

A	GESCHICHTE, SYSTEMATISCHER ORT UND ZIELSETZUNG DER MASCHINELLEN LEMMATISIERUNG	
1	Notwendigkeit der maschinellen Lemmatisierung	
1.1	Der Terminus 'Lemma'	1
1.2	Lemmatisierte Textwörterbücher	3
1.3	Lemmatisierung als Teilprozess der maschinellen Sprachübersetzung	6
2	Allgemeine Zielsetzung	7
3	Mögliche Verfahrensweisen	9
3.1	Inputintensive Verfahren	9
3.2	Programmintensive Verfahren	11
3.2.1	Informationsintensive Verfahren	12
3.2.2	Regelintensive Verfahren	14
B	LINGUISTISCHE GRUNDLAGEN DER LEMMATISIERUNG	
1	Lemma als Menge linguistischer Einheiten	19
1.1	Wort - Lemma	19
2	Semantische Merkmale der Lemmatisierung	23
2.1	Hypothese	24
2.2	Problematische Konsequenzen	
2.2.1	Unterschiedliche g_c bei gleicher Bedeutung	
2.2.1.1	Trennung synonyme Wortformen	25
2.2.1.2	Orthographische Varianten	26
2.2.1.3	Flexionsbedingter Wechsel von g_c	27
2.2.2	Gleiche g_c bei verschiedener Bedeutung	27
2.2.3	Verbalkomposita	31
3	Die Relevanz des Funktionsteils für die Lemmatisierung	
3.1	Die paradigmatischen Eigenschaften	32

VI

3.1.1	Wortklassen	33
3.1.2	Homographie	38
3.1.2.1	Die Homographie Verb/Adjektiv	42
3.1.3	Flexionsklassen	46
3.1.4	Präfix - Verbzusatz	49
3.1.5	Reflexivität als E_p	51
3.2	Die strukturalen Eigenschaften	60
3.2.1	‡ Vollverb	61
3.2.2	Der syntaktische Kontext	62
3.2.3	Der semantische Kontext	75
3.2.3.1	Verbklassifizierung nach Denkmodellen	80
3.2.3.2	Die Subkategorisierung der 'Aspekte...'	84
3.2.3.3	Die Kasustheorie von Fillmore	95
3.2.3.3.1	Die inhärente Subkategorisierung	115
3.2.3.3.2	Die Subkategorisierungsmerkmale der Verben	121
4	Einige Analysegrundsätze der Kasusgrammatik	123
4.1	Zusatzinformationen	128
4.2	Zur Strategie des Reduktionsteils	137
C	LEMMATISIERUNG ALS PROJEKT DER LINGUISTISCHEN DATENVERARBEITUNG	146
1	Die Erkennung der Wortklasse (E_p (WK)) durch Endungsanalyse	147
1.1	<i>-keit, -keiten</i>	148
1.2	<i>-heit, -heiten</i>	149
1.3	<i>-ung, -ungen</i>	149
2	Die Erkennung der Flexionsformen	153
2.1	Regelmäßige und unregelmäßige Verben	156
2.2	Besonderheiten in der morphologischen Analyse von R- und U-Verben	160
2.2.1	Zur Methode des Wörterbuchvergleichs	162
2.2.2	Eingeschränkte Flektierbarkeit	165
2.2.3	PART II ohne <i>ge-</i>	166
2.2.4	Besonderheiten in der Flexion von R-Verben	170
2.2.4.1	Endung des INF = <i>n</i>	170
2.2.4.2	g_c -Morph endet auf <i>s</i> , <i>z</i> oder <i>x</i>	173
2.2.4.3	Eingeschobenes <i>e</i>	176

2.2.5	Die Flexion und Analyse der U-Verben	179
2.2.5.1	Paradigmen und Paradigmenklassen	180
2.2.5.2	Allgemeine Reduktionsregeln	182
2.2.5.3	Abweichungen	
2.2.5.3.1	Die Gruppe <i>weisen, fressen, wachsen, geniessen, sitzen</i> und andere	184
2.2.5.3.2	Das Verb <i>bersten</i>	186
2.2.5.3.3	Die Gruppe <i>bieten, binden, braten, reiten</i> und andere	186
2.2.5.4	Die Paradigmenklassen der U-Verben	187
2.2.5.5	Die Kodierung der unregelmäßigen Verben	199
2.2.5.6	Die Erzeugung des Lemmanamens bei unregelmäßigen Verben	204
3	Zur Zusammenfassung zu Wortformen	209
3.1	Getrennte Flexionselemente	209
3.2	Getrennte Verbzusätze	210
3.2.1	Die Wortklasse Verbzusatz	210
3.2.2	Die Wortklassenmehrdeutigkeit der VZS	211
3.2.3	Kriterien zur Auflösung der Homographie	213
4	Der Aufbau der Lexikoneinheit	214
D	SCHLUSSBEMERKUNGEN	219
E	REGISTER	220
F	BIBLIOGRAPHIE	226
G	VERZEICHNIS DER ABKÜRZUNGEN	231