

2.3.3	Grouping with Word2Vec	31
2.4	Experiments	32
2.4.1	Setup	32
2.4.2	Evaluation	33
2.4.3	Results	34
2.5	Conclusions and Future Work	36
3	DISCOVer: DIStributional Approach Based on Syntactic Dependencies for Discovering COnstructions	39
3.1	Introduction	40
3.2	Related Work	42
3.3	Methodology for Discovering Constructions	45
3.3.1	Description of the Task	46
3.3.2	The Corpus	46
3.3.3	Matrix	47
3.3.4	Clustering	50
3.3.5	Generalization: Linking and Filtering Clusters	53
3.3.6	Pattern Generation	55
3.4	Evaluation	58
3.4.1	Clustering Evaluation	58
3.4.2	Pattern Evaluation	58
3.5	Conclusions and Future Work	65
II	Paraphrase Typology and Paraphrase Identification	67
4	WARP-Text: a Web-Based Tool for Annotating Relationships between Pairs of Texts	69
4.1	Introduction	70
4.2	Related Work	70
4.3	WARP-Text	71
4.3.1	Annotation Scheme	72
4.3.2	Administrator Interface	72
4.3.3	Annotator Interface	73
4.4	Conclusions and Future Work	75
5	ETPC - a Paraphrase Identification Corpus Annotated with Extended Paraphrase Typology and Negation	77
5.1	Introduction	78
5.2	Related Work	79

5.3	Extended Paraphrase Typology	80
5.3.1	Basic Terminology	81
5.3.2	From Atomic to Textual Paraphrases	81
5.3.3	Objectives of EPT and Research Questions.	83
5.3.4	The Extended Paraphrase Typology	83
5.4	Annotation Scheme and Guidelines	86
5.4.1	Non-Sense Preserving Atomic Phenomena	87
5.4.2	Sense Preserving Atomic Phenomena	88
5.4.3	Inter-Annotator Agreement	89
5.4.4	Annotation of Negation	91
5.5	The ETPC Corpus	91
5.5.1	Non-Sense Preserving Atomic Phenomena	91
5.5.2	Sense Preserving Atomic Phenomena	92
5.5.3	Discussion	93
5.5.4	Applications of ETPC	94
5.6	Conclusions and Future Work	95
6	A Qualitative Evaluation Framework for Paraphrase Identification	97
6.1	Introduction	98
6.2	Related Work	99
6.3	Qualitative Evaluation Framework	100
6.3.1	The ETPC Corpus	100
6.3.2	Evaluation Methodology	100
6.4	PI Systems	101
6.5	Results	103
6.5.1	Overall Performance	103
6.5.2	Full Performance Profile	104
6.5.3	Comparing Performance Profiles	107
6.5.4	Comparing Performance by Phenomena	108
6.6	Discussion	109
6.7	Conclusions and Future Work	110
III	Paraphrasing, Textual Entailment, and Semantic Similarity	113
7	Annotating and Analyzing the Interactions between Meaning Relations	115
7.1	Introduction	116
7.2	Related Work	117

7.2.1	Interactions between Relations	117
7.2.2	Corpora with Multiple Semantic Layers	118
7.3	Corpus Creation	119
7.3.1	Sentence Pool	119
7.3.2	Pair Generation	120
7.3.3	Relation Annotation	121
7.3.4	Final Corpus	123
7.4	Interactions between Relations	126
7.4.1	Correlations between Relations	126
7.4.2	Overlap of Relation Labels	128
7.4.3	Discussion	130
7.5	Conclusion and Further Work	131
8	Decomposing and Comparing Meaning Relations: Paraphrasing, Textual Entailment, Contradiction, and Specificity	133
8.1	Introduction	134
8.2	Related Work	136
8.3	Shared Typology for Meaning Relations	137
8.3.1	Decomposing Meaning Relations	138
8.3.2	The SHARel Typology	138
8.3.3	Research Questions	141
8.4	Corpus Annotation	142
8.4.1	Choice of Corpus	142
8.4.2	Annotation Setup	143
8.4.3	Agreement	143
8.5	Analysis of the Results	145
8.5.1	Type Frequency	145
8.5.2	Decomposing Specificity	149
8.6	Discussion	151
8.7	Conclusions and Future Work	152
9	Conclusions	155
9.1	Contributions and Discussion of the Results	155
9.2	Resources	160
9.3	Future Research Directions	160
A	Annotation Guidelines for ETPC	183
A.1	Presentation	183
A.1.1	Credits	183
A.2	The task	184