

Noah Bubenhofer

Sprachgebrauchsmuster

Korpuslinguistik als Methode
der Diskurs- und Kulturanalyse



Walter de Gruyter · Berlin · New York

Inhalt

1	Einleitung	I
1.1	Eingrenzung	4
1.2	Hypothesenbildung	5
1.3	Terminologie und allgemeine Hinweise	8
I	Muster als Effekte des Sprechens: Theoretische Perspektive	II
2	Grundbegriffe	15
2.1	Sprachgebrauch korpuslinguistisch	15
2.2	Muster	18
2.2.1	Bedeutungsvielfalt	18
2.2.2	Klärung: Beispiele und Definition	21
2.2.3	Muster als Muster erkennen	24
2.2.4	Abgrenzungen und Fazit	29
2.3	Diskurse	31
2.3.1	Foucault'sche Diskursanalyse	31
2.3.2	Linguistische Diskursanalyse	32
2.3.3	Kritik	35
2.3.4	Fazit	38
2.4	Kultur	39
2.4.1	Kultur als Bedeutungsgewebe	39
2.4.2	Kultur als semiotisches System	40
2.4.3	Fazit	42
3	Die Bedeutung musterhaften Sprachgebrauchs	43
3.1	Sprachgebrauchsmuster als Effekte sozialen Sprachhandelns	43
3.1.1	Formelhafte Textoberfläche – Kontextualisierung	43
3.1.2	Sozialität und Kulturalität von Sprache	46

3.1.3	Common Sense	50
3.2	Sprache lesen, Kultur lesen	52
4	Was sind Sprechweisen?	55
4.1	Stil	55
4.2	Kommunikative Gattungen	57
4.3	Mentalitäten und diskurssemantische Grundfiguren	60
4.4	Typik des Verhaltens und idiomatische Prägungen	62
4.5	Argumentationsfiguren, Topoi	67
4.6	Metaphern	77
4.7	Sprachgebrauchsgeschichte als Kulturanalyse	84
4.8	Die Einordnung der Konzepte	87
II	Mustererkennung als Basis für die Analyse:	
	Methodische Herleitung	95
5	Grundsätze	99
5.1	Daten lesen: ‚corpus-based‘ und ‚corpus-driven‘	99
5.2	Daten, Muster, Beschreibung: Die Heuristik im Überblick	102
5.3	Wörter, Sätze, Texte? Die Elemente und Grenzen eines Diskurses bestimmen	105
6	Musterhafte Strukturen finden: Die Klumpen im Text	111
6.1	Kollokationen, Kookkurrenzen, n-Gramme	111
6.2	Syntagmatische Muster	118
6.3	Zwischenfazit: Terminologische Festlegung	121
6.4	Kollokationen, corpus-driven	123
6.5	Lemmata und Wortarten: Gewinn oder Verlust an Information?	124
7	Der statistische Zugang zu sprachlichen Daten	131
7.1	Deskriptive Statistik	132
7.2	Testen von Hypothesen	134
7.2.1	Klumpen: Wörter sind nie zufällig verteilt	134
7.2.2	Kontingenztafel erstellen	136
7.2.3	Chi-Quadrat-Test	137

7.2.4	Log-Likelihood G^2	139
7.2.5	Weitere Tests	139
7.3	Weitere Hilfsmittel	142
7.3.1	Dispersion: Juillands D	142
7.3.2	Differenzkoeffizient	144
7.4	Fazit	144
8	Von corpus-driven zu corpus-based	149
8.1	Typikprofile erstellen	149
8.2	Interpretieren: Diskurse modellieren	152
8.2.1	Grad der Typik	155
8.2.2	Kontextualisierungsprofile	160
8.2.3	Paradigmatische Konstellationen	163
8.2.4	Pragmatische Funktionen	168
8.3	Überprüfen: Zurück in die Daten	171
9	Ressourcen der Korpus- und Computerlinguistik	175
9.1	Öffentlich verfügbare Referenzkorpora	175
9.2	Software zur Korpusabfrage	180
9.3	Software für die Berechnung von Kollokationen	181
III Anwendungsbeispiele		185
10	Die Datengrundlage: Das NZZ-Korpus	189
10.1	Publizistische Daten zur ‚Neuen Zürcher Zeitung‘	189
10.2	Eckdaten des Korpus	190
11	Aufbereitung der Daten	193
12	Berechnung der typischen Mehrworteinheiten	197
12.1	Festlegung von Parametern und Teilkorpora	197
12.2	Selektion der Mehrworteinheiten für die Analyse	200
13	Beispielanalysen	209
13.1	Die Auslandsberichterstattung im NZZ-Korpus	209
13.1.1	Grobanalyse	209
13.1.2	Die <i>bosnischen Serben</i> : Die Verwendung von Ethnienbezeichnungen	212

13.1.3	Kontextualisierungsprofile: <i>Kampf gegen X, Kampf dem X oder Kampf mit X?</i>	220
13.1.4	Was ist wert, gezählt zu werden? <i>Die Zahl der X</i>	229
13.2	Die Inlandsberichterstattung im NZZ-Korpus	240
13.2.1	Grobanalyse	241
13.2.2	Routinisierung der medialen Selbstreflexion: <i>vor den Medien</i>	242
13.2.3	Das Sorgenbarometer auf der Textoberfläche: <i>Kampf gegen X, Bekämpfung von X</i>	248
13.2.4	Argumentieren und bewerten: <i>nicht nur ... sondern auch</i>	257
13.3	Weitere Ressorts des NZZ-Korpus	268
13.3.1	<i>Zum ersten Mal und die moderne, westliche, bürgerliche Gesellschaft</i>	269
13.3.2	<i>Es ist nicht mehr wie früher.</i>	269
13.3.3	Blick in die Vergangenheit: <i>die -iger Jahre</i>	272
13.3.4	Veränderungen in der Wirtschaftswelt	273
13.3.5	Geografische Referenzen	274
13.3.6	Der nördliche Nachbar: <i>Die deutsche(n) X</i>	280
13.3.7	Vom Schweizer Skiverband zu <i>Swiss Ski</i>	285
13.3.8	Die <i>Damen</i> und die <i>Frauen</i> im Sport	287
13.3.9	Ressorts im synchronen Vergleich: Sprachgebrauch in Leserbriefen	291
14	Hypothesenbildung	295
14.1	Sprechen über KRIEG UND GEWALT	296
14.2	Sprechen über SORGEN UND PROBLEME	298
14.3	Sprachgebrauch und Welt	300
14.4	Sprachgebrauch und Text	302
IV	Fazit und Ausblick	305
15	Theorie und Praxis	309
15.1	Sprachgebrauchsmuster als Kristallisationkerne von Diskursen	309
15.2	Semantische Konzepte in der korpuslinguistischen Operationalisierung	312

16	Praxis und Theorie	317
16.1	Der Gewinn des korpuslinguistischen Blicks für die Theoriebildung	317
16.1.1	Diskurse, Sprechweise, Themen	317
16.1.2	Kontingente Diskursdefinitionen	318
16.1.3	Semantische Matrizen typischer Sprechweisen	319
16.2	Nutzen einer corpus-driven-Analyse für die Diskurs- und Kulturanalyse	320
16.2.1	Korpuslinguistik und Diskursanalyse	320
16.2.2	Korpuslinguistik und Kulturanalyse	322
17	Ausblick	325
17.1	Andere Daten, ergänzte Methodik: Ein Beispiel	325
17.2	Schlüsse, Möglichkeiten, Desiderata	332
17.2.1	Korpora	332
17.2.2	Statistik	333
17.2.3	Semantische Kategorisierungen und Annotation	335
17.2.4	Sprachgebrauchsmuster als kleinste semantische Einheiten	336
V	Anhang	339
A	Typikprofile	341
A.1	Ressort ‚Ausland‘	341
A.1.1	Periode 1995–1997	341
A.1.2	Periode 2003–2005	343
A.2	Ressort ‚Inland‘	345
A.2.1	Periode 1995–1997	345
A.2.2	Periode 2003–2005	347
A.3	Ressort ‚Feuilleton‘	350
A.3.1	Periode 1995–1997	350
A.3.2	Periode 2003–2005	351
A.4	Ressort ‚Wirtschaft‘	352
A.4.1	Periode 1995–1997	352
A.4.2	Periode 2003–2005	355
A.5	Ressort ‚Sport‘	357
A.5.1	Periode 1995–1997	357

A.5.2 Periode 2003–2005	358
A.6 Ressort ‚Leserbriefe‘	359
Tabellenverzeichnis	361
Abbildungsverzeichnis	365
Literatur	369
Nachschlagewerke	381
Quellentexte und Korpora	383
Stichwortverzeichnis	385