

---

## Contents

---

<b>Preface</b> . . . . .	ix
<b>Introduction</b> . . . . .	xi
<b>Chapter 1. From Opinion Analysis to Figurative Language Treatment</b> . . . . .	1
1.1. Introduction . . . . .	1
1.2. Defining the notion of opinion . . . . .	3
1.2.1. The many faces of opinion . . . . .	3
1.2.2. Opinion as a structured model . . . . .	4
1.2.3. Opinion extraction: principal approaches . . . . .	5
1.3. Limitations of opinion analysis systems . . . . .	7
1.3.1. Opinion operators . . . . .	8
1.3.2. Domain dependency . . . . .	9
1.3.3. Implicit opinions . . . . .	10
1.3.4. Opinions and discursive context above phrase level . . . . .	11
1.3.5. Presence of figurative expressions . . . . .	12
1.4. Definition of figurative language . . . . .	13
1.4.1. Irony . . . . .	13
1.4.2. Sarcasm . . . . .	18
1.4.3. Satire . . . . .	20
1.4.4. Metaphor . . . . .	21
1.4.5. Humor . . . . .	22
1.5. Figurative language: a challenge for NLP . . . . .	23
1.6. Conclusion . . . . .	23

<b>Chapter 2. Toward Automatic Detection of Figurative Language</b> . . . . .	25
2.1. Introduction . . . . .	25
2.2. The main corpora used for figurative language . . . . .	27
2.2.1. Corpora annotated for irony/sarcasm . . . . .	28
2.2.2. Corpus annotated for metaphors . . . . .	33
2.3. Automatic detection of irony, sarcasm and satire . . . . .	36
2.3.1. Surface and semantic approaches . . . . .	36
2.3.2. Pragmatic approaches . . . . .	39
2.4. Automatic detection of metaphor . . . . .	51
2.4.1. Surface and semantic approaches . . . . .	52
2.4.2. Pragmatic approaches . . . . .	53
2.5. Automatic detection of comparison . . . . .	58
2.6. Automatic detection of humor . . . . .	58
2.7. Conclusion . . . . .	61
<b>Chapter 3. A Multilevel Scheme for Irony Annotation in Social Network Content</b> . . . . .	63
3.1. Introduction . . . . .	63
3.2. The FrIC . . . . .	65
3.3. Multilevel annotation scheme . . . . .	66
3.3.1. Methodology . . . . .	66
3.3.2. Annotation scheme . . . . .	69
3.4. The annotation campaign . . . . .	79
3.4.1. Glozz . . . . .	79
3.4.2. Data preparation . . . . .	80
3.4.3. Annotation procedure . . . . .	81
3.5. Results of the annotation campaign . . . . .	83
3.5.1. Qualitative results . . . . .	83
3.5.2. Quantitative results . . . . .	84
3.5.3. Correlation between different levels of the annotation scheme . . . . .	89
3.6. Conclusion . . . . .	93
<b>Chapter 4. Three Models for Automatic Irony Detection</b> . . . . .	95
4.1. Introduction . . . . .	95
4.2. The FrIC <sup>Auto</sup> corpus . . . . .	97

---

4.3. The SurfSystem model: irony detection based on surface features . . . . .	99
4.3.1. Selected features . . . . .	99
4.3.2. Experiments and results . . . . .	101
4.4. The PragSystem model: irony detection based on internal contextual features . . . . .	104
4.4.1. Selected features . . . . .	104
4.4.2. Experiments and results . . . . .	109
4.4.3. Discussion . . . . .	116
4.5. The QuerySystem model: developing a pragmatic contextual approach for automatic irony detection . . . . .	118
4.5.1. Proposed approach . . . . .	118
4.5.2. Experiments and results . . . . .	122
4.5.3. Evaluation of the query-based method . . . . .	123
4.6. Conclusion . . . . .	124
<b>Chapter 5. Towards a Multilingual System for Automatic Irony Detection . . . . .</b>	<b>127</b>
5.1. Introduction . . . . .	127
5.2. Irony in Indo-European languages . . . . .	128
5.2.1. Corpora . . . . .	128
5.2.2. Results of the annotation process . . . . .	130
5.2.3. Summary . . . . .	139
5.3. Irony in Semitic languages . . . . .	140
5.3.1. Specificities of Arabic . . . . .	142
5.3.2. Corpus and resources . . . . .	143
5.3.3. Automatic detection of irony in Arabic tweets . . . . .	146
5.4. Conclusion . . . . .	149
<b>Conclusion . . . . .</b>	<b>151</b>
<b>Appendix . . . . .</b>	<b>155</b>
<b>References . . . . .</b>	<b>169</b>
<b>Index . . . . .</b>	<b>189</b>