Glottometrics ist eine unregelmäßig erscheinende Zeitchrift (2-3 Ausgaben pro Jahr) für die quantitative Erforschung von Sprache und Text.

Beiträge in Deutsch oder Englisch sollten an einen der Herausgeber in einem gängigen Textverarbeitungssystem (vorrangig WORD) geschickt werden. Glottometrics kann aus dem Internet heruntergeladen, auf CD-ROM (in PDF Format) oder in Buchform bestellt werden.

Herausgeber – Editors

G. Altmann
Univ. Bochum (Germany)
ram-verlag@t-online.de

K.-H. Best
Univ. Göttingen (Germany)
kbest@gwdg.de

R. Čech
Univ. Ostrava (Czech Republic)
cechradek@gmail.com

F. Fan
Univ. Dalian (China)
Fanfengxiang@yahoo.com

P. Grzybek
Univ. Graz (Austria)
peter.grzybek@uni-graz.at

E. Kelih
Univ. Vienna (Austria)
emmerich.kelih@univie.ac.at

R. Köhler
Univ. Trier (Germany)
koehler@uni-trier.de

H. Liu
Univ. Zhejiang (China)
lhtzju@gmail.com

J. Mačutek
Univ. Bratislava (Slovakia)
jmacutek@yahoo.com

G. Wimmer
Univ. Bratislava (Slovakia)
wimmer@mat.savba.sk

P. Zörnig
Univ. Brasilia (Brazilia)
peter@unb.br

External academic peers for Glottometrics

Prof. Dr. Haruko Sanada
Rissho University, Tokyo, Japan (http://www.ris.ac.jp/en/);
Link to Prof. Dr. Sanada: http://researchmap.jp/read0128740/?lang=english;
mailto:hsanada@ris.ac.jp

Prof. Dr. Thorsten Roelcke
TU Berlin, Berlin, Germany (http://www.tu-berlin.de/)
Link to Prof. Dr. Roelcke: http://www.daf.tu-berlin.de/menne/deutsch_als_fremd
_und_fachsprache/personal/professoren_und_pds/prof_dr_thorsten_roelcke/
mailto:Thorsten Roelcke (roelcke@tu-berlin.de)

Bestellungen der CD-ROM oder der gedruckten Form sind zu richten an
Orders for CD-ROM or printed copies to RAM-Verlag RAM-Verlag@t-online.de

Herunterladen/ Downloading: https://www.ram-verlag.eu/journals-e-journals/glottometrics/

Die Deutsche Bibliothek – CIP-Einheitsaufnahme
Diese elektronische Ressource ist im Internet (Open Access) unter der Adresse

Bibliographische Deskription nach 37 (2017) ISSN 2625-8226
Abstracts

Ramon Ferrer-i-Cancho
Random crossings in dependency trees

Abstract. It has been hypothesized that the rather small number of crossings in real syntactic dependency trees is a side-effect of pressure for dependency length minimization. Here we answer a related important research question: what would be the expected number of crossings if the natural order of a sentence was lost and replaced by a random ordering? We show that this number depends only on the number of vertices of the dependency tree (the sentence length) and the second moment about zero of vertex degrees. The expected number of crossings is minimum for a star tree (crossings are impossible) and maximum for a linear tree (the number of crossings is of the order of the square of the sequence length).

Jianwei Yan, Siqi Liu
The distribution of dependency relations in Great Expectations and Jane Eyre

Abstract. This study explores features of specific literary works, Great Expectations and Jane Eyre (hereafter referred to as GE and JE), based on the theoretical framework of dependency grammar. Both works are masterpieces of critical realism in Victorian era. This study is a descriptive analysis, which investigates the dependency relations of both works, including dependency distance, dependency direction and dependency type. The results indicate that: 1) The difference of syntactic difficulty between GE and JE is not statistically significant in accordance with MDDs (mean dependency distances); 2) There is a similar trend in the distribution of ADDs (absolute dependency distances), but the differences between GE and JE in ADDs are highly significant; 3) there is no significant difference in the distribution of dependency directions between GE and JE; 4) Both GE and JE have forty-three same dependency types; Meanwhile, although the differences of the distribution of dependency types are highly significant, there is no significant difference between GE and JE in MDDs of dependency types.

Kateřina Pelegrinová, Gabriel Altmann
The study of adverbials in Czech

Abstract. Each well defined linguistic concept can be studied quantitatively. Though this way has no end, one must perform the study stepwise. Here we analyze the behavior of adverbs and adverbial expressions and apply the models to Czech texts. The adverbials are classified in 13 classes and we study the class size, the length in individual classes, the placing of adverbials, the runs of left and right adverbials and the gaps between right adverbials. Further problems are sketched in the Introduction.

Hans J. Holm
Steppe homeland of Indo-Europeans favored by a Bayesian approach with revised data and processing

Abstract. Despite dozens of hypotheses, the origin and development of the Indo-European language family are still under debate. A well-known glottochronological approach to this problem using Bayesian computation of language divergence dates claims to have provided evidence for the period of Neolithic expansion known as the
“Anatolian hypothesis.” The dates have met with considerable criticism from other disciplines. I decided to investigate the evidence for these dates by replicating and analyzing the approach. During this process, a further approach located a date of origin from between 3950 – 4740 BC. One of the insights of this study was that previous results were significantly disrupted by poorly attested languages, which were consistently removed step by step.

This paper supports this finding using data from the previous approaches and my own updated dataset. The resulting date is around 4800 BC. However, the topology of the trees differed considerably over the course of several hundreds of tests. This problem was avoided in previous approaches by rigorous topological forcing. Here we apply a west–east dichotomy from a previous purely lexicostatistical (i.e. without times) approach based on the best available Indo-European dataset of approx. 1,100 verbal roots, which produces dates around 4100 BC. These dates reflect the most recent state of knowledge in linguistics, archeology and genetics in favor of the Steppe hypothesis. A new synopsis of the wheel problem, a primary argument for the divergence date, shows that not one but three different Indo-European denotations coincide in different areas with different types of wheel–axle constructions. Archeological cultures likely to have been affected by the migrations are presented visually at the end of this paper.

**Poiret Rafael, Haitao Liu**

Mastering the measurement of text’s frequency structure: an investigation on Lambda’s reliability

**Abstract.** Lambda is a measure of frequency structure that has been presented to be independent of text size (Popescu, Čech & Altmann, 2011). We demonstrate in this study that Lambda is obviously dependent on text size, confirming the findings of Čech (2015). Based on the assumption that Lambda was independent of text size, Popescu, Čech & Altmann (2011) investigated its capacity to detect text genre. We find that Lambda is still able to distinguish genres, but only very different ones. We also propose an experimental method based on Chinese to observe if Lambda is really able to measure the degree of analytism/synthetism of a text (Popescu, Čech & Altmann, 2011). We find that this method is promising. Moreover, our results corroborate with the assumption that Lambda has this property.

**Book Reviews**

**Kelih, Emmerich,** *Phonologische Diversität–Wechselbeziehungen zwischen Phonologie, Morphologie und Syntax.*

Frankfurt am Main: Peter LangVerlag 2016, 272 pages.

Reviewed by Gabriel Altmann

**Liu, Haitao & Liang, Junying** (eds.), *Motifs in Language and Text.*


Reviewed by Hanna Gnatchuk