book were very much in accordance with what has been stated by scholars such as Peter Brown, Arnold Angenendt, and Christoph Auffarth about other early Christian cultures and their application of *holy*. The several similarities between the Armenian and different early Christian cultures, such as the focus on *holy* in the context of persons, as well as differences, such as the close relationship between *surb* and purity,

demonstrate that the case-study of the Armenian notion of *surb* is a pertinent addition to the overall discussion of *holy* in early Christian cultures in which the Armenian case remains at the periphery (Colpe 1988b; Gemeinhardt and Heyden 2012; Wokart 2018).

The second contribution of this book is related to the comparative notion of *holy* in the study of religion. In this regard, the examinations and discussions in this book have yielded the following observations and results:

- The comparative notion of *holy* employed as a *tertium comparationis* was applicable to both the Armenian notion of *surb* and to the very diverse use cases of the contemporary notion of *holy*, thereby demonstrating that both notions are interrelated. This capacity of the comparative notion of *holy* also demonstrated its general potential for comparative research in the study of religion. The comparative notion of *holy* was able to identify the unique applications of *surb* and *holy* as well as the few but strong interrelations between the two semantic fields. This differentiated result illustrates that the extensive research on a comparative notion of *holy* during the last two centuries is able to cope with a great variety of historical and contemporary examples of notions related to *holy*. Yet, this book has also demonstrated the necessity to combine the individual approaches that form the overall field of *holy* in the study of religion since none of the more specific approaches can efficiently cover the often diverse material alone. A case in point are the works by Eliade and Otto, which had a major influence on the modern understanding of *holy* in the study of religion and channel individual approaches towards the *holy*. Even though both corresponded very well to some modern intellectual or academic discussions on *holy*, they were hardly able to relate to the restricted and concrete use of *surb* in the Armenian texts.
- Despite the comparative notion of *holy*'s capacities in the comparative part, it still needs to be adjusted in some respects. A telling example is the focus of the Armenian terminology on purity. Even though the relationship between *holy* and purity is also briefly recognized in contemporary discussions on *holy*, it is neglected in relation to other aspects such as "Holy and Feelings" or "Holy as a Synonym for Religion." Therefore, I hope that the results of this book have demonstrated that an emphasis on purity with respect to the comparative notion of *holy* can also prove to be instructive for other settings. For instance, the knowledge of the importance of the interrelation between *surb* and purity in the Armenian material has also revealed such a connection within the non-religious use of *holy* in the Twitter data.

7.4 Critical Review of the Methodology

The methodological approach of this book consisted of a unique combination of tools and methods from the field of corpus linguistics (macroanalysis), which were complemented by the examination of single sentences and passages (microanalysis). The approach yielded meaningful, consistent, and transparent results that were generally in accordance with the already existing literature while also generating several new insights into the material.

The combination of macro- and microanalyses proved to be especially useful in the course of the examination. The macroanalyses yielded valuable statistical insights into the broader semantic fields and applications of *surb* and *holy*, thereby facilitating a comparison of the individual semantic fields. The advantage of methods and tools from the field of corpus linguistics lies in their potential to generalize large amounts of data that cannot be handled by an individual researcher. Furthermore, the results of these corpus analyses are strictly based on empirical observations and on transparent theoretical presuppositions (distributional semantics). In addition, the fact that the results of the statistical analyses are expressed with numbers makes them easier to compare with one another. Yet, the potential to evaluate word relations on the basis of numbers alone presents the risk of conceiving misleading assumptions by taking them as absolute representations. Needless to say, these numbers should never be regarded as representing “hard” scientific values. Instead, they serve as abstracted and approximated indications of the relations between words and their potential spectrum of meaning. The final evaluation of the relationship between words, however, should also consider qualitative analyses as a complementary tool.

To avoid the false impression of a scientification of research in humanities, the investigations of shorter sentences or passages in this book proved to be very helpful. This more established hermeneutical⁶ approach did indeed uncover several blind spots in the macroanalysis. The absence of purity in the macroanalyses of the Armenian data was certainly the most obvious one. Besides giving interesting new insights into the data, the microanalysis consequently also safeguarded against misleading purely data-driven research in the humanities.

The successful combination of quantitative and qualitative methods in this book should be regarded as one of its major achievements in addition to the contributions to the historical and systematic research in the study of religion. It will hopefully prove to be useful for future research as well.

6 Even though there was no explicit qualitative approach in this part of the book. Yet, future expansion of the research questions in this book should also consider approaching the examinations in the micro section in a more guided manner. For instance, by adopting principles and ideas from “objective hermeneutics,” see Wernet (2009); Wagner and Oevermann (2001).

Bibliography

- Adler, Hans. 1992. "Lexikologie, Lexematik." In *Historisches Wörterbuch der Philosophie*, edited by Joachim Ritter, Karlfried Gründer, Gottfried Gabriel, Eckart Scheerer, Stephan Meier-Oeser, Benedikt Haller, Oliver R. Scholz, and Kerstin Behnke. Basel: Schwabe Verlag.
- Agamben, Giorgio. 2015. *Homo Sacer: Die souveräne Macht und das nackte Leben*. 10. Auflage. Vol. 2068. Edition Suhrkamp. Frankfurt am Main: Suhrkamp.
- Agat'angeghos, and Robert W. Thomson. 1970. *The Teaching of Saint Gregory: An Early Armenian Catechism*. Vol. 3. Harvard Armenian Texts and Studies. Cambridge, Mass.: Harvard University Press.
- Agat'angelos, and Robert W. Thomson. 1976. *History of the Armenians*. 1st ed. Albany: State University of New York Press.
- Aleksidze, Nikoloz. 2018. "Caucasia: Albania, Armenia, and Georgia." In *A Companion to Religion in Late Antiquity*, edited by Josef Lössl and Nicholas Baker-Brian, 135–57. Blackwell Companions to the Ancient World. Hoboken, NJ: John Wiley & Sons Inc.
- Alles, Gregory D. 2010. "Rudolf Otto (1869-1937)." In *Klassiker der Religionswissenschaft*, edited by Axel Michaels, 198–211. München: Beck.
- Angenendt, Arnold. 2007. *Heilige und Reliquien: Die Geschichte ihres Kultes vom Frühen Christentum bis zur Gegenwart*. 2., überarb. Aufl., genehmigte Lizenzausg. Hamburg: Nikol.
- Archer, Dawn, Andrew Wilson, and Paul Rayson. 2002. "Introduction to the USAS Category System." http://ucrel.lancs.ac.uk/usas/usas_guide.pdf.
- Asad, Talal. 1997. *Genealogies of Religion: Discipline and Reasons of Power in Christianity and Islam*. 2. ed. A Johns Hopkins Paperback. Baltimore, Md.[u.a.]: Johns Hopkins Univ. Press.
- . 2003. *Formations of the Secular: Christianity, Islam, Modernity*. Cultural Memory in the Present. Stanford: Stanford University Press.
- Beck, Andrea, and Andreas Berndt. 2013. *Sakralität und Sakralisierung: Perspektiven des Heiligen*. 1. Aufl. Stuttgart: Franz Steiner Verlag.

- Beck, Andrea, Klaus Herbers, and Andreas Nehring, eds. 2017. *Heilige und geheiligte Dinge: Formen und Funktionen*. Vol. Band 20. Beiträge Zur Hagiographie. Stuttgart: Franz Steiner Verlag.
- Berner, Ulrich. 2010. "Mircea Eliade." In *Klassiker der Religionswissenschaft*, edited by Axel Michaels, 343–56. München: Beck.
- Biber, Douglas, Randi Reppen, Douglas Biber, and Randi Reppen, eds. 2015. *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139764377>.
- Boyce, Mary. 2011. "AMŠa SPNTA." In *Encyclopaedia Iranica*. <http://www.iranicaonline.org/articles/amesa-spenta-beneficent-divinity>.
- Bräunlein, Peter J. 2017. "Was ist uns heilig? Kulturwissenschaftliche Anmerkungen zu sakralen Dingen." In *Heilige und geheiligte Dinge*, edited by Andrea Beck, Klaus Herbers, and Andreas Nehring, 20:9–28. Beiträge zur Hagiographie. Stuttgart: Franz Steiner Verlag.
- Brennecke, Hanns Christof. 1998. "Chalkedonense und Henotikon." In *Chalkedon*, edited by Johannes Roldanus and Johannes van Oort, 24–53. Studien der Patristischen Arbeitsgemeinschaft (SPA). Louvain: Peeters.
- Brezina, Vaclav. 2018. *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge; New York, NY: Cambridge University Press; Cambridge University Press.
- Brezina, Vaclav, Tony McEnery, and Stephen Wattam. 2015. "Collocations in Context: A New Perspective on Collocation Networks." *International Journal of Corpus Linguistics* 20 (2): 139–73.
- Brown, Peter. 1986. *Die letzten Heiden: Eine kleine Geschichte der Spätantike*. Berlin: Klaus Wagenbach.
- . 2015. *The Cult of the Saints: Its Rise and Function in Latin Christianity*. Enlarged ed., 2. ed. Chicago: The University of Chicago Press.
- Burkert, Walter, Christel Matthias Schröder, and Peter Antes. 2011. *Griechische Religion der Archaischen und Klassischen Epoche*. 2., überarb. und erw. Aufl. Vol. 15. Die Religionen der Menschheit. Stuttgart: Kohlhammer.
- Calzolari, Valentina. 2014, cop. 2014. "Écriture et mémoire religieuse dans l'Arménie ancienne (Ve siècle ap. J.-C.)." In *Le savoir des religions*, edited by Daniel Barbu, Philippe Borgeaud, and Mélanie Lozat, 373–92. Gollion: Infolio editions.
- . 1997. "Réécriture des textes apocryphes en Arménien: L'exemple de la légende de l'apostolat de Thaddée en Arménie." *Apocrypha*, no. 8: 97–110.
- Canal, Héctor, Caroline Sauter, Hans-Joachim Schott, and Maik Neumann. 2013. *Das Heilige (in) der Moderne*. Transcript Verlag.
- Chaumont, Marie-Louise. 1969. *Recherches sur l'histoire d'Arménie de l'avènement des Sassanides à la conversion du royaume*. Paris.
- Clark, Stephen. 2015. "Vector Space Models of Lexical Meaning." In *The Handbook of Contemporary Semantic Theory*, edited by Chris Fox and Shalom Lappin, 493–523. Blackwell Handbooks in Linguistics. Hoboken: John Wiley & Sons.

- Colpe, Carsten, ed. 1977. *Die Diskussion um das Heilige*. Vol. Bd. 305. Wege der Forschung. Darmstadt: Wissenschaftliche Buchgesellschaft [Abt. Verl.].
- . 1988a. "Das Heilige." In *Handbuch religionswissenschaftlicher Grundbegriffe*, 3:80–99.
- . 1988b. "Heilig (Sprachlich)." In *Handbuch religionswissenschaftlicher Grundbegriffe*, 3:74–80.
- . 1990. *Über das Heilige: Versuch, seiner Verkennung kritisch vorzubeugen*. Anton Hain. Frankfurt am Main: A. Hain.
- Cowe, S. Peter. 2012. "The Armenian Version of the New Testament." In *The Text of the New Testament in Contemporary Research*, edited by Bart D. Ehrman and Michael W. Holmes, 253–92. Brill.
- Dhawan, Nikita, and María do Mar Castro Varela. 2015. *Postkoloniale Theorie*. Transcript Verlag.
- Dihle, Albrecht. 1988. "Heilig." In *Reallexikon für Antike und Christentum*, 14:1–63.
- Douglas, Mary. 2001. *Purity and Danger: An Analysis of the Concepts of Pollution and Taboo*. London: Routledge.
- DuCharme, Bob. 2013. *Learning SPARQL: Querying and Updating with SPARQL 1.1*. Second edition. Beijing: O'Reilly.
- Durkheim, Émile. 2017. *Die elementaren Formen des religiösen Lebens*. 4. Auflage. Vol. 2. Taschenbuch / Verlag der Weltreligionen. Berlin: Verlag der Weltreligionen.
- Eliade, Mircea, and Willard R. Trask. 1959. *The Sacred and the Profane: The Nature of Religion*. Harcourt.
- Firth, J. R. 1957. "A Synopsis of Linguistic Theory 1930-1955." In *Studies in Linguistic Analysis*, 1–32. Oxford.
- Frawley, William, and William Bright. 2003. *International Encyclopedia of Linguistics*. 2nd ed. New York: Oxford University Press.
- Freiberger, Oliver. 2019. *Considering Comparison: A Method for Religious Studies*.
- Frevel, Christian, and Christophe Nihan, eds. 2013. *Purity and the Forming of Religious Traditions in the Ancient Mediterranean World and Ancient Judaism*. Vol. 3. Dynamics in the History of Religion. Leiden; Boston: Brill.
- Fuchs-Heinritz, Werner. 2013. *Lexikon zur Soziologie*. 5., überarb. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gantke, Wolfgang und Serikov, Vladislav, ed. 2015. *Das Heilige als Problem der gegenwärtigen Religionswissenschaft*. Peter Lang Pub Inc.
- Garsoïan, Nina. 2000. "MOVSĒS XORENACI." In *Encyclopaedia Iranica*. <http://www.iranicaonline.org/articles/movses-xorenaci>.
- Garsoïan, Nina G. 1999. *L'Église Arménienne et le Grand Schisme d'orient*. Vol. t. 100. Corpus Scriptorum Christianorum Orientalium Subsidia. Lovanii: Éditions Peeters.
- . 2004a. "Janus: The Formation of the Armenian Church from the IVth to the VIIth Century." In *The Formation of a Millennial Tradition*, edited by Garegin,

- Robert F. Taft, and Gabriele Winkler, 79–95. *Orientalia Christiana Analecta*. Roma: Pontificio istituto orientale.
- . 2004b. “L’histoire attribuée à Movsès Xorenac’i: Que reste-t-il à en dire?” *Revue des Etudes Arméniennes* 2003-2004 (29): 29–48.
- , ed. 2010. *Studies on the Formation of Christian Armenia*. Vol. CS959. Variorum Collected Studies Series. Farnham: Ashgate/Variorum.
- Garsoïan, Nina G., and Jean-Pierre Mahé. 1997. *Des Parthes au Califat: Quatre leçons sur la formation de l’identité arménienne*. Vol. 10. Travaux et mémoires du Centre de Recherche d’Histoire et Civilisation de Byzance. Paris: De Boccard.
- Garsoïan, Nina G., Thomas F. Mathews, and Robert W. Thomson. 1982. *East of Byzantium: Syria and Armenia in the Formative Period*. Washington D.C.: Dumbarton Oaks Center for Byzantine Studies Trustees for Harvard University.
- Gemeinhardt, Peter, and Katharina Heyden, eds. 2012. *Heilige, Heiliges und Heiligkeit in spätantiken Religionskulturen*. Vol. Bd. 61. Religionsgeschichtliche Versuche und Vorarbeiten. Berlin; Boston: De Gruyter.
- Géron, Aurélien. 2019. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. Second edition. Sebastopol, CA: O’Reilly Media, Inc.
- Gries, Stefan. 2015. “Quantitative Designs and Statistical Techniques.” In *The Cambridge Handbook of English Corpus Linguistics*, edited by Douglas Biber, Randi Reppen, Douglas Biber, and Randi Reppen, 50–72. Cambridge: Cambridge University Press.
- Harris, Zellig S. 2015. “Distributional Structure.” *WORD* 10 (2-3): 146–62.
- Hutter, Manfred. 2013. “Concepts of Purity in Anatolian Religion.” In *Purity and the Forming of Religious Traditions in the Ancient Mediterranean World and Ancient Judaism*, edited by Christian Frevel and Christophe Nihan, 159–75. Dynamics in the History of Religion. Leiden; Boston: Brill.
- Hübschmann, Heinrich. 1897. *Armenische Grammatik: Armenische Etymologie*. 2nd ed. Leipzig.
- Illert, Martin. 2007. *Doctrina Addai, de Imagine Edessena: [Griechisch, Lateinisch, Deutsch]*. Vol. Bd. 45. Fontes Christiani. Turnhout: Brepols.
- Jannidis, Fotis, Hubertus Kohle, and Malte Rehbein. 2017. *Digital Humanities: Eine Einführung*. Stuttgart: J.B. Metzler.
- Jong, Albert de. 2015. “Armenian and Georgian Zoroastrianism.” In *The Wiley Blackwell Companion to Zoroastrianism*, edited by Michael Stausberg and Yuhan Sohrab-Dinshaw Vevaina, 119–28. The Wiley Blackwell Companions to Religion. Hoboken: John Wiley & Sons Inc.
- Kamper, Dietmar, ed. 1997. *Das Heilige: Seine Spur in der Moderne*. Unveränd. Nachdr. Bodenheim: Syndikat-Buchges.
- Karrer, M., Williams, R.D., Hauschild, W., a.o. 2019. “Christologie.” In *Religion in Geschichte Und Gegenwart*.

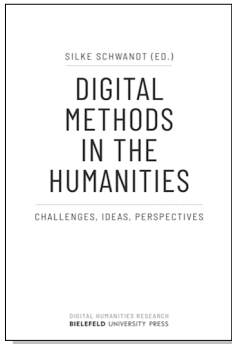
- Kouymjian, Dickran. 2015. "Introduction. 3.2. Armenian Manuscripts." In *Comparative Oriental Manuscript Studies*, edited by Alessandro Bausi and Eugenia Sokolinski, 38–40. Hamburg: Tredition.
- Lanczkowski, Günter. 1980. "Heiligkeit." In *Theologische Realenzyklopädie*, edited by Gerhard Krause, Siegfried M. Schwertner, and Gerhard Müller, 14:695–712. Berlin [u.a.]: De Gruyter.
- Luhrmann, T. M. 1996. *The Good Parsi: The Fate of a Colonial Elite in a Postcolonial Society*. Cambridge, Mass.: Harvard University Press.
- Lutzky, Harriet. 1993. "On a Concept Underlying Indo-European Terms for the Sacred." *The Journal of Indo-European Studies* 21: 283–301.
- Mahé, Annie, and Jean-Pierre Mahé. 2012. *Histoire de l'Arménie, Des Origines à Nos Jours*. Pour l'histoire.
- Malinar, Angelika, ed. 2009. *Un-Reinheit im Kulturvergleich*. Paderborn; München: Fink.
- McEnery, Tony, and Andrew Hardie. 2012. *Corpus Linguistics: Method, Theory and Practice*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Michaels, Axel, and Michael Bergunder. 2019. "Phenomenology of Religion." In *Religion in Geschichte Und Gegenwart*.
- Morani, Moreno. 2001. "La Terminologia del sacro in lingue indoeuropee antiche: riflessioni e problemi." In *Pensiero e Istituzioni del Mondo Classico nelle Culture del Vicino Oriente*, edited by R. B. Finazzi and Alfredo Valvo, 165–96. L'Eredità Classica nel Mondo Orientale. Alessandria: Edizioni dell'Orso.
- Oeldemann, Johannes. 2011. *Die Kirchen des christlichen Ostens: Orthodoxe, orientalische und mit Rom unierte Ostkirchen*. 3. Aufl. Vol. 577. Topos-Taschenbücher. Kevelaer: Verl.-Gemeinschaft Topos Plus.
- Otto, Rudolf. 2004 [1917]. *Das Heilige: Über das Irrationale in der Idee des Göttlichen und sein Verhältnis zum Rationalen*. Nachdr. Beck'sche Reihe. München: Beck.
- Otto, Rudolf, and John W. Harvey. 1923. *The Idea of the Holy*. Oxford: Oxford University Press.
- Paden, William E. 2007. "Sacred and Profane." In *Religion Past & Present*, edited by Hans Dieter Betz. Leiden; Boston: Brill.
- Parker, Robert. 2003. *Miasma: Pollution and Purification in Early Greek Religion*. Reprinted. Oxford: Clarendon Press.
- Perikhanian, Anahit. 1986. "Sur Arménien Buzand." In *Armenian Studies*, edited by Dickran Kouymjian, 653–57. Lisbon.
- Pulman, Stephen. 2012. "Distributional Semantic Models." In *Compositional Methods in Physics and Linguistics*, edited by Sadrzadeh Heunen Grefenstette. Oxford.
- Pylajew, Maxim, and Konstantin Antonow. 2015. "Die Kategorie des Heiligen in der Religionsphänomenologie, Theologie und Philosophie des 20. Jahrhunderts." *Neue Zeitschrift für Systematische Theologie und Religionsphilosophie* 57 (1).

- P'awstos, and Nina G. Garsoïan. 1989. *The Epic Histories Attributed to P'awstos Buzand (Buzandaran Patmut'iwnk')*. Vol. 8. Harvard Armenian Texts and Studies. Cambridge Mass.: Distributed for the Dept. of Near Eastern Languages and Civilizations Harvard University by Harvard University Press.
- Rayson, Paul. 2015. "Computational Tools and Methods for Corpus Compilation and Analysis." In *The Cambridge Handbook of English Corpus Linguistics*, edited by Douglas Biber, Randi Reppen, Douglas Biber, and Randi Reppen, 32–49. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139764377.003>.
- Robertson, Noel. 2013. "The Concept of Purity in Greek Sacred Laws." In *Purity and the Forming of Religious Traditions in the Ancient Mediterranean World and Ancient Judaism*, edited by Christian Frevel and Christophe Nihan, 195–245. Dynamics in the History of Religion. Leiden; Boston: Brill.
- Russell, James R. 1987. *Zoroastrianism in Armenia*. Vol. 5. Harvard Iranian Series. Cambridge MA: Harvard University Dept. of Near Eastern Languages and Civilizations; National Association for Armenian Studies and Research.
- . 1990. "Pre-Christian Armenian Religion." *Aufstieg Und Niedergang Der römischen Welt*, 2679–92.
- Russell, James R. 2012. "Faustus." In *Encyclopaedia Iranica*, IX:449–51. <http://www.iranicaonline.org/articles/faustus->
- Sahle, Patrick. 2015. "Digital Humanities? Gibt's doch gar nicht!" HAB - Herzog August Bibliothek.
- Sakr, Sherif, Marcin Wylot, Raghava Mutharaju, Danh Le Phuoc, and Irini Fundulaki. 2018. *Linked Data: Storing, Querying, and Reasoning*. Cham: Springer.
- Sarkissian, Karekin. 1965. *The Council of Chalcedon and the Armenian Church*. London.
- Schalk, Peter. 2013. *Religion in Asien? Studien zur Anwendbarkeit des Religionsbegriffs*. Vol. 32. Acta Universitatis Upsaliensis: Historia Religionum. Uppsala: Univ.
- Schleiermacher, Friedrich. 1980 [1806]. *Über die Religion: Reden an die Gebildeten unter ihren Verächtern*. Stuttgart: Reclam.
- Schlette, Magnus, and Volkhard Krech. 2017. "Sakralisierung." In *Handbuch Religionssoziologie*, edited by Detlef Pollack, Volkhard Krech, Olaf Müller, and Markus Hero, 437–63. Veröffentlichungen der Sektion Religionssoziologie der Deutschen Gesellschaft für Soziologie. Springer Science and Business Media; Springer VS.
- Schmidt, Werner H., Peter Schäfer, and Klaus Berger. 1980. "Geist/Heiliger Geist/Geistesgaben." In *Theologische Realenzyklopädie*, edited by Gerhard Krause, Siegfried M. Schwertner, and Gerhard Müller, 12:170–254. Berlin [u.a.]: De Gruyter.
- Schröder, Stefan. 2012. "Darstellungen der Religionsphänomenologie in der deutschen religionswissenschaftlichen Einführungsliteratur: Ein Vergleich." *Zeitschrift für Junge Religionswissenschaft*, no. 7: 20–39.

- Schütze, Hinrich. 1998. "Automatic Word Sense Discrimination." *Computational Linguistics*, no. 24 (1): 97–123.
- Shapiro, Susan O. 2013. "Pasolini's Medea: A Twentieth-Century Tragedy." In *Ancient Greek Women in Film*, edited by Konstantinos P. Nikoloutsos, 95–116. Classical Presences. Oxford: Oxford Univ. Press.
- Söderblom, Nathan. 1913. "Holiness." In *Encyclopaedia of Religion and Ethics*, edited by James Hastings, 6:731–59.
- Stempel, R. 1988. "Armenisch Surb Heilig, Rein: Erbwort oder Lehnwort?" In *Nubia Et Oriens Christianus*, edited by Caspar Detlef Gustav Müller, Piotr O. Scholz, and Reinhard Stempel, 239–42. Bibliotheca Nubica. Köln: J. Dinter.
- Streng, Frederick J. 2018. "Sacred." In *Encyclopædia Britannica*. <https://www.britannica.com/topic/sacred>.
- The Oxford Companion to English Language*. 2017. Oxford University Press.
- Thewrewk, Emil, and Sextus Pompeius Festus. 1889. *De verborum significatu quae supersunt cum Pauli Epitome*. Budapest.
- Thomson, Robert W. 1994. *Studies in Armenian Literature and Christianity*. Vol. CS451. Collected Studies Series. Aldershot, Hampshire, Great Britain; Brookfield, Vt., USA: Variorum.
- . 1995. *A Bibliography of Classical Armenian Literature to 1500 AD*. Corpus Christianorum. Turnhout: Brepols.
- . 2006. *History of the Armenians*. Rev. ed. Ann Arbor: Caravan Books.
- Trenter, Kate. 2013. "Corpus Linguistics of Historical Texts." In *The Encyclopedia of Applied Linguistics*, edited by Carol Chapelle. Malden, Mass.; Hoboken, USA: Wiley-Blackwell; John Wiley.
- Tyrell, Hartmann. 2018. "Das Religioide und der Glaube." In *Georg Simmel und das Leben in der Gegenwart*, edited by Rüdiger Lautmann and Hanns Wienold, 347–62. Wiesbaden: Springer VS, Springer Fachmedien Wiesbaden.
- Underwood, Ted. 2017. "A Genealogy of Distant Reading." *Digital Humanities Quarterly* 011 (2).
- van Esbroeck, Michel. 1983. "La naissance du culte de Saint Barthélémy en Arménie." *Revue des Études Arméniennes*, 171–95.
- Wagner, Hans-Josef, and Ulrich Oevermann. 2001. *Objektive Hermeneutik und Bildung des Subjekts: Mit einem Text von Ulrich Oevermann: Die Philosophie von Charles Sanders Peirce als Philosophie der Krise*. Weilerswist: Velbrück Wissenschaft.
- Wernet, Andreas. 2009. *Einführung in die Interpretationstechnik der Objektiven Hermeneutik*. 3. Aufl. Qualitative Sozialforschung. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Williams, Alan V. 2015. "Purity and Pollution / the Body." In *The Wiley Blackwell Companion to Zoroastrianism*, edited by Michael Stausberg and Yuhan Sohrab-Dinshaw Vevaina, 345–61. The Wiley Blackwell Companions to Religion. Hoboken: John Wiley & Sons Inc.

- Winkler, Gabriele, and Vardapet Koriun. 1994. *Koriwns Biographie des Mesrop Maštoc': Übersetzung und Kommentar*. Vol. 245. *Orientalia Christiana Analecta*. Roma: Pontificio istituto orientale.
- Wittgenstein, Ludwig, and Joachim Schulte. 2003. *Philosophische Untersuchungen*. 1. Aufl. Vol. Bd. 1372. Bibliothek Suhrkamp. Frankfurt am Main: Suhrkamp.
- Wokart, Norbert. 2018. *Heilig, Heiligkeit*. *Historisches Wörterbuch der Philosophie*. Basel: Schwabe Verlag.

Bielefeld University Press



Silke Schwandt (ed.)

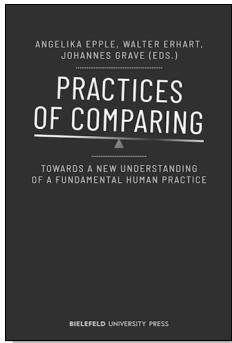
Digital Methods in the Humanities Challenges, Ideas, Perspectives

2020, 312 p., pb., col. ill.

38,00 € (DE), 978-3-8376-5419-6

E-Book: available as free open access publication

PDF: ISBN 978-3-8394-5419-0



Angelika Epple, Walter Erhart, Johannes Grave (eds.)

Practices of Comparing Towards a New Understanding of a Fundamental Human Practice

2020, 406 p., pb., col. ill.

39,00 € (DE), 978-3-8376-5166-9

E-Book: available as free open access publication

PDF: ISBN 978-3-8394-5166-3



Haun Saussy

Are We Comparing Yet? On Standards, Justice, and Incomparability

2019, 112 p., pb.

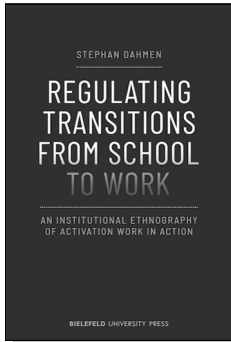
19,99 € (DE), 978-3-8376-4977-2

E-Book: available as free open access publication

PDF: ISBN 978-3-8394-4977-6

**All print, e-book and open access versions of the titles in our list
are available in the online shop www.bielefeld-university-press.de**

Bielefeld University Press



Stephan Dahmen

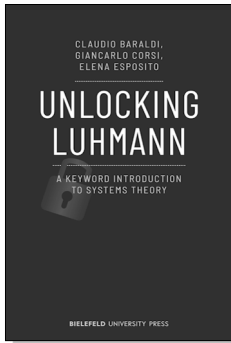
Regulating Transitions from School to Work An Institutional Ethnography of Activation Work in Action

June 2021, 312 p., pb., ill.

36,00 € (DE), 978-3-8376-5706-7

E-Book: available as free open access publication

PDF: ISBN 978-3-8394-5706-1



Claudio Baraldi, Giancarlo Corsi, Elena Esposito

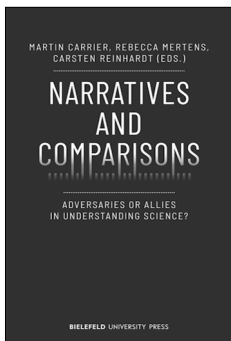
Unlocking Luhmann A Keyword Introduction to Systems Theory

April 2021, 276 p., pb.

40,00 € (DE), 978-3-8376-5674-9

E-Book: available as free open access publication

PDF: ISBN 978-3-8394-5674-3



Martin Carrier, Rebecca Mertens, Carsten Reinhardt (eds.)

Narratives and Comparisons Adversaries or Allies in Understanding Science?

January 2021, 206 p., pb., col. ill.

35,00 € (DE), 978-3-8376-5415-8

E-Book: available as free open access publication

PDF: ISBN 978-3-8394-5415-2

**All print, e-book and open access versions of the titles in our list
are available in the online shop www.bielefeld-university-press.de**