

Qualico 2009

Text and Language

**Structures · Functions · Interrelations
Quantitative Perspectives**

Edited by

**Peter Grzybek
Emmerich Kelih
Ján Mačutek**

Advisory Editor

Eric S. Wheeler

**Praesens Verlag
Wien 2010**

Contents

Preface <i>Peter Grzybek, Emmerich Kelih, Ján Mačutek</i>	vii
Quantitative analysis of Keats' style: genre differences <i>Sergej Andreev</i>	1
Word-length-related parameters of text genres in the Ukrainian language. A pilot study <i>Solomija Buk, Olha Humenchyk, Liliya Mal'tseva, Andrij Rovenchak</i>	13
On the quantitative analysis of verb valency in Czech <i>Radek Āech, Ján Mačutek</i>	21
A link between the number of set phrases in a text and the number of described facts <i>Łukasz Dębowski</i>	31
Modeling word length frequencies by the Singh-Poisson distribution <i>Gordana Đuraš, Ernst Stadlober</i>	37
How do I know if I am right? Checking quantitative hypotheses <i>Sheila Embleton, Dorin Uritescu, Eric S. Wheeler</i>	49
Text difficulty and the Arens-Altman law <i>Peter Grzybek</i>	57
Parameter interpretation of the Menzerath law: evidence from Serbian <i>Emmerich Kelih</i>	71
A syntagmatic approach to automatic text classification. Statistical properties of <i>F</i> - and <i>L</i> -motifs as text characteristics <i>Reinhard Köhler, Sven Naumann</i>	81
Probabilistic reading of Zipf <i>Jan Králík</i>	91
Revisiting Tertullian's authorship of the <i>Passio Perpetuae</i> through quantitative analysis <i>Jerónimo Leal, Giulio Maspero</i>	99
Textual typology and interactions between axes of variation <i>Sylvain Loiseau</i>	109

Rank-frequency distributions: a pitfall to be avoided <i>Ján Mačutek</i>	119
Measuring lexical richness and its harmony <i>Gregory Martynenko</i>	125
Measuring semantic relevance of words in synsets <i>Ivan Obradović, Cvetana Krstev, Duško Vitas</i>	133
Distribution of canonical syllable types in Serbian <i>Ivan Obradović, Aljoša Obuljen, Duško Vitas, Cvetana Krstev, Vanja Radulović</i>	145
Statistical reduction of the feature space of text styles <i>Vasilij V. Poddubnyj, Anastasija S. Kravcova</i>	159
Quantitative properties of the Nko writing system <i>Andrij Rovenchak, Valentin Vydrin</i>	171
Distribution of motifs in Japanese texts <i>Haruko Sanada</i>	183
Quantitative data processing in the ORD speech corpus of Russian everyday communication <i>Tatiana Sherstinova</i>	195
Complex investigation of texts with the system “StyleAnalyzer” <i>O.G. Shevelyov, V.V. Poddubnyj</i>	207
Retrieving collocational information from Japanese corpora: its methods and the notion of “circumcollocate” <i>Tadaharu Tanomura</i>	213
Diachrony of noun-phrases in specialized corpora <i>Nicolas Turenne</i>	223
Subject index	237
Author index	243
Authors’ addresses	247