

# **Structure, Function and Process in Texts**

Edited by

Lu Wang, Reinhard Köhler, Arjuna Tuzzi.

**RAM-Verlag 2018**

**Wang, Lu**  
**Köhler, Reinhard**  
**Tuzzi, Arjuna**

# **Structure, Function and Process in Texts**

© Copyright 2018 by RAM-Verlag

Publisher: RAM-Verlag  
Stüttinghauser Ringstr. 44  
D-58515 Lüdenscheid  
Germany  
[RAM-Verlag@t-online.de](mailto:RAM-Verlag@t-online.de)  
<http://ram-verlag.eu>

The publisher cannot be held responsible for any linguistic errors in the book:  
Such responsibility is only up to the authors.

**ISBN: 978-3-942303-56-9**

# Contents

## Preface

Michele A. Cortelazzo, Paolo Nadalutti, Stefano Ondelli, Arjuna Tuzzi <b>Authorship Attribution and Text Clustering in Contemporary Italian Novels.....</b>	<b>1</b>
Martina Benešová, Dan Faltýnek, Lukáš Zámečník <b>Functional Explanation in Synergetic Linguistics.....</b>	<b>15</b>
Sheila Embleton, Dorin Uritescu, Eric S. Wheeler <b>An Expanded Quantitative Study of Linguistic vs. Geographic Distance Using Romanian Dialect Data.....</b>	<b>25</b>
Adam Pawłowski, Krzysztof Topolski, Piotr Malak, Jan Kocoń, Michał Marcińczuk <b>Statistical Distributions of Parts of Speech Frequencies in Polish. Big Data Analysis.....</b>	<b>34</b>
Lu Wang, Yahui Guo <b>Polyfunctionality Studies in German, Dutch, English and Chinese.....</b>	<b>55</b>
Relja Vulcanović, Oliver Ruff <b>Measuring the Degree of Violation of the One-Meaning–One-Form Principle.....</b>	<b>67</b>
Haruko Sanada <b>Quantitative Interrelations of Properties of Complement and Adjunct..</b>	<b>78</b>
Eduard Klyshinsky, Varvara Logacheva, Joakim Nivre <b>Multilingual Quantitative Analysis of Morphological Ambiguity.....</b>	<b>100</b>

Hao Sun, Mingzhe Jin <b>Collaborative Writing of Otome no minato.....</b>	<b>116</b>
Alfiya Galieva, Olga Nevzorova <b>Corpus-Based Assessment of Linguistic Complexity of the Tatar Language: Methodology and Preliminary Results.....</b>	<b>128</b>
Reinhard Köhler <b>QUALICO 2016 in Trier. Conference Report.....</b>	<b>142</b>

## Preface

Founded in 1994, the International Quantitative Linguistics Association (IQLA) aims at promoting the development of theoretical, empirical, and applied Quantitative Linguistics (QL) and fostering an effective communication among scientists that work in this highly interdisciplinary field. This multidisciplinary background enables us to share methods and findings across the boundaries of the individual sub-disciplines, languages, and methodological areas. Quantitative linguistics cannot only be characterised as a strongly co-operative science but this property can be considered as an essential principle.

IQLA organizes on a regular basis the QUAntitative LInguistics CONference (QUALICO) that represents the most important meeting for IQLA members and for QL scholars. The first eight conferences in this series took place in: Trier (1991), Moscow (1994), Helsinki (1997), Prague (2000), Athens, Georgia (2003), Graz (2009), Belgrade (2012) and Olomouc (2014). The 9th meeting, Qualico2016, took again place in Trier. The scientific program was organized over 3 days in August 24-28 in two parallel sessions. The lecture halls were rented in the Deutsche Richterakademie (German Academy of European Law).

This book includes a selection of the papers presented at Qualico2016 from various areas of QL. The contributions highlight the richness of this branch of research and the wide range of different topics, methods and approaches available in QL.

A substantial amount of contributions are focussed on research in morphology. *Klyshinsky, Logacheva and Nivre* study morphological ambiguity, aiming at assisting natural language processing, by defining six types of word ambiguity and computing their distributions in nine languages. In order to measure grammatical ambiguity, *Wang and Guo* introduce polyfunctionality, which refers to the number of parts of speech a word can be attributed to. They suggest that a universal model can capture the distributions on data from German, Dutch, English and Chinese. *Galieva and Nevzorova* report an attempt to assess the morphological complexity of the Tatar language. *Pawłowski, Topolski, Malak, Kocoń and Marcińczuk* compare the differences between the distributions of vocabulary in the entire corpus and in its subcorpora (consisting of individual parts of speech), on data from Polish and explain the reasons. *Vulanović and Ruff* present a formula, based on set theory, to measure how much a linguistic system violates the One-Meaning-One-Form principle.

Two contributions to this volume involve stylometrics. *Cortelazzo, Nadalutti, Ondelli and Tuzzi* work on an Italian case of authorship attribution and thorough methods of

text clustering investigate the distinctive traits of Elena Ferrante's writing style with a specific focus on the presumed similarity with other novels written by Italian authors from Naples and its surroundings. The study proposed by *Sun and Jin* on a Japanese novel *Otome no minato*, which is published under the name of Yasunari Kawabata but disputed to be written by Tsuneko Nakazato. This study identifies the authorship by examining stylometric features such as part-of-speech bi-grams, particle bi-grams and phrase patterns and classifiers such as correspondence analysis, hierarchical cluster analysis, etc.

The contribution by *Embleton, Uritescu and Wheeler* deal with a dialectometric problem. They test the relationship between the linguistic distance and geographic distance on data from a Romanian dialect. Their study involves as an innovative feature, taking the popularity of telecommunication into consideration: The researchers measure geographic distance by both travel distance and travel time.

*Sanada* demonstrates a syntactic research on Japanese complement and adjunct. The quantitative properties of the two sentence components including frequency, length and position are investigated in various aspects.

*Benešová, Faltýnek and Zámečník* propose a discussion on functional explanation in quantitative linguistics, especially in synergetic linguistics. The authors focus on the philosophic reflection on epistemological and methodological bases and suggest a reconstruction of the functional model of explanation in synergetic linguistics.

The conference was evaluated both by the organisers and the participants as another successful event.

The Editors,

Lu Wang, Reinhard Köhler, Arjuna Tuzzi