
Contents

Preface	xi
Chapter 1. Delayed Interpretation, Shallow Processing and Constructions: the Basis of the “<i>Interpret Whenever Possible</i>” Principle	1
Philippe BLACHE	
1.1. Introduction.	1
1.2. Delayed processing	3
1.3. Working memory	5
1.4. How to recognize chunks: the segmentation operations	8
1.5. The delaying architecture.	10
1.5.1. Segment-and-store	11
1.5.2. Aggregating by cohesion	12
1.6. Conclusion	16
1.7. Bibliography	17
Chapter 2. Can the Human Association Norm Evaluate Machine-Made Association Lists?	21
Michał KORZYCKI, Izabela GATKOWSKA and Wiesław LUBASZEWSKI	
2.1. Introduction.	21
2.2. Human semantic associations	23
2.2.1. Word association test.	23
2.2.2. The author’s experiment.	24
2.2.3. Human association topology	25
2.2.4. Human associations are comparable.	26

2.3. Algorithm efficiency comparison	29
2.3.1. The corpora	29
2.3.2. LSA-sourced association lists.	29
2.3.3. LDA-sourced lists	31
2.3.4. Association ratio-based lists	31
2.3.5. List comparison	32
2.4. Conclusion	37
2.5. Bibliography	38
Chapter 3. How a Word of a Text Selects the Related Words in a Human Association Network	41
Wiesław LUBASZEWSKI, Izabela GATKOWSKA and Maciej GODNY	
3.1. Introduction.	41
3.2. The network	44
3.3. The network extraction driven by a text-based stimulus	46
3.3.1. Sub-graph extraction algorithm.	46
3.3.2. The control procedure	48
3.3.3. The shortest path extraction.	48
3.3.4. A corpus-based sub-graph.	50
3.4. Tests of the network extracting procedure.	50
3.4.1. The corpus to perform tests	50
3.4.2. Evaluation of the extracted sub-graph.	51
3.4.3. Directed and undirected sub-graph extraction: the comparison	52
3.4.4. Results per stimulus	53
3.5. A brief discussion of the results and the related work	58
3.6. Bibliography	60
Chapter 4. The Reverse Association Task	63
Reinhard RAPP	
4.1. Introduction.	63
4.2. Computing forward associations	67
4.2.1. Procedure.	67
4.2.2. Results and evaluation	69
4.3. Computing reverse associations.	71
4.3.1. Problem.	71
4.3.2. Procedure.	71
4.3.3. Results and evaluation	76

4.4. Human performance	78
4.4.1. Dataset	78
4.4.2. Test procedure	80
4.4.3. Evaluation	81
4.5. Performance by machine	82
4.6. Discussion, conclusions and outlook	84
4.6.1. Reverse associations by a human	84
4.6.2. Reverse associations by a machine	85
4.7. Acknowledgments	87
4.8. Bibliography	88
Chapter 5. Hidden Structure and Function in the Lexicon	91
Philippe VINCENT-LAMARRE, Mélanie LORD, Alexandre BLONDIN-MASSÉ, Odile MARCOTTE, Marcos LOPES and Stevan HARNAD	
5.1. Introduction	91
5.2. Methods	92
5.2.1. Dictionary graphs	92
5.2.2. Psycholinguistic variables	96
5.2.3. Data analysis	96
5.3. Psycholinguistic properties of Kernel, Satellites, Core, MinSets and the rest of each dictionary	97
5.4. Discussion	101
5.4.1. Limitations	104
5.5. Future work	104
5.6. Bibliography	106
Chapter 6. Transductive Learning Games for Word Sense Disambiguation	109
Rocco TRIPODI and Marcello PELILLO	
6.1. Introduction	109
6.2. Graph-based word sense disambiguation	111
6.3. Our approach to semi-supervised learning	113
6.3.1. Graph-based semi-supervised learning	113
6.3.2. Game theory and game dynamics	114
6.4. Word sense disambiguation games	116
6.4.1. Graph construction	116
6.4.2. Strategy space	117
6.4.3. The payoff matrix	118
6.4.4. System dynamics	119

6.5. Evaluation	120
6.5.1. Experimental setting	120
6.5.2. Evaluation results	121
6.5.3. Comparison with state-of-the-art algorithms	124
6.6. Conclusion	124
6.7. Bibliography	125
Chapter 7. Use Your Mind and Learn to Write: The Problem of Producing Coherent Text	129
Michael ZOCK and Debela Tesfaye GEMECHU	
7.1. The problem	129
7.2. Suboptimal texts and some of the reasons	131
7.2.1. Lack of coherence or cohesion	132
7.2.2. Faulty reference	133
7.2.3. Unmotivated topic shift	134
7.3. How to deal with the complexity of the task?	135
7.4. Related work	136
7.5. Assumptions concerning the building of a tool assisting the writing process	138
7.6. Methodology	141
7.6.1. Identification of the syntactic structure	143
7.6.2. Identification of the semantic seed words	144
7.6.3. Word alignment	145
7.6.4. Determination of the similarity values of the aligned words	146
7.6.5. Determination of the similarity between sentences	150
7.6.6. Sentence clustering based on their similarity values	151
7.7. Experiment and evaluation	151
7.8. Outlook and conclusion	154
7.9. Bibliography	155
Chapter 8. Stylistic Features Based on Sequential Rule Mining for Authorship Attribution	159
Mohamed Amine BOUKHALED and Jean-Gabriel GANASCIA	
8.1. Introduction and motivation	159
8.2. The authorship attribution process	162
8.3. Stylistic features for authorship attribution	163
8.4. Sequential data mining for stylistic analysis	165

8.5. Experimental setup	166
8.5.1. Dataset	166
8.5.2. Classification scheme	167
8.6. Results and discussion	169
8.7. Conclusion	173
8.8. Bibliography	173
Chapter 9. A Parallel, Cognition-oriented Fundamental Frequency Estimation Algorithm	177
Ulrike GLAVITSCH	
9.1. Introduction.	177
9.2. Segmentation of the speech signal	180
9.2.1. Speech and pause segments	180
9.2.2. Voiced and unvoiced regions	182
9.2.3. Stable and unstable intervals	183
9.3. F0 estimation for stable intervals	184
9.4. F0 propagation	186
9.4.1. Control flow	187
9.4.2. Peak propagation	189
9.5. Unstable voiced regions	191
9.6. Parallelization	191
9.7. Experiments and results	192
9.8. Conclusions.	194
9.9. Acknowledgments.	195
9.10. Bibliography	195
Chapter 10. Benchmarking n-grams, Topic Models and Recurrent Neural Networks by Cloze Completions, EEGs and Eye Movements	197
Markus J. HOFMANN, Chris BIEMANN and Steffen REMUS	
10.1. Introduction	198
10.2. Related work	199
10.3. Methodology	200
10.3.1. Human performance measures	200
10.3.2. Three flavors of language models	201
10.4. Experiment setup.	203
10.5. Results	204
10.5.1. Predictability results	204
10.5.2. N400 amplitude results.	206
10.5.3. Single-fixation duration (SFD) results.	208

10.6. Discussion and conclusion	210
10.7. Acknowledgments	212
10.8. Bibliography	212
List of Authors	217
Index	219