

Inhalt

Danksagung	IX
English Summary	XI
1. Einleitung	1
1.1 Gegenstand und Zielsetzung	3
1.2 Aufbau der Arbeit	5
2. Das Bedeutungspotential sprachlicher Einheiten	9
2.1 Die Variabilität sprachlicher Bedeutung	9
2.1.1 Ambiguität	10
2.1.2 Bedeutungsvariation	13
2.2 Zusammenfassung	14
3. Bedeutungspotential und Bedeutungskonstitution	17
3.1 Kognitive Linguistik	17
3.1.1 Holistischer Ansatz	19
3.1.2 Sprache als semantisches Wissen	22
3.2 Kognitive Semantik	23
3.2.1 Bedeutung als Potential	24
3.2.2 Bedeutung als Prozess	28
3.2.3 Implikationen für die Modellierung	32
3.3 Zusammenfassung	35
4. Das Word Space Model	39
4.1 Grundkonzeption des Modells	40
4.1.1 Der Wortraum	40
4.1.2 Wörter als Vektoren	42
4.1.3 Kontextvektoren und Kookkurrenzen zweiter Ordnung	47
4.1.4 Zusammenfassung	49
4.2 Theoretische Grundlagen des Modells	50
4.2.1 Der Word Space als semantischer Raum	51
4.2.2 Die distributionelle Hypothese	52
4.2.3 Diskussion	54
4.3 Zusammenfassung	56
5. Bedeutungskonstitution im Vektorraum	59
5.1 Repräsentation von Input und Output	59
5.1.1 Bedeutungspotential im Vektorraum	60
5.1.2 Input und Output als Vektoren	61

5.2	Bedeutungskonstitution als Transformation von Vektoren	63
5.2.1	Transformation durch den Kontext	64
5.2.2	Gewichtung der Kontexte	68
5.2.3	Mehrdeutigkeit im Vektorraum	71
5.3	Diskussion	73
6.	Softwaretechnologische Umsetzung	77
6.1	Das Text Engineering Software Laboratory (Tesla)	78
6.1.1	Experimente im virtuellen Labor	79
6.1.2	Arbeiten im virtuellen Labor	80
6.1.3	Das Tesla Role System	84
6.2	Verfahrensschritte und Komponenten	86
6.2.1	Korpora	87
6.2.2	Vorverarbeitung	88
6.2.3	Kookkurrenzvektoren	89
6.2.4	Normalisierung	90
6.2.5	Gewichtung	91
6.2.6	Token-Vektoren	92
6.2.7	Clusteranalyse	93
6.2.8	Visualisierung	95
6.2.9	Beispielwörter für die Experimente	97
6.3	Zusammenfassung	101
7.	Experimente zur Bedeutungskonstitution	105
7.1	Repräsentation der Eingabeinformation	105
7.1.1	Aufbau des Experiments	106
7.1.2	Parametrisierung	108
7.1.3	Referenzräume	110
7.2	Bedeutungskonstitution in Einzelkontexten	114
7.2.1	Aufbau des Experiments	115
7.2.2	Parametrisierung	117
7.2.3	Beispielanalysen	118
7.3	Semantische Profile	126
7.3.1	Aufbau des Experiments	127
7.3.2	Parametrisierung	127
7.3.3	Beispielanalysen	130
7.4	Zusammenfassung	135
8.	Fazit: Muster und Bedeutung	139
A.	Komponenten	143
A.1	Reader	143
A.1.1	LCC Reader	143

A.1.2	SdeWaC Reader	143
A.2	Vorverarbeitung	144
A.2.1	Simple Tokenizer	144
A.2.2	Tree Tagger Wrapper	144
A.2.3	Snowball Stemmer Wrapper	145
A.3	Vektorerstellung	145
A.3.1	Sentence Based Vector Generator	145
A.3.2	Punctuation Filter	146
A.3.3	Frequency Range Filter	146
A.3.4	POSFilter	147
A.3.5	Wordlist Filter	148
A.4	Normalisierung und Gewichtung	148
A.4.1	VectorNormalization	148
A.4.2	VectorWeighting	148
A.5	Repräsentation von Einzelvorkommen	149
A.5.1	CollocationVectors	149
A.5.2	Context Vectors	150
A.5.3	Sentence Vectors	151
A.6	Clustering	152
A.6.1	Distanzbasierte Verfahren	152
A.6.2	Dichthebasierte Verfahren	154
A.6.3	ClusterFilter	155
A.7	Visualisierung	155
B.	Experimente	157
B.1	Kookkurrenzvektoren und Referenzräume	158
B.2	Bedeutungskonstitution in Einzelkontexten	160
B.3	Semantische Profile	161
C.	Assoziationsmaße	163
C.1	Pointwise Mutual Information	163
C.2	Log-Likelihood-Ratio	164
	Abbildungsverzeichnis	167
	Literaturverzeichnis	169