Issues in Quantitative Linguistics

4

edited by

Emmerich Kelih
Róisín Knight
Ján Mačutek
Andrew Wilson

Dedicated to Reinhard Köhler on the occasion of his 65th birthday

2016

RAM-Verlag

Studies in Quantitative Linguistics

Editors

Fengxiang Fan (fanfengxiang@yahoo.com)
Emmerich Kelih (emmerich.kelih@univie.ac.at)
Reinhard Köhler (koehler@uni-trier.de)
Ján Mačutek (jmacutek@yahoo.com)
Eric S. Wheeler (wheeler@ericwheeler.ca)

1. U. Strauss, F. Fan, G. Altmann, *Problems in quantitative linguistics 1*. 2008, VIII + 134 pp.
2. V. Altmann, G. Altmann, *Anleitung zu quantitativen Textanalysen. Methoden und Anwendungen.* 2008, IV+193 pp.
3. I.-I. Popescu, J. Mačutek, G. Altmann, *Aspects of word frequencies*. 2009, IV +198 pp.
4. R. Köhler, G. Altmann, *Problems in quantitative linguistics 2.* 2009, VII + 142 pp.
5. R. Köhler (ed.), *Issues in Quantitative Linguistics.* 2009, VI + 205 pp.
6. A. Tuzzi, I.-I. Popescu, G.Altmann, *Quantitative aspects of Italian texts*. 2010, IV+161 pp.
7. F. Fan, Y. Deng, *Quantitative linguistic computing with Perl.* 2010, VIII + 205 pp.
8. I.-I. Popescu et al., *Vectors and codes of text*. 2010, III + 162 pp.
9. F. Fan, *Data processing and management for quantitative linguistics with Foxpro***.** 2010, V + 233 pp.
10. I.-I. Popescu, R. Čech, G. Altmann, *The lambda-structure of texts*. 2011, II + 181 pp
11. E. Kelih et al. (eds.), *Issues in Quantitative Linguistics Vol. 2*. 2011, IV + 188 pp.
12. R. Čech, G. Altmann, *Problems in quantitative linguistics 3*. 2011, VI + 168 pp.
13. R. Köhler, G. Altmann (eds.), *Issues in Quantitative Linguistics Vol 3.* 2013, IV + 403 pp.
14. R. Köhler, G. Altmann, *Problems in Quantitative Linguistics Vol. 4.* 2014, VIII+148 pp.
15. Best, K.-H., Kelih, E. (eds.), *Entlehnungen und Fremdwörter: Quantitative Aspekte.* 2014. VI + 163 pp.
16. I.-I. Popescu, K.-H. Best, G. Altmann, G. *Unified Modeling of Length in Language.* 2014, VIII + 123 pp.

17. G. Altmann, R. Čech, J. Mčutek, L. Uhlířová (eds.), *Empirical Approaches to Text and Language Analysis dedicated to Luděk Hřebíček on the occasion of his 80th birthday*. 2014. VI + 231 pp.
18. M. Kubát, V. Matlach, R. Čech, *QUITA Quantitative Index Text Analyzer*. 2014, VII + 106 pp.
19. K.-H. Best, *Studien zur Geschichte der Quantitativen Linguistik.* 2015. III + 158 pp.
20. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, P. Mohanty, E. Kelih, R. Chen, G. Altmann, *Descriptiveness, Activity and Nominality in Formalized Text Sequences.* 2015. IV+120 pp.
21. G. Altmann, *Problems in Quantitative Linguistics Vol. 5*. 2015. III+146 pp.
22. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, R. Chen, G. Altmann, *Positional occurrences in texts: Weighted Consensus Strings.* 2015. II+178 pp.

ISBN: 978-3-942303-44-6

© Copyright 2016 by RAM-Verlag, D-58515 Lüdenscheid

RAM-Verlag
Stüttinghauser Ringstr. 44
D-58515 Lüdenscheid
Germany
RAM-Verlag@t-online.de
http://ram-verlag.eu

Contents

Gabriel Altmann
The first steps 1

Gabriel Altmann
On Köhlerian Motifs 2 - 8

Haitao LIU; Yu FANG
Quantitative Aspects of Hierarchical Motifs 9 - 26

Jiří Milička
Key Length Motifs in Czech and Arabic Texts 27 - 42

Germán Coloma
A Synergetic Regression Model of Language Complexity Trade-Offs 43 - 60

Lu Wang
Synergetic Studies on Chinese Lexical Structure 61 - 81

Haruko Sanada
A Measurement of Parts of Speech in Texts Using the Noun-Based Proportion 82 - 93

Hanna Gnatchuk
Testing Hypotheses on English Compounds 94 - 103

Maria Rukk
Context Specific Distribution of Word Meanings 104 - 112

Hans Goebl, Pavel Smečka
The Quantitative Nature of *Working Maps* (WM) and *Taxatorial Areas* (TA). 113 - 127

Sheila Embleton, Dorin Uritescu and Eric S. Wheeler
Play with the Data! 128 - 134

Michele A. Cortelazzo; Arjuna Tuzzi
The First End-Of-Year Address by the New President of the Italian Republic Sergio Mattarella 135 - 149

Miroslav Kubát; Radek Čech
Thematic Concentration and Vocabulary Richness 150 - 159

Solomija Buk, Andrij Rovenchak
Probing the "Temperature" Approach on Ukrainian Texts: Long-prose Fiction by Ivan Franko 160 - 175

Sergey Andreev
Can Pronouns Change the Dynamic Visualization of the Poetic World? 176 - 182

Fan Fengxiang
A Study on Segmental *TTR*, Word Length and Sentence Length 183 - 195

Xiaxing PAN; Haitao LIU
Statistical Analysis of the Diachronic Development of Terminal Rhyme in Chinese Poetry 196 - 216

Gejza Wimmer, Ján Mačutek
Lexical Text Compactness with Link Length Taken into Account 217 - 227

Andrew Wilson
Continuous Modelling of Verse Lengths in Welsh and Gaelic Metrical Psalmody 228 - 236

Kamil Stachowski
German Loanwords in Polish and Remarks on the Piotrowski-Altmann Law 237 - 259

Emmerich Kelih; Ján Mačutek
Probleme der Modellierung von Lehnbeziehungen (am Beispiel von Serbokroatismen im Slowenischen) 260 - 272

Antoni Lozano; Bernardino Casas; Chris Bentz; Ramon Ferrer-i-Cancho
Fast Calculation of Entropy with Zhang's Estimator 273 - 285

Gabriel Altmann
Der Emeritus 286 - 287

The first steps

To learn quantitative linguistics means to collect the works written by Reinhard Köhler and read them thoroughly. Needless to say, one must already know what other (= normal) linguists know, i.e. definitions, classifications, rules, history, etc., but, if one ventures to take a step further, one necessarily bumps against the door of a world one wanted to avoid. It is surely not the infinite paradise but rather the hell in which Köhler, dressed as Lucifer, leads the innocent linguist through the labyrinth of new concepts, formulas, systems, synergetics, statistics, hypotheses, theories. The world of language begins to change its form and the linguist begins (very) slowly to see that there is more light in this world than outside. Köhler loses slowly his Luciferian shape and at once he seems to be the angel responsible for this world. He leads the linguist along ways that join the individual properties of language. One imagines that one is in phonology – very far from syntax - but the angel shows one that there is a very short way between them, and, what is more, one can express it by a formula. That means one need not walk the way; it is sufficient to think it.

Of course, this is rather an esoteric world that exists also in physics and biology, and there is nobody who could prohibit you from entering it (the only exception is the dean of your faculty or head of your department!), but fortunately there is an angel looking sharply at the dean (or head) and leading you into a world in which human language looks like a self-organized system. You may settle on a concept and the system shows you immediately all the links to other concepts. What is more, it brings you to concepts developed by the "Lucifer" himself. As a (pure) linguist, you never heard of them before.

If you succeed in abandoning this world, you will realize that, in front of the door, the same Lucifer stays and wishes you – smiling politely – good moral conscience and much success in repeating what you must teach the students in order to become an "expert" yourself. Linguistics is not about teaching a language but an immersion into worlds which are abstract and similar to the fifth (or higher) dimension of physics. At each step, you can see the smiling Lucifer who created it himself – perhaps only in order to irritate classical linguists who thought that they knew everything already.

But, if you want to stay in this world, you will soon see that you must learn a lot of mathematics, a lot of philosophy of science, forget grammatical rules and all degenerative drawing of trees, and ask the masked Lucifer for help. He will merely smile and, since he wants to retire this year, he will show you a monumental heap of paper and say: "Read all the papers and books in this heap. You can read them more quickly than I wrote them!"

Nevertheless, we hope that he will make the heap higher.

Gabriel Altmann