

Anhang

Zum Artikel „Korpusanalyse“ auf

<https://www.bibelwissenschaft.de/wirelex/wirelex/>

Stefan Altmeyer

(erstellt: Sept. 2018)

1. Allgemeine Korpusdaten

2. Typische Wörter

2.1. Worthäufigkeiten

2.2. Häufigste Namen im Literaturverzeichnis

2.3. Schlüsselwörter

2.4. Kontrastive Schlüsselwörter

2.5. Zweistufige Schlüsselwortanalyse

3. Wortverbindungen

3.1. Kollokationen

3.2. Konkordanz

Zur Veranschaulichung verschiedener Verfahren der Korpusanalyse wurde unter dem Namen „WiReLex2018“ ein Korpus aller im August 2018 online verfügbaren 368 Artikel des WiReLex erstellt und exemplarisch untersucht. Hier werden die Daten dokumentiert, auf die sich die im Artikel genannten Beispiele beziehen.

1. Allgemeine Korpusdaten

Diese Tabelle zeigt die allgemeine Statistik des Korpus, insbesondere Korpusgröße und Verteilung der verschiedenen Texteinheiten (Text, Satz, Wort).

tokens (running words)	1.318.526
types (distinct words)	91.218
standardised type/token ratio (STTR)	51,25
STTR std.dev.	48,81
STTR basis	1.000
mean word length (in characters)	6,61
word length std.dev.	4,38
sentences	88.587
mean (in words)	14,89
std.dev.	13,08
texts	368
mean (in words)	3.582,95
std.dev.	1.146,88

2. Typische Wörter

2.1. Worthäufigkeiten

Die Liste zeigt die häufigsten Wörter im Korpus und deren Verteilung. Zwei Ausschnitte sind beispielhaft herausgegriffen: die am häufigsten gebrauchten sowie eine Vergleichsgruppe deutlich seltener gebrauchter Nomen.

N	Word	Freq.	Texts	%	Dispersion
1	und	42499	368	100	0,98
2	der	39024	368	100	0,98
3	die	34172	368	100	0,98
4	in	20425	368	100	0,98
5	zu	13398	368	100	0,98
6	von	13352	368	100	0,98
7	den	11439	368	100	0,98
8	des	10253	368	100	0,98
9	ist	9398	368	100	0,97
...					
44	Religionsunterricht	2203	281	76,36	0,89
49	Menschen	2077	307	83,42	0,93
55	Bildung	1965	239	64,95	0,84
58	Gott	1750	221	60,05	0,85
59	Religion	1695	258	70,11	0,74

68	Schülerinnen	1456	242	65,76	0,92
71	Kirche	1396	217	58,97	0,93
73	Schüler	1314	237	64,40	0,90
76	Kinder	1266	236	64,13	0,92
...					
657	Gruppen	184	86	23,37	0,83
659	Kraft	184	105	28,53	0,87
662	Erwachsenenbildung	183	64	17,39	0,64
666	Christi	182	84	22,83	0,84
668	Voraussetzungen	182	104	28,26	0,89
672	Reformation	181	64	17,39	0,69
673	Sünde	181	33	8,97	0,57
674	Unterscheidung	181	108	29,35	0,90
...					

2.2. Häufigste Namen im Literaturverzeichnis

Die Liste bezieht sich auf die Literaturverzeichnisse aller Artikel und zeigt die zehn am häufigsten genannten Namen. Sie ist nicht nach absoluten Nennungen (Freq.) sortiert, sondern nach Anzahl der Texte, in denen die Namen genannt werden.

N	Word	Freq.	Texts	%	Dispersion
1	Schweitzer	236	106	28,8	0,89
2	Adam	153	86	23,37	0,87
3	Lachmann	146	81	22,01	0,83
4	Englert	133	77	20,92	0,89
5	Zimmermann	249	73	19,84	0,81
6	Schröder	154	70	19,02	0,65
7	Ziebertz	176	69	18,75	0,86
8	Büttner	180	69	18,75	0,86
9	Rothgangel	138	65	17,66	0,74
10	Leimgruber	96	61	16,58	0,84

2.3. Schlüsselwörter

Die folgenden Listen zeigen jeweils eine Auswahl aus den ersten 200 Schlüsselwörtern. Dazu wurden inhaltliche Klassen gebildet. Einige Einträge wurden gelöscht. Insgesamt wurden die Schlüsselwörter nach folgenden Kriterien berechnet:

Korpus: Alle Texte aus WiReLex2018 ohne Literaturverzeichnisse

Referenzkorpus: Wikipedia deutsch, Stand 11/2008

Min. Freq.: 20

Min. % of texts: 5%

P max.: $p \leq 10^{-7}$

Auswahl: Konzepte

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
163	Elementarisierung	93	41	0	961,4	135,4
164	Jugendtheologie	92	28	0	951,1	135,4
175	Kindertheologie	98	35	16	920,8	10,1
185	Korrelation	157	36	543	884,1	5,7

Auswahl: Grundbegriffe

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
6	Bildung	1965	239	14420	8459,2	4,6
7	Religion	1695	258	9612	8077,0	5,0
8	Lernen	1172	209	2060	7906,2	6,6
20	Lernens	495	165	405	3883,2	7,7
27	Religionen	626	141	2718	3278,0	5,3
40	Kompetenzen	478	120	1794	2624,3	5,5
41	Praxis	813	225	11994	2485,5	3,6
49	Handeln	514	180	3583	2260,0	4,7
50	Erfahrungen	676	225	8507	2257,3	3,8
52	Religiosität	299	87	383	2160,3	7,1
56	Reflexion	382	163	1386	2119,5	5,6
59	Lernprozesse	235	105	123	1970,2	8,4
64	Gerechtigkeit	403	95	2563	1838,1	4,8
68	Zugänge	279	111	607	1787,3	6,3
70	Erziehung	468	126	4706	1750,4	4,1
78	Bildungsprozesse	170	71	21	1625,4	10,5
85	Verständnis	466	213	6131	1519,2	3,7
86	Sinne	753	251	20493	1509,5	2,7
93	Kompetenz	299	111	1700	1423,4	4,9
96	Sozialisation	205	73	328	1412,8	6,8
104	Dialog	322	123	2568	1338,2	4,5
105	Identität	406	133	5257	1335,1	3,8
106	Hermeneutik	201	61	378	1334,6	6,5
108	Freiheit	480	135	8412	1322,2	3,3
109	Pluralität	161	85	104	1310,6	8,1
119	Tradition	556	209	13924	1191,6	2,8
120	Erfahrung	410	163	6638	1186,3	3,4

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
122	Kompetenzorientierung	113	47	0	1168,2	135,7
123	Religionsgemeinschaften	197	48	576	1165,6	5,9
129	Methoden	437	133	8606	1115,3	3,2
133	Wirklichkeit	344	145	4653	1104,5	3,7
136	Erzählungen	321	71	3857	1098,4	3,9
150	Anforderungssituationen	99	22	0	1023,5	135,5
152	Erschließung	131	81	110	1023,3	7,7
153	Bildungsprozessen	98	57	0	1013,1	135,5
155	Pluralisierung	109	65	22	1008,5	9,8
156	Dimensionen	233	119	1699	1005,7	4,6
158	Verstehen	176	81	598	996,4	5,7
160	Lernprozessen	113	66	43	985,0	8,8
161	Subjekt	215	92	1361	982,4	4,8
174	Orientierung	261	136	2948	921,9	4,0
179	Subjekte	136	72	254	904,5	6,6
181	Vielfalt	262	134	3170	893,2	3,9
183	Begegnung	257	116	3033	887,6	3,9
196	Bildungsstandards	102	30	66	830,1	8,1
198	Inhalte	334	153	6877	826,8	3,1

Auswahl: Inhalte

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
9	Gott	1750	221	13279	7428,8	4,5
15	Gottes	1138	193	7196	5202,0	4,8
39	Jesus	683	128	5938	2732,7	4,3
44	Jesu	558	119	4214	2373,5	4,5
46	Bibel	610	126	5857	2332,0	4,2
58	Glaubens	436	171	2532	2059,2	4,9
80	Glauben	476	164	6099	1574,2	3,8
82	Glaube	341	145	2186	1550,4	4,8
87	Christus	443	125	5494	1490,8	3,8
95	Ethik	319	86	2162	1418,7	4,7
97	Auferstehung	255	55	961	1398,2	5,5
100	Kirchengeschichte	248	51	902	1375,0	5,6
126	Paulus	270	53	2099	1134,4	4,5
132	Mt	197	50	690	1105,0	5,6
134	Kirche	1396	217	85615	1101,7	1,5
141	Texte	482	128	11352	1082,8	2,9
142	Segen	201	24	819	1074,8	5,4
146	Kirchen	501	151	12993	1044,5	2,8
178	Sünde	181	33	900	904,6	5,1
180	Schöpfung	206	72	1481	894,4	4,6

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
184	Bergpredigt	121	19	145	886,0	7,2
194	Katechismus	150	31	527	840,6	5,6
195	Christologie	113	26	131	832,8	7,2

Auswahl: Lernorte

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
3	Religionsunterricht	2203	281	537	20069,1	9,5
14	Religionsunterrichts	573	149	58	5536,9	10,8
35	Unterricht	726	191	6338	2899,4	4,3
57	Katechese	235	60	75	2087,2	9,1
69	Unterrichts	278	97	605	1780,8	6,3
79	Sekundarstufe	249	93	567	1576,7	6,3
89	Schule	980	227	37034	1449,3	2,2
130	Schulen	549	121	14845	1106,4	2,7
139	Gemeindepädagogik	105	28	0	1085,5	135,6
144	Erwachsenenbildung	183	64	568	1064,3	5,8
199	Jugendarbeit	192	44	1417	824,6	4,6
200	Lernorte	91	41	24	823,2	9,4

2.4. Kontrastive Schlüsselwörter

Für diese Berechnung wurde das Korpus in zwei Teilkorpora geteilt und eine Schlüsselwortanalyse durchgeführt, indem das eine mit dem anderen verglichen wurde. Einzelne Einträge wurden gelöscht.

Korpus: Artikel, in denen der Name Schweitzer im Literaturverzeichnis vorkommt

Referenzkorpus: Rest

Min. Freq.: 20

Min. % of texts: 5%

P max.: $p \leq 10^{-7}$

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
1	KIRCHENGESCHICHTE	379	53	33	129,2	2,4
2	PAULUS	294	45	21	113,5	2,7
3	MUSIK	270	45	17	111,6	2,8
4	MEDIEN	415	91	51	106,1	1,9
5	GEIST	254	58	16	104,9	2,8
6	SEGEN	222	21	11	102,2	3,2

N	Key word	Freq.	Texts	RC. Freq.	Log_L	Log_R
7	FRAUEN	302	64	27	101,1	2,3
8	SÜNDE	202	29	12	85,9	2,9
9	SEELE	185	27	9	85,8	3,2
10	FILM	163	45	8	75,4	3,2
11	MUTTER	118	28	7	50,2	2,9
12	DIAKONIE	109	18	6	48,0	3,0
13	SCHULD	119	42	8	47,5	2,7
15	JUGENDARBEIT	206	38	30	43,6	1,6
16	ERLÖSUNG	139	25	15	40,1	2,1
17	MITTELALTER	111	41	9	39,7	2,5
18	FRAU	101	43	7	39,7	2,7
19	LITURGIE	137	34	15	39,1	2,0
20	INTERNET	108	37	9	37,9	2,4
21	GEBET	159	46	21	37,7	1,8
23	KIRCHENGESCHICHTLICHE	91	30	7	33,6	2,6
24	GEMEINDEN	161	62	24	33,1	1,6
25	NÄCHSTENLIEBE	113	27	12	33,1	2,1
26	RITUALE	113	30	12	33,1	2,1
27	DTN	74	22	4	32,9	3,1
28	MÄDCHEN	155	33	23	32,1	1,6
29	KIRCHENGESCHICHTSDIDAKTISCH	76	31	5	30,7	2,8
32	PAPST	122	35	16	29,2	1,8

2.5. Zweistufige Schlüsselwortanalyse

Es wurden Schlüsselwörter für jeden einzelnen Text berechnet (mit Referenzkorpus Wikipedia). Die Tabelle führt auf, welche Schlüsselwörter in wie vielen Texten typisch sind. Gelistet werden alle Schlüsselwörter aus mind. 10 Texten. Einige Listeneinträge wurden gelöscht.

N	KW	Texts	%	Overall Freq.
1	Religionsunterricht	133	36,2	1878
2	Schülerinnen	101	27,5	1151
6	Bildung	81	22,1	1554
7	Gott	73	19,9	1412
8	Lernen	68	18,5	902
9	Gottes	66	18,0	870
10	Schüler	63	17,2	839
12	Religionspädagogik	56	15,3	767
13	Menschen	55	15,0	1034
14	Theologie	48	13,1	849

N	KW	Texts	%	Overall Freq.
15	Kontext	46	12,5	334
16	Unterricht	44	12,0	429
17	Kirche	41	11,2	779
18	Praxis	39	10,6	417
19	Kinder	39	10,6	689
20	Jugendlichen	38	10,4	411
22	Blick	36	9,8	277
23	Religionsunterrichts	36	9,8	389
24	Jesus	34	9,3	504
25	Perspektive	33	9,0	226
26	Fragen	31	8,4	295
27	Religionen	30	8,2	409
28	Lernens	29	7,9	264
29	Erfahrungen	29	7,9	275
31	Handeln	29	7,9	255
32	Schule	28	7,6	432
33	Jugendliche	27	7,4	290
34	Auseinandersetzung	26	7,1	170
35	Glaubens	26	7,1	173
36	Glauben	25	6,8	200
37	Kompetenzen	25	6,8	308
38	Jesu	25	6,8	393
40	Christus	25	6,8	256
41	Lernenden	25	6,8	224
42	Erziehung	25	6,8	295
43	Frage	24	6,5	324
44	Texte	23	6,3	305
45	Kindern	22	6,0	218
46	Kirchen	22	6,0	240
47	Forschung	21	5,7	184
48	Bibel	21	5,7	424
49	Reflexion	20	5,4	127
50	Wirklichkeit	20	5,4	136
52	Freiheit	19	5,2	258
53	Methoden	19	5,2	220
55	Schülern	18	4,9	136
56	Erzählungen	18	4,9	243
57	Schulen	18	4,9	339
59	Zugänge	17	4,6	118
60	Sekundarstufe	17	4,6	120
62	Perspektiven	17	4,6	116
63	Religiosität	16	4,4	175
64	Grundschule	16	4,4	132

N	KW	Texts	%	Overall Freq.
65	Christen	16	4,4	161
66	Mt	15	4,1	138
67	Pädagogik	15	4,1	133
70	Kompetenz	15	4,1	160
71	Mensch	15	4,1	161
72	Dimension	15	4,1	146
73	EKD	14	3,8	145
75	Ethik	14	3,8	207
76	Sinne	14	3,8	113
77	Leben	14	3,8	181
78	Verständnis	14	3,8	96
80	Kirchengeschichte	14	3,8	186
81	Kommunikation	13	3,5	97
82	Texten	13	3,5	101
84	Glaube	13	3,5	116
85	Unterrichts	13	3,5	151
86	Tradition	13	3,5	147
87	Entwicklung	13	3,5	165
88	Inhalte	12	3,3	83
89	Gesellschaft	12	3,3	180
90	Didaktik	12	3,3	111
91	Islam	12	3,3	131
92	Modell	12	3,3	155
93	Studien	12	3,3	116
94	Gegenwart	12	3,3	88
95	Gerechtigkeit	12	3,3	265
96	Judentum	12	3,3	127
97	Medien	12	3,3	210
98	Analyse	12	3,3	108
101	Staat	12	3,3	183
102	Identität	12	3,3	193
104	Ansatz	11	3,0	70
105	Medium	11	3,0	83
106	Lebenswelt	11	3,0	115
109	Religionsgemeinschaften	11	3,0	133
110	Wissen	11	3,0	92
113	Erzählung	11	3,0	114
116	Theorie	11	3,0	135
117	Herausforderungen	11	3,0	63
119	Konzept	11	3,0	98
122	Umgang	10	2,7	74
123	Begegnung	10	2,7	85
125	Orientierung	10	2,7	62

N	KW	Texts	%	Overall Freq.
127	Christentum	10	2,7	72
128	Traditionen	10	2,7	64
129	Geschichten	10	2,7	88
130	Geschichte	10	2,7	152

Verbindungen zwischen Schlüsselwörtern

Für das Schlüsselwort „Islam“ wurden Associates berechnet, also solche Wörter, die in den gleichen Texten typisch sind, in denen „Islam“ ein Schlüsselwort ist. Es wurde gefordert, dass dies in mindestens drei Texten der Fall sein muss. Einzelne Einträge wurden gelöscht.

N	Associate	Strength	Texts	%
1	Muslime	9,1	5	41
2	Dialog	7,8	4	33
3	Begegnung	7,6	4	33
4	Judentum	7,4	4	33
5	Religionen	7,0	5	41
6	Gott	6,5	6	50
7	Religion	6,2	6	50
8	Glaube	6,0	3	25
9	Gottes	5,9	5	41
10	Lernen	4,9	4	33
11	Jesus	4,6	3	25
14	Religionsunterricht	3,9	4	33
15	Schülerinnen	3,0	3	25

3. Wortverbindungen

3.1. Kollokationen

Es wurden Kollokationen des Suchwortes „Gott“ in einem Horizont von L5-R5 berechnet. Die Tabelle zeigt die 10 häufigsten Nomen sortiert nach Anzahl der Texte, in denen die Kollokation auftritt. Zusätzlich ist vermerkt, mit welchen weiteren Nomen das Wort im Umfeld von „Gott“ zusammen vorkommt (n>=5)

N	Word	Texts	Total	Shared Collocates
1	MENSCHEN	75	154	LEBEN, JESUS, VERHÄLTNIS
2	MENSCH	45	72	JESUS, CHRISTUS, VERHÄLTNIS

3	JESUS	35	57	MENSCH, MENSCHEN, KIRCHE, RELIGIONEN, FREIHEIT
4	WELT	32	38	MENSCH, MENSCHEN
5	GLAUBEN	28	36	
6	BEZIEHUNG	28	38	MENSCHEN
7	CHRISTUS	27	37	MENSCH, FREIHEIT, RELIGIONEN, KIRCHE
8	FRAGE	25	30	
9	LEBEN	23	26	MENSCHEN
10	GLAUBE	17	21	

3.2. Konkordanz

Hier ein beispielhafter Auszug aus einer Konkordanz des Suchwortes „Dialog“, der zeigt, wie sprachliche Qualifizierung auf unterschiedliche Konzepte verweist.

N	Concordance	Set	Tag	Word #	Sent	Sent #
23	Rede vom Altern in den so bestimmten Dialog mit den älteren Menschen und			2.425	90	20
24	zum einen vom Wort beziehungsweise Dialog und lässt zum anderen die			2.591	...	21
25	lösen, der christlich-islamische Dialog nicht die aktuellen Kriege im			2.151	...	17
26	heute. Für den christlich-islamischen Dialog zeigt sich diese Kontextualität u.a			1.578	97	5
27	oder christlich-islamischer Dialog geführt. In diesen fließen dann			330	24	6
28	Wirkung. Der christlich-jüdische Dialog in einer Stadt in Deutschland			2.138	...	4
29	Akzente gesetzt: Der christlich-jüdische Dialog insgesamt ist seinerseits			860	43	12
30	Taktgebern, christlich-jüdischem Dialog und Religionsdidaktik, haben die			899	45	8
31	Beziehungen: „ Dem christlich-jüdischen Dialog kann man sich als Christ relativ			1.301	56	44
32	, die den christlich-jüdischen Dialog , speziell seine Belastungen,			2.439	...	14
33	Fragen, die im christlich-jüdischen Dialog eine Rolle spielen: vom			2.328	91	9
34	„ christlich-jüdischer Dialog “ nennt, eine nicht minder große			774	41	13
35	, so als würde sich etwa aus dem ? Dialog der Religionen automatisch der			151	11	2
36	Erkenntnisgewinn erwächst aus dem Dialog zwischen den Erzählungen und			4.437	...	7

concordance collocates plot patterns clusters timeline filenames source text notes

322 entries Row 35 0% T S