Studies in quantitative linguistics

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


Gedruckt mit Unterstützung der Karl-Franzens-Universität Graz

ISBN: 978-3-942303-07-1

© Copyright 2011 by RAM-Verlag, D-58515 Lüdenscheid

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

This volume contains 14 contributions from international scholars from eight countries (Germany, India, Austria, Rumania, Slovakia, Czech Republic, Ukraine, USA). All contributions have their methodological background in quantitative linguistics, although with some overlap to other branches of linguistics like corpus linguistics, cognitive linguistics, and psycholinguistics.

In the first section, “Lexical Semantics”, the contributions are devoted to the statistical analysis of collocations (substantives plus modal verbs) in German, to selected problems of the polysemy of semantic fields in English, to the diachronic analysis of semantic concepts in English, and to the statistical analysis of concept structures by means of associative analysis. The common background of these contributions is an attempt at analyzing concepts of cognitive linguistics and cross-cultural linguistics from a statistical and quantitative point of view. All in all, the contributions give a good overview of the methodological and theoretical state of the art in this particular branch of linguistics. This work is a starting point for the formulation of some general deductive hypotheses as usually required in quantitative linguistics.

The second section, “Text and script analysis”, contains papers about the frequency structure of texts, based on the analysis of word form frequencies. On the one hand some text indicators are tested empirically, and on the other hand some new quantitative indicators of the thematic concentration in texts are presented. Syntactical issues are discussed on the level of sentence lengths (with a focus on data homogeneity) and selected problems of context comparison are analyzed by the means of combinatorial methods. Two papers are devoted to research of phonetics/semantics (sound symbolism) and to the quantitative analysis of an African script.

The third section, “Words and word classes”, contains papers devoted to some morphological and morphosyntactical problems. One paper is a continuation of works within the Göttingen Project on Length Distribution in Linguistics and focuses on word length frequencies in Norwegian. Another paper discusses selected problems of a quantitative classification of part-of-speech systems. Furthermore, German substantive compounds are analyzed with regard to their frequency of occurrence and their combinations with some lexical classes. Finally, one further paper is devoted to the morphological and phonological word structure in Ukrainian and Indonesian, where an interpretation of the parameters of Menzerath’s law is offered.

As a whole, the omnibus volume gives a good overview of the recent development in selected branches of quantitative linguistics and quantitative text analysis. We would like to thank our contributors for their effectiveness and productiveness in the process of editing this second omnibus volume of the “Studies in Quantitative Linguistics”.

Emmerich Kelih
Contents

Preface 1

Lexical semantics

Yesypenko Nadia
Cognitive semantics and the corpus-based approach to concept structure 1

Soloviova, Olga
Diachronic investigation of the semantic field of ‘size’ in English 22

Drebet, Viktor
Der Gebrauch der Substantive mit Modalverben in gegenwärtigen deutschsprachigen Prosawerken 39

Kantemir, Sergej
Vergleichende Analyse der polysemen Mikrostrukturen von Land – Volk – Staat und zemlia – narod – derzhava in der deutschen und der ukrainischen Presse 57

Text and script analysis

Čech, Radek
Frequency structure of New Year’s presidential speeches in Czech. The authorship analysis 75

Inamdar, Atul S.; Prabhu-Ajgaonkar, S.G.
A stylostatistical study using the sentence length criterion 88

Lvova, Nadija L.
Determining phonetic symbolism of the text 94

Milička, Jiří
A combinatorial method for a context comparison 104

Popescu, Ioan-Iovitz; Altmann, Gabriel
Thematic concentration in texts 110

Rovenchak, Andrij; Riley, Charles; Sherman, Tombekai
Complexity of the Vai script revisited: a frequency study of the syllabary 117
Words and word classes

Best, Karl-Heinz
Wortlängen im Norwegischen 129

Mačutek, Ján; Rovenchak, Andrij
Canonical word forms: Menzerath–Altmann law, phonemic length and syllabic length 136

Matskulyak, Yuliya
Semantische Besonderheiten der Komponenten in Komposita mit dem Modell N + N 148

Vulanović, Relja
Classifying parts-of-speech systems by their quantitative properties 170