

# Springer Complexity

---

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems—cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse “real-life” situations like the climate, the coherent emission of light from lasers, chemical reaction–diffusion systems, biological cellular networks, the dynamics of stock markets and of the internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series “Understanding Complex Systems” focusing on the various applications of complexity, and the “Springer Series in Synergetics”, which is devoted to the quantitative theoretical and methodological foundations, and the “SpringerBriefs in Complexity” which are concise and topical working reports, case-studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

## Editorial and Programme Advisory Board

Henry Abarbanel, Institute for Nonlinear Science, University of California, San Diego, USA

Dan Braha, New England Complex Systems, Institute and University of Massachusetts, Dartmouth, USA

Péter Érdi, Center for Complex Systems Studies, Kalamazoo College, USA and Hungarian Academy of Sciences, Budapest, Hungary

Karl Friston, Institute of Cognitive Neuroscience, University College London, London, UK

Hermann Haken, Center of Synergetics, University of Stuttgart, Stuttgart, Germany

Viktor Jirsa, Centre National de la Recherche Scientifique (CNRS), Université de la Méditerranée, Marseille, France

Janusz Kacprzyk, System Research, Polish Academy of Sciences, Warsaw, Poland

Kunihiko Kaneko, Research Center for Complex Systems Biology, The University of Tokyo, Tokyo, Japan

Scott Kelso, Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, USA

Markus Kirkilionis, Mathematics Institute and Centre for Complex Systems, University of Warwick, Coventry, UK

Jürgen Kurths, Nonlinear Dynamics Group, University of Potsdam, Potsdam, Germany

Andrzej Nowak, Department of Psychology, Warsaw University, Poland

Hassan Qudrat-Ullah, School of Administrative Studies, York University, Canada

Linda Reichl, Center for Complex Quantum Systems, University of Texas, Austin, USA

Peter Schuster, Theoretical Chemistry and Structural Biology, University of Vienna, Vienna, Austria

Frank Schweitzer, System Design, ETH Zürich, Zürich, Switzerland

Didier Sornette, Entrepreneurial Risk, ETH Zürich, Zürich, Switzerland

Stefan Thurner, Section for Science of Complex Systems, Medical University of Vienna, Vienna, Austria

# Understanding Complex Systems

---

**Founding Editor: Scott Kelso**

Future scientific and technological developments in many fields will necessarily depend upon coming to grips with complex systems. Such systems are complex in both their composition—typically many different kinds of components interacting simultaneously and nonlinearly with each other and their environments on multiple levels—and in the rich diversity of behavior of which they are capable.

The Springer Series in Understanding Complex Systems series (UCS) promotes new strategies and paradigms for understanding and realizing applications of complex systems research in a wide variety of fields and endeavors. UCS is explicitly transdisciplinary. It has three main goals: First, to elaborate the concepts, methods and tools of complex systems at all levels of description and in all scientific fields, especially newly emerging areas within the life, social, behavioral, economic, neuro- and cognitive sciences (and derivatives thereof); second, to encourage novel applications of these ideas in various fields of engineering and computation such as robotics, nano-technology and informatics; third, to provide a single forum within which commonalities and differences in the workings of complex systems may be discerned, hence leading to deeper insight and understanding.

UCS will publish monographs, lecture notes and selected edited contributions aimed at communicating new findings to a large multidisciplinary audience.

More information about this series at <http://www.springer.com/series/5394>

Alexander Mehler · Andy Lücking  
Sven Banisch · Philippe Blanchard  
Barbara Frank-Job  
Editors

# Towards a Theoretical Framework for Analyzing Complex Linguistic Networks

 Springer

*Editors*

Alexander Mehler  
Goethe-University Frankfurt am Main  
Department of Computer Science  
and Mathematics  
Frankfurt am Main  
Germany

Philippe Blanchard  
Department of Physics  
University of Bielefeld  
Bielefeld  
Germany

Andy Lücking  
Goethe-University Frankfurt am Main  
Department of Computer Science  
and Mathematics  
Frankfurt am Main  
Germany

Barbara Frank-Job  
Faculty of Linguistics & Literary Studies  
University of Bielefeld  
Bielefeld  
Germany

Sven Banisch  
Max Planck Institute for Mathematics  
in the Sciences  
Inselstrasse 22  
D-04103 Leipzig  
Germany

ISSN 1860-0832

ISSN 1860-0840 (electronic)

Understanding Complex Systems

ISBN 978-3-662-47237-8

ISBN 978-3-662-47238-5 (eBook)

DOI 10.1007/978-3-662-47238-5

Library of Congress Control Number: 2015940024

Springer Heidelberg New York Dordrecht London

© Springer-Verlag Berlin Heidelberg 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer-Verlag GmbH Berlin Heidelberg is part of Springer Science+Business Media  
([www.springer.com](http://www.springer.com))

# Introduction

Alexander Mehler, Andy Lücking, Sven Banisch,  
Philippe Blanchard, and Barbara Frank-Job

## 1 On the Content of This Book

Currently, we observe an advent of approaches to analyzing linguistic networks with the methods of stochastic physics and graph theory. Generally speaking, a linguistic network is represented by a graph whose vertices denote linguistic units (e.g., words, sentences, or textual units) and whose edges model linguistic (e.g. syntactic, semantic or pragmatic) relations of these units. The aim of models operating on such networks is to capture the synchronic, topological or evolutionary dynamics of linguistic systems, say, on the phonological, morphological, syntactic, semantic or pragmatic level. What these approaches have in common is that they model the structural or temporal dynamics of linguistic systems in order to test information-theoretical or linguistic hypotheses on the grounds of complex network theory. This is partly done in terms of a *strong network perspective* according to which the network approach is seen to be indispensable to test the focal hypotheses. Apparently, the area of language evolution provides a good test case for such an approach.

---

Alexander Mehler · Andy Lücking

Goethe-University Frankfurt am Main, Department of Computer Science and Mathematics,  
Frankfurt am Main, Germany

e-mail: {Mehler, Luecking}@em.uni-frankfurt.de

Sven Banisch

Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22,  
D-04103 Leipzig, Germany

e-mail: Sven.Banisch@mis.mpg.de

Philippe Blanchard

Faculty of Physics, Bielefeld University, Germany

e-mail: Blanchard@physik.uni-bielefeld.de

Barbara Frank-Job

Faculty of Linguistics & Literary Studies, University of Bielefeld, Bielefeld, Germany

e-mail: Barbara.Job@uni-bielefeld.de

Language evolution can be seen as a meso system that connects language as a macro system with the micro system of cognitive processes of language processing. Starting from such a unified approach to language structure, language change and processing, network approaches try to gain insights into the laws of linguistic information processing in communities of social agents.

In spite of the remarkable success regarding the development of expressive graph models of linguistic systems, these approaches are still in need of a unifying framework. To date, the models are connected by a common methodical stance based on complex network theory in addition to quantitative linguistics. Thus, we face a range of diverse network models that focus on laws of information processing without clarifying their synergetic interdependencies. This is partly due to the lack of shared standards of data modeling, of the interoperability of algorithmic graph models and of the sustainability of the underlying linguistic resources and corpora. Obviously, interdisciplinary research across the boarder of computer science, linguistics and stochastic physics may profit from the availability of such standards.

This book aims at making first steps into the direction of filling this gap. It presents theoretical and empirical results in support of a unifying approach to linguistic networks that may help to overcome bottleneck problems of this field of research. To this end, the book comprises recent research efforts in the area of linguistic networks. It brings together scientists with diverse backgrounds ranging from linguistics to text-technology, from computational humanities to statistical network theory. The book is organized, roughly, into six parts including semantic and syntactic networks, the interplay of language and cognition, the simulation of socio-linguistic dynamics and text-technological resources of network modeling. Special emphasis is put on critical articles and articles that review recent developments in the field. This includes the following fields of research:

- Resources of linguistic network analysis.
- Principles of linguistic network induction.
- Topological models of language structure.
- Models of language dynamics: evolution, diachrony, change.
- Unified models from stochastic physics.
- Network models from cognitive linguistics.
- Network models of phonological, lexical, syntactic, semantic or pragmatic systems.
- Network models of text systems in contrast to language systems.

Dealing with these and related topics, the aim of the book is to advocate and promote network models of linguistic systems that are both based on thorough mathematical models and substantiated in terms of linguistic interpretations. In this way, the book contributes first steps towards establishing a statistical network theory as a theoretical basis of linguistic network analysis across the boarder of the natural sciences and the humanities.

## 2 Overview of the Book

### 2.1 *Part I: Cognition*

Successful applications of network analysis with a particular focus on the interplay of language and cognition are reviewed in the chapter of Beckage and Colunga. Concentrating on semantic and phonological networks, it explores network features and their relation to human language performance including the application to cognitive impairment and atypical behavior.

The chapter by Vitevitch, Goldstein and Johnson combines network tools and data from a psycholinguistic experiment to explore speech perception errors with the aim to understand better what is perceived when a spoken word is misperceived. The experimental results of their phonological association task are evaluated in terms of path' on a network of phonological similarity.

The chapter by De Deyne, Verheyen and Storms compares semantic networks derived from text corpora with networks obtained through word association experiments by looking at macro- and mesoscopic properties of both types of graphs. While the analysis reveals structural similarities at the global level, significant differences between text and word association graphs emerge at a lower level of community structure or centrality. The chapter also presents a comparison with human relatedness judgments.

### 2.2 *Part II: Topology*

The chapter by Biemann, Krumov, Roos and Weihe presents a statistical analysis of the motif signatures of co-occurrence graphs including co-authorship networks, communication networks and linguistic co-occurrence graphs of natural and artificial languages. Based on the hypothesis that different word classes serve different functions in a language an analysis of co-occurrence graphs for different word classes (verbs vs. nouns vs. adjectives etc.) is performed which shows that especially verbs are distinguishable from other word classes by their motif signature – across different languages.

The chapter by Araújo and Banisch highlights the need to consider different ways of network induction in network-based analysis of language and reasons that induction and analysis are strongly interdependent tasks. Based on a framework comprising different abstraction levels along with levels of statistical analysis, the authors argue that the field of linguistic networks is challenged by the fact that an interpretation of topological indicators used in network analysis becomes the harder, the higher the abstraction level of the network.

The chapter by Masucci, Kalampokis, Eguíluz and Hernández-García presents an information-theoretic approach to derive a directed network of semantic flow between Wikipedia articles using a complete snapshot of the English Wikipedia. The authors show that the resulting semantic space is characterized by a scale-free behavior at different scales which implies a hierarchical organization of semantic spaces.

The chapter by Zweig confronts the physically-inspired context-free quest for universal structures with the need of contextual interpretations in sociology and in linguistics. Zweig questions the usefulness of network representations of word-adjacency relations, because most of the well-known topological indicators rely on a rather specific network process and they may therefore be misleading if this process is not known or not adequately modeled by the process underlying the method.

### **2.3 *Part III: Syntax***

The chapter by Čech, Mačutek and Liu presents a critical review of the application of complex network tools to the analysis of syntax and points out the main challenges for further research. Among many other things, the article discusses the impact of syntax on network properties, the preprocessing of data, and the application of network studies to language typology and acquisition.

A second chapter dealing with syntactic dependency networks is by Chen and Liu. Based on two syntactic dependency networks from different genres this chapter analyses the syntactic status of three function words in Chinese. The importance (the authors propose the notion of syntactic centrality) of the words is analyzed by independently removing them from the network and comparing their statistical characteristics before and after removal.

The chapter by Ferrer i Cancho challenges the existing theory of syntax by confronting the observation that syntactic dependencies between the words of a sentence rarely cross when drawn over a sentence with two null hypotheses for the expected number of crossings by chance. Relying on the trade-off between parsimony and explanatory power, the chapter argues that the minimization of syntactic dependency length (as a principle that derives from limited computational resources of the brain) can explain uncrossing dependencies and that this explanation is, from an economic point of view, preferable over explanations relying on grammar.

### **2.4 *Part IV: Dynamics***

The role of cultural transmission in language change across three generations is analyzed on the basis of an extended simulation model by Gong and Shuai. While transmission within the offspring generation and between the offspring and the parent generation fosters language change and leads, at the same time, to mutual understandability within generations and across consecutive generations, interaction between children and their grandparent's generation plays an important role in preserving mutual cross-generational understandability in the long run.

Another simulation study is presented by Baxter who complements his numerical results with analytical arguments. Drawing on an evolutionary approach to language change, the author looks in detail to the convergence behavior of the model on different social networks and with heterogeneous patterns of mutual influence that, taken together, may encode a variety of social structures.



The chapter by Maity and Mukherjee presents a simulation study of the effect of inflexible individuals on the dynamics of the naming game and shows that rigid minorities lead to the emergence of dominant states in the population. The model is analyzed on a series of static networks of different complexity ranging from the complete graph to scale-free topologies and a dynamic network obtained from real-world time-varying face-to-face interaction data is also considered.

## **2.5 *Part V: Resources***

The requirements of a data format applicable to the wide range of linguistic network data are discussed in the chapter by Stührenberg, Diewald and Gleim. The authors analyze various existing graph formats in relation to their expressivity and support by common tools for network analysis and propose an extension of GraphML as a possibly complex data model of a graph which allows to quickly extract views for specific tasks, rather than extracting incoherent different views from raw data. It is noteworthy, that this chapter grew out of a working group that was constituted at the MLN conference.

The book concludes with the chapter by Mehler and Gleim who present the LN system, an online platform for the automatic generation of lexical networks from texts. It addresses two communities: on the one hand humanities scholars (e.g., historical semanticists) who aim at studying the change of language use as an indicator of social-semantic change. On the other hand, network theorists who are in need of null models for making linguistic networks comparable. The workflow of the LN system – using GraphML as an output standard for linguistic networks – is explained and exemplified.

# Acknowledgements

The book collects selected and extended contributions to the workshop on *Modeling Linguistic Networks: from Language Structures to Communication Processes* held at Goethe University Frankfurt, December 10-11, 2012, as part of the research project *Linguistic Networks: Text Technological Representation, Computational Linguistic Synthesis and Physical Modeling* funded by the German Federal Ministry of Education (BMBF). Financial support by the BMBF is gratefully acknowledged.

# Contents

<b>Introduction</b> .....	<b>V</b>
<i>Alexander Mehler, Andy Lücking, Sven Banisch, Philippe Blanchard, Barbara Frank-Job</i>	
1 On the Content of This Book .....	V
2 Overview of the Book .....	VII
2.1 Part I: Cognition .....	VII
2.2 Part II: Topology .....	VII
2.3 Part III: Syntax .....	VIII
2.4 Part IV: Dynamics .....	VIII
2.5 Part V: Resources .....	IX
<b>Part I: Cognition</b>	
<b>Language Networks as Models of Cognition: Understanding Cognition through Language</b> .....	<b>3</b>
<i>Nicole M. Beckage, Eliana Colunga</i>	
1 Introduction .....	3
2 Language as a Network .....	5
2.1 Semantic Networks .....	5
2.2 Phonological Networks .....	6
3 Global Level Network Structure .....	7
3.1 Small-World Structure .....	8
3.2 Scale-Free Networks .....	10
4 Human Performance in Relation to Network Structure .....	11
4.1 Spreading Activation .....	11
4.2 Frequency Effects .....	15
5 Network Models within Linguistic Networks .....	16
5.1 Acquisition .....	17
5.2 Network Navigation .....	21
6 Understanding Atypical Processes .....	23
7 The Future of Language Networks .....	25
References .....	26

**Path-Length and the Misperception of Speech: Insights from Network Science and Psycholinguistics . . . . . 29**

*Michael S. Vitevitch, Rutherford Goldstein, Elizabeth Johnson*

1	Introduction . . . . .	29
2	Network Analysis: What Can Be Perceived When Speech Is Misperceived? . . . . .	31
3	Psycholinguistic Experiment: What Is Perceived When Speech Is Misperceived? . . . . .	34
3.1	Method . . . . .	35
3.2	Results . . . . .	37
4	Conclusion . . . . .	40
	References . . . . .	43

**Structure and Organization of the Mental Lexicon: A Network Approach Derived from Syntactic Dependency Relations and Word Associations . . . . . 47**

*Simon De Deyne, Steven Verheyen, Gert Storms*

1	Introduction . . . . .	47
1.1	Macro-, Meso-, and Microscopic Properties of the Mental Lexicon . . . . .	48
1.2	Acquiring a Mental Lexicon through Language . . . . .	50
1.3	Chapter Outline . . . . .	51
2	Constructing the Networks . . . . .	53
2.1	Mental Networks . . . . .	53
2.2	Language Networks . . . . .	54
3	Exploring the Structure of Language and Mental Networks . . . . .	56
3.1	Macroscopic Structure . . . . .	56
3.2	Mesoscopic Structure . . . . .	59
3.3	Semantic Relatedness Evaluation . . . . .	66
4	Discussion . . . . .	70
4.1	Relationship between Language and Word Associations . . . . .	72
4.2	Final Words . . . . .	73
	References . . . . .	74

**Part II: Topology**

**Network Motifs Are a Powerful Tool for Semantic Distinction . . . . . 83**

*Chris Biemann, Lachezar Krumov, Stefanie Roos, Karsten Weihe*

1	Introduction . . . . .	84
2	Related Work . . . . .	86
3	The Case Studies . . . . .	87
3.1	Co-occurrence Graphs from Natural Vs. Artificial Language . . . . .	87
3.2	Co-occurrence Graphs from Verbs Vs. Other Word Classes . . . . .	94

3.3	Peer-to-Peer Streaming Networks . . . . .	99
3.4	Co-Authorship Networks from Two Subdisciplines of Physics . . . . .	100
3.5	Mailing Networks . . . . .	102
4	Conclusions and Outlook . . . . .	103
	References . . . . .	103
<b>Multidimensional Analysis of Linguistic Networks . . . . .</b>		<b>107</b>
<i>Tanya Araújo, Sven Banisch</i>		
1	Introduction . . . . .	107
2	Linguistic Networks Are Special . . . . .	109
2.1	Three Types of Networks . . . . .	109
2.2	Network Induction . . . . .	112
3	Three Levels of Statistical Analysis . . . . .	114
3.1	A Brief Note on Signal Processing on Graphs . . . . .	115
3.2	The Statistical Levels . . . . .	115
3.3	Stylized Facts in Network Analysis . . . . .	116
3.4	Levels in the Statistical Analysis of Networks . . . . .	118
4	On the Intelligibility of Statistical Indicators in Linguistic Networks . . . . .	120
4.1	Path-Based Measures . . . . .	120
4.2	Links and Flows, Structure and Function . . . . .	121
4.3	Types of Network Flow . . . . .	122
4.4	Flow in Linguistic Networks . . . . .	122
5	Examples . . . . .	124
6	Discussion . . . . .	124
7	Concluding Remarks . . . . .	126
	References . . . . .	127
<b>Semantic Space as a Metapopulation System: Modelling the Wikipedia Information Flow Network . . . . .</b>		<b>133</b>
<i>A. Paolo Masucci, Alkiviadis Kalampokis, Victor M. Eguíluz, Emilio Hernández-García</i>		
1	Introduction . . . . .	133
2	The Dataset . . . . .	136
3	Topology of the Semantic Space . . . . .	136
4	Modelling the Semantic Space . . . . .	141
5	Discussion . . . . .	143
	Appendix . . . . .	145
	References . . . . .	148
<b>Are Word-Adjacency Networks Networks? . . . . .</b>		<b>153</b>
<i>Katharina Anna Zweig</i>		
1	Introduction . . . . .	153
1.1	Perspectives of Network Analysis . . . . .	154

2	De nitions .....	156
2.1	De nition of Word-Adjacency Networks .....	156
3	Walk-Based Methods and Network Flows .....	157
3.1	Models of Walks .....	159
4	Word-Adjacency Networks in the Literature .....	160
5	Summary .....	162
	References .....	163
 <b>Part III: Syntax</b>		
<b>Syntactic Complex Networks and Their Applications .....</b>		<b>167</b>
<i>Radek Čech, Ján Mačutek, Haitao Liu</i>		
1	Introduction .....	167
2	Basic Characteristics of Syntactic Networks .....	168
3	Early Development of Syntactic Complex Network Analysis ....	169
4	Role of Syntax in Syntactic Dependency Complex Networks ....	172
5	Preprocessing of Data for a Syntactic Complex Network Analysis Pitfalls to be Avoided .....	177
6	Applications of Syntactic Complex Networks to Language Typology and Acquisition .....	179
6.1	Language Typology .....	180
6.2	Language Acquisition .....	181
7	Conclusion .....	182
	References .....	183
 <b>Function Nodes in Chinese Syntactic Networks .....</b>		<b>187</b>
<i>Xinying Chen, Haitao Liu</i>		
1	Introduction .....	187
2	The Chinese Dependency Networks for This Study .....	189
3	Chinese Function Words .....	192
4	Chinese Function Words in the Language Networks .....	193
4.1	Network Properties of Chinese Function Words .....	193
4.2	Network Manipulation .....	196
5	Conclusion .....	198
	References .....	199
 <b>Non-crossing Dependencies: Least Effort, Not Grammar .....</b>		<b>203</b>
<i>Ramon Ferrer-i-Cancho</i>		
1	Introduction .....	203
2	The Syntactic Dependency Structure of Sentences .....	207
3	The Null Hypothesis .....	208
4	Alternative Hypotheses .....	212
4.1	A Principle of Minimization of Dependency Crossings .....	213
4.2	A Principle of Minimization of Dependency Lengths ...	214

4.3	The Relationship between Minimization of Crossings and Minimization of Dependency Lengths . . . . .	216
5	A Stronger Null Hypothesis . . . . .	219
5.1	The Probability That Two Edges Cross . . . . .	220
5.2	The Expected Number of Edge Crossings . . . . .	221
6	Another Stronger Null Hypothesis . . . . .	224
7	Predictions, Testing and Selection . . . . .	224
8	Discussion . . . . .	227
	Appendix . . . . .	229
	References . . . . .	231

## Part IV: Dynamics

### **Simulating the Effects of Cross-Generational Cultural Transmission on Language Change . . . . . 237**

*Tao Gong, Lan Shuai*

1	Introduction . . . . .	237
2	Modified Acquisition Framework . . . . .	240
3	Simulation Results . . . . .	242
4	Discussions and Conclusions . . . . .	244
	Appendix . . . . .	248
	References . . . . .	254

### **Social Networks and Beyond in Language Change . . . . . 257**

*Gareth J. Baxter*

1	Introduction . . . . .	257
2	Utterance Selection Model of Language Change . . . . .	258
3	Numerical Model . . . . .	260
4	Analysis . . . . .	264
5	Social Networks in the Neutral Model . . . . .	267
6	Weighted Interactor Selection . . . . .	268
6.1	Asymmetry Independent of Network Structure . . . . .	269
6.2	Asymmetry Depends on Speakers Degree . . . . .	272
7	Conclusions . . . . .	274
	Appendix . . . . .	275
	References . . . . .	276

### **Emergence of Dominant Opinions in Presence of Rigid Individuals . . . . . 279**

*Suman Kalyan Maity, Animesh Mukherjee*

1	Introduction . . . . .	279
2	Related Work . . . . .	282
3	The Model Description . . . . .	283
4	Results and Discussion . . . . .	283
4.1	The Mean-Field Case . . . . .	283
4.2	Scale-Free Networks . . . . .	287

5	Time-Varying Networks . . . . .	290
5.1	Dataset Description . . . . .	291
5.2	The Model Adaptation in the Time-Varying Setting . . . .	291
5.3	Results and Discussion . . . . .	292
6	Conclusions and Future Works . . . . .	294
	References . . . . .	294

## Part V: Resources

### **Considerations for a Linguistic Network Markup Language . . . . . 299**

*Maik Stührenberg, Nils Diewald, Rüdiger Gleim*

1	Introduction . . . . .	299
2	Data Formats . . . . .	299
2.1	Data Models . . . . .	300
2.2	Data Structures . . . . .	301
2.3	Data Serialization . . . . .	302
3	Existing Formats . . . . .	304
3.1	GML . . . . .	305
3.2	XGMML . . . . .	306
3.3	GraphXML . . . . .	307
3.4	GraphML . . . . .	309
3.5	GXL . . . . .	311
3.6	GrAF . . . . .	313
3.7	Summary . . . . .	315
4	Network Tools . . . . .	315
5	Proposal for a Linguistic Network Markup Language . . . . .	318
5.1	Extending GraphML by Redefinition . . . . .	320
5.2	Extending GraphML by XML Namespaces . . . . .	322
5.3	Example Instance . . . . .	325
6	Conclusion . . . . .	327
	References . . . . .	327

### **Linguistic Networks – An Online Platform for Deriving Collocation**

### **Networks from Natural Language Texts . . . . . 331**

*Alexander Mehler, Rüdiger Gleim*

1	Introduction . . . . .	331
2	On the Parameter Space of LN . . . . .	334
3	The Software Architecture of LN . . . . .	336
4	Summary . . . . .	340
	References . . . . .	340

### **Author Index . . . . . 343**



# List of Contributors

Tanya Araújo

ISEG (Lisboa School of Economics and Management) of the University of Lisbon  
and Research Unit on Complexity in Economics (UECE), Portugal

e-mail: tanya@iseg.utl.pt

Sven Banisch

Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

e-mail: Sven.Banisch@mis.mpg.de

Gareth Baxter

Department of Physics, University of Aveiro, Portugal

e-mail: gjbaxter@ua.pt

Nicole Beckage

University of Colorado Boulder, USA

e-mail: nicole.beckage@colorado.edu

Chris Biemann

Computer Science Department, Technische Universität Darmstadt, Germany

e-mail: biem@cs.tu-darmstadt.de

Radek Čech

Department of Czech Language, Faculty of Arts, University of Ostrava,  
Czech Republic

e-mail: cechradek@gmail.com

Xinying Chen

School of International Study, Xi'an Jiaotong University, Xi'an, China

e-mail: chenxinying@mail.xjtu.edu.cn

Eliana Colunga

University of Colorado Boulder, USA

e-mail: eliana.colunga@colorado.edu

Simon De Deyne

University of Adelaide, Australia

e-mail: [simon.dedeyne@adelaide.edu.au](mailto:simon.dedeyne@adelaide.edu.au)

Nils Diewald

Institut für Deutsche Sprache, Mannheim, Germany

e-mail: [diewald@ids-mannheim.de](mailto:diewald@ids-mannheim.de)

Víctor M. Eguíluz

Instituto de Física Interdisciplinar y Sistemas Complejos IFISC,  
Palma de Mallorca, Spain

e-mail: [victor@ifisc.uib-csic.es](mailto:victor@ifisc.uib-csic.es)

Ramon Ferrer i Cancho

Complexity and Quantitative Linguistics Lab, LARCA Research Group,  
Departament de Ciències de la Computació, Universitat Politècnica de Catalunya  
(UPC), Barcelona, Catalonia, Spain

e-mail: [rferrericancho@lsi.upc.edu](mailto:rferrericancho@lsi.upc.edu)

Rüdiger Gleim

Text Technology Lab, Department of Computer Science and Mathematics,  
Goethe-University Frankfurt, Frankfurt am Main, Germany

e-mail: [gleim@em.uni-frankfurt.de](mailto:gleim@em.uni-frankfurt.de)

Rutherford Goldstein

University of Kansas, Lawrence, KS 66045 USA

e-mail: [rgolds@ku.edu](mailto:rgolds@ku.edu)

Tao Gong

Haskins Laboratories, New Haven, CT, USA

e-mail: [gtojty@gmail.com](mailto:gtojty@gmail.com)

Emilio Hernández-García

Instituto de Física Interdisciplinar y Sistemas Complejos IFISC,  
Palma de Mallorca, Spain

e-mail: [emilio@ifisc.uib-csic.es](mailto:emilio@ifisc.uib-csic.es)

Elizabeth Johnson

University of Kansas, Lawrence, KS 66045 USA

e-mail: [libbyjohnson54@gmail.com](mailto:libbyjohnson54@gmail.com)

Alkiviadis Kalampokis

Dipartimento di Scienze e Tecnologie, Università degli Studi di Napoli  
“Parthenope”, Centro Direzionale di Napoli, Napoli, Italy

e-mail: [alkis@marine.aegean.gr](mailto:alkis@marine.aegean.gr)

Lachezar Krumov

Computer Science Department, Technische Universität Darmstadt, Germany

e-mail: [lakkrumov17@yahoo.com](mailto:lakkrumov17@yahoo.com)

Haitao Liu

Department of Linguistics, Zhejiang University, Hangzhou, CN-310058, China  
e-mail: lhtzju@gmail.com

Ján Mačutek

Department of Applied Mathematics and Statistics, Comenius University,  
Mlynská dolina, Bratislava, Slovakia  
e-mail: jmacutek@yahoo.com

Suman Kalyan Maity

Dept. of Computer Science and Engineering, Indian Institute of Technology,  
Kharagpur, India  
e-mail: sumankalyan.maity@cse.iitkgp.ernet.in

Paolo Masucci

Centre for Advanced Spatial Analysis, University College of London,  
London, UK  
e-mail: paolo\_masucci@yahoo.it

Alexander Mehler

Text Technology Lab, Department of Computer Science and Mathematics,  
Goethe-University Frankfurt, Frankfurt am Main, Germany  
e-mail: mehler@em.uni-frankfurt.de

Animesh Mukherjee

Dept. of Computer Science and Engineering, Indian Institute of Technology,  
Kharagpur, India  
e-mail: animeshm@cse.iitkgp.ernet.in

Stefanie Roos

Computer Science Department, Technische Universität Darmstadt, Germany  
e-mail: stefanie.roos@tu-dresden.de

Lan Shuai

Haskins Laboratories, New Haven, CT, USA  
e-mail: susan.shuai@gmail.com

Gert Storms

University of Leuven, Belgium  
e-mail: gert.storms@ppw.kuleuven.be

Maik Stührenberg

Institut für Deutsche Sprache, Mannheim, Germany  
e-mail: stuehrenberg@ids-mannheim.de

Steven Verheyen

University of Leuven, Belgium  
e-mail: steven.verheyen@ppw.kuleuven.be

Michael Vitevich  
University of Kansas, Lawrence, KS 66045 USA  
e-mail: mvitevit@ku.edu

Karsten Weihe  
Computer Science Department, Technische Universität Darmstadt, Germany  
e-mail: weihe@cs.tu-darmstadt.de

Katharina Zweig  
Department of Computer Science, TU Kaiserslautern, Germany  
e-mail: zweig@cs.uni-kl.de