

1. Record Nr.	UNINA9910139008903321
Autore	Schuller Bjorn
Titolo	Computational paralinguistics : emotion, affect and personality in speech and language processing / / Bjorn W. Schuller, Anton M. Batliner
Pubbl/distr/stampa	Hoboken, New Jersey : , : John Wiley & Sons, , [2014] ©2014
ISBN	1-118-70662-5 1-118-70666-8 1-118-70663-3
Edizione	[First edition.]
Descrizione fisica	1 online resource (xxi, 321 pages) : illustrations
Altri autori (Persone)	BatlinerAnton
Disciplina	401/.90285
Soggetti	Computational linguistics Emotive (Linguistics) Human-computer interaction Language and emotions Linguistic models - Data processing Paralinguistics Psycholinguistics - Data processing Speech processing systems
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Preface xiii Acknowledgements xv List of Abbreviations xvii Part I Foundations 1 Introduction 3 1.1 What is Computational Paralinguistics? A First Approximation 3 1.2 History and Subject Area 7 1.3 Form versus Function 10 1.4 Further Aspects 12 1.4.1 The Synthesis of Emotion and Personality 12 1.4.2 Multimodality: Analysis and Generation 13 1.4.3 Applications, Usability and Ethics 15 1.5 Summary and Structure of the Book 17 References 18 2 Taxonomies 21 2.1 Traits versus States 21 2.2 Acted versus Spontaneous 25 2.3 Complex versus Simple 30 2.4 Measured versus Assessed 31 2.5 Categorical versus Continuous 33 2.6 Felt versus Perceived 35 2.7 Intentional versus Instinctual 37 2.8 Consistent versus Discrepant 38 2.9 Private versus Social 39 2.10 Prototypical versus Peripheral 40 2.11

Universal versus Culture-Specific	41
2.12 Unimodal versus Multimodal	43
2.13 All These Taxonomies - So What?	44
2.13.1 Emotion Data: The FAU AEC	45
2.13.2 Non-native Data: The C-AuDIT corpus	47
References	48
3 Aspects of Modelling	53
3.1 Theories and Models of Personality	53
3.2 Theories and Models of Emotion and Affect	55
3.3 Type and Segmentation of Units	58
3.4 Typical versus Atypical Speech	60
3.5 Context	61
3.6 Lab versus Life, or Through the Looking Glass	62
3.7 Sheep and Goats, or Single Instance Decision versus Cumulative Evidence and Overall Performance	64
3.8 The Few and the Many, or How to Analyse a Hamburger	65
3.9 Reifications, and What You are Looking for is What You Get	67
3.10 Magical Numbers versus Sound Reasoning	68
References	74
4 Formal Aspects	79
4.1 The Linguistic Code and Beyond	79
4.2 The Non-Distinctive Use of Phonetic Elements	81
4.2.1 Segmental Level: The Case of /r/ Variants	81
4.2.2 Supra-segmental Level: The Case of Pitch and Fundamental Frequency - and of Other Prosodic Parameters	82
4.2.3 In Between: The Case of Other Voice Qualities, Especially Laryngealisation	86
4.3 The Non-Distinctive Use of Linguistics Elements	91
4.3.1 Words and Word Classes	91
4.3.2 Phrase Level: The Case of Filler Phrases and Hedges	94
4.4 Disfluencies	96
4.5 Non-Verbal, Vocal Events	98
4.6 Common Traits of Formal Aspects	100
References	101
5 Functional Aspects	107
5.1 Biological Trait Primitives	109
5.1.1 Speaker Characteristics	111
5.2 Cultural Trait Primitives	112
5.2.1 Speech Characteristics	114
5.3 Personality	115
5.4 Emotion and Affect	119
5.5 Subjectivity and Sentiment Analysis	123
5.6 Deviant Speech	124
5.6.1 Pathological Speech	125
5.6.2 Temporarily Deviant Speech	129
5.6.3 Non-native Speech	130
5.7 Social Signals	131
5.8 Discrepant Communication	135
5.8.1 Indirect Speech, Irony, and Sarcasm	136
5.8.2 Deceptive Speech	138
5.8.3 Off-Talk	139
5.9 Common Traits of Functional Aspects	140
References	141
6 Corpus Engineering	159
6.1 Annotation	160
6.1.1 Assessment of Annotations	161
6.1.2 New Trends	164
6.2 Corpora and Benchmarks: Some Examples	164
6.2.1 FAU Aibo Emotion Corpus	165
6.2.2 aGender Corpus	165
6.2.3 TUM AVIC Corpus	166
6.2.4 Alcohol Language Corpus	168
6.2.5 Sleepy Language Corpus	168
6.2.6 Speaker Personality Corpus	169
6.2.7 Speaker Likability Database	170
6.2.8 NKI CCRT Speech Corpus	171
6.2.9 TIMIT Database	171
6.2.10 Final Remarks on Databases	172
References	173
Part II Modelling	7
Computational Modelling of Paralinguistics: Overview	179
References	183
8 Acoustic Features	185
8.1 Digital Signal Representation	185
8.2 Short Time Analysis	187
8.3 Acoustic Segmentation	190
8.4 Continuous Descriptors	190
8.4.1 Intensity	190
8.4.2 Zero Crossings	191
8.4.3 Autocorrelation	192
8.4.4 Spectrum and Cepstrum	194
8.4.5 Linear Prediction	198
8.4.6 Line Spectral Pairs	202
8.4.7 Perceptual Linear Prediction	203
8.4.8 Formants	205
8.4.9 Fundamental Frequency and Voicing Probability	207
8.4.10 Jitter and Shimmer	212
8.4.11 Derived Low-Level Descriptors	214
References	214
9 Linguistic Features	217
9.1 Textual Descriptors	217
9.2 Preprocessing	218
9.3 Reduction	218
9.3.1 Stopping	218
9.3.2 Stemming	219
9.3.3 Tagging	219
9.4 Modelling	220
9.4.1 Vector Space Modelling	220
9.4.2 On-line Knowledge	222
References	227
10 Supra-segmental Features	230
10.1 Functionals	231
10.2 Feature Brute-Forcing	232
10.3 Feature Stacking	233
References	234
11 Machine-Based Modelling	235
11.1 Feature Relevance Analysis	235
11.2 Machine Learning	238
11.2.1 Static Classification	238
11.2.2 Dynamic Classification: Hidden Markov Models	256
11.2.3 Regression	262
11.3 Testing Protocols	264
11.3.1 Partitioning	264
11.3.2 Balancing	266
11.3.3 Performance Measures	267
11.3.4 Result Interpretation	272
References	277
12 System Integration and Application	281
12.1 Distributed Processing	281
12.2	

Autonomous and Collaborative Learning 284 12.3 Confidence Measures 286 References 287 13 'Hands-On': Existing Toolkits and Practical Tutorial 289 13.1 Related Toolkits 289 13.2 openSMILE 290 13.2.1 Available Feature Extractors 293 13.3 Practical Computational Paralinguistics How-to 294 13.3.1 Obtaining and Installing openSMILE 295 13.3.2 Extracting Features 295 13.3.3 Classification and Regression 302 References 303 14 Epilogue 304 Appendix 307 A.1 openSMILE Feature Sets Used at Interspeech Challenges 307 A.2 Feature Encoding Scheme 310 References 314 Index 315

Sommario/riassunto

"This book is a guide through the contemporary field of automatically detecting speaker states/traits in speech via acoustic and linguistic properties. The authors will first introduce the general topic covering definitions, usability and application, and then discuss the psychological underpinnings of emotions, affect and personality and how they are expressed and categorized in speech. Reflecting the multidisciplinary character of the field, the authors switch to aspects of human speech and language containing speech production and perception, and linguistic and paralinguistic aspects. The authors will also focus on the signal processing and machine learning aspects of the actual computational modelling of emotion and personality and will explain the detection process from corpus collection through feature extraction and model testing to system integration. After a general introduction into computational modelling of emotion and personality including pre-processing, feature extraction and machine learning algorithms, acoustic and linguistic analyses will each be handled in separate chapters. Once emotion and personality have been recognised by a technical system, the question arises how to best integrate this information in a system context, in particular dealing with uncertainty - an aspect often handled with lower attention, neglecting its high importance. The authors will cover this providing an extra chapter on aspects in this context as standards for emotion and personality, dealing with error-prone prediction results, real-time issues, application design, and real-life evaluation of systems. The book will end with a tutorial enabling the reader to build an emotion detection model on an existing corpus. This hands-on approach by integrating actual data sets, software, and open-source utilities will make the book invaluable as a teaching tool and similarly useful for those professionals already in the field"--

"In this book, we will focus on analysis, basically excluding generation and synthesis"--
