

Contents of Glottometrics 12, 2006 (including abstracts)

Shizuo Mizutani

Fan, Fengxiang

Models for dynamic inter-textual type-token relationship **1-10**

Abstract. This paper examines the inter-textual type-token relationship and tests some existing quantitative models describing the vocabulary size and text length relationship. 8,334 samples were randomly drawn from the British National Corpus, totaling 8,001,000 tokens. The result shows that the models by Herdan and Brunet can capture the dynamic inter-textual type-token relationship, and the latter is also robust in extrapolation.

Peust, Carsten

Script complexity revisited **11-15**

Abstract: A simple method for quantifying the complexity of graphical signs is suggested. The complexity is defined as the number of crossing points which can maximally be achieved with an overlapping straight line. In signs composed of several disconnected elements, the complexity is to be computed for each element separately. This method is compared with another complexity measure serving a similar purpose recently proposed by Altmann.

Jayaram, B.D.; Vidya, M.N.

Word length distribution in Indian languages **16-38**

Abstract. The paper investigates Indian Language pattern with respect to word length distribution. The languages selected are Assamese, Marathi and Hindi belonging to Indo-Aryan family and Kannada and Tamil belonging to Dravidian Family. It examines data of different registers like Aesthetics, Social Science, Natural, Physical and Professional Sciences, Official and Media languages and Translated Material. It is observed that no single distribution fits across languages while across registers a single distribution fits majority of samples.

Best, Karl-Heinz

Gesetzmäßigkeiten im Erstspracherwerb **39-54**

Abstract. Language acquisition abides by laws. These laws seem to be the same as in language change and text production. If one considers the stages of language acquisition separately then there are further models controlling distributions, rank orders and processes. This paper yields some evidence for these laws.

Dzurjuk, Tetjana

Sentence length as a feature of style (applied to works of German writers)

55-62

Abstract. Sentence length in German is studied in three formal length categories and three genres. We try to characterize the individual writer by means of a vector of properties which can be used both for classification and the study of development.

Yokoyama, Shoichi; Wada, Yukiko

A logistic regression model of variant preference in Japanese kanji: an integration of mere exposure effect and generalized matching law

63-74

Abstract. The word *hinoki* or ‘cypress’ can be transcribed in two variant forms, 檜 (the so-called “traditional” variant) and 桧 (the “simplified” variant), in Japanese kanji. Such variant forms are called *kanji variants*. The present paper reviews a series of studies on Japanese kanji recognition (Yokoyama, 2006a, 2006b, 2006c), and proposes a logistic regression model which accounts for performance in a preference judgment task based on kanji frequency data. Yokoyama (2006a) administers preference and familiarity judgment tasks in which the participants were presented with 263 pairs of traditional and simplified variants and asked to choose the more preferable or familiar variant of each pair. The analyses indicate a positive contribution of frequency to variant preferences, supporting the so-called “mere exposure effect” theory of Zajonc (1968). This finding leads to a logistic regression model that describes preference behavior in kanji recognition, based on Fechner’s law. Yokoyama (2006b) shows that the model is comparable to the so-called “generalized matching law” of Baum (1974) and to the “ideal free distribution theory” of Fagen (1987). Yokoyama (2006c) further examines the predictive validity of the model with empirical data obtained from a preference judgment task, administered in the Tokyo and Kyoto areas. Logistic regression analyses are performed with the ratio of preference for the given variants and the logit of the character frequencies, yielding significant correlations between the predicted probabilities and the observed responses ($r = .804$ for Asahi newspaper data). The present paper synthesizes these studies and proposes a logistic regression model that efficiently describes preference behavior in Japanese kanji recognition, integrating the theoretical perspectives of the mere exposure effect and the generalized matching law.

History of Quantitative Linguistics**Best, Karl-Heinz**

XV. Jean Paul (1763-1825)

75-77**Best, Karl-Heinz**

XVI. Ernst Wilhelm Förstemann (1822-1906)

77-86**Best, Karl-Heinz**

XVII. Karl Knauer (1906-1966)

86-94**Best, Karl-Heinz**

XVIII. August Friedrich Pott (1802-1887)

94-96

Book reviews

A.A. Polikarpov, G.G. Sil'nickij, V.V. Poddubnyj (eds.), *Kvantitativnaja Lingvistika: Issledovanija i modeli* (Klim-2005). Materialy Vserossijskoj naučnoj konferencii (6-10 ijunja 2005 g.). Novosibirsk: Novosibirskij Go sudarstvennyj Pedagogičeskij Universitet. Reviewed by **Emmerich Kelih** 97-106

G. Altmann, V. Levickij, V. Perebyinis (eds.). *Problemy kvantitativnoji lingvistyky: zbirnyk naukovych pracj* (Problems of Quantitative Linguistics). Černivci: Ruta, 2005. 352 S. Von **Juri Kijko** 106-108

Glottometrics ist eine unregelmäßig erscheinende Zeitschrift 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** (PDF-Format) oder in **Buchform** bestellt werden.

Glottometrics is a scientific journal for the quantitative research on language and text published at irregular intervals

Contributions in English or German written with a common text processing system (preferably WORD) should be sent to one of the editors

Glottometrics can be downloaded from the **Internet**, obtained on **CD-ROM** (in PDF) or in form of **printed copies**

Herausgeber/Editors:

G. Altmann	02351973070-0001@t-online.de
K.-H. Best	kbest@gwdg.de
P. Grzybek	peter.grzybek@uni-graz.at
A. Hardie	a.hardie@lancaster.ac.uk
L. Hrebicek	hrebicek@orient.cas.cz
R. Köhler	koehler@uni-trier.de
V. Kromer	kromer@newmail.ru
O. Rottmann	otto.rottmann@t-online.de
A. Schulz	reuter.schulz@t-online.de
G. Wimmer	wimmer@mat.savba.sk
A. Ziegler	arneziegler@compuserve.de

Herunterladen/ Downloading: <http://www.ram-verlag.de>

ISSN 1617-8351