

# Table of contents

 <https://doi.org/10.1075/cilt.356.toc>

Pages v–vi of

## **Language and Text: Data, models, information and applications**

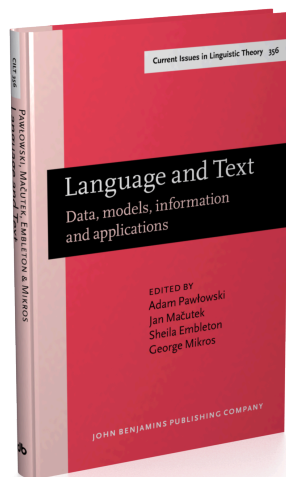
**Edited by Adam Pawłowski, Jan Mačutek, Sheila Embleton and George Mikros**

[*Current Issues in Linguistic Theory*, 356] 2021. vi, 280 pp.

© John Benjamins Publishing Company

This electronic file may not be altered in any way. For any reuse of this material written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

For further information, please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website at [benjamins.com/rights](http://benjamins.com/rights)



# Table of contents

Introduction	1
<i>Adam Pawłowski, Sheila Embleton, Jan Mačutek and George Mikros</i>	
<b>Part I. Theory and models</b>	
On the impact of the initial phrase length on the position of enclitics in Old Czech	9
<i>Radek Čech, Pavel Kosek, Olga Navrátilová and Ján Mačutek</i>	
Term distance, frequency and collocations	21
<i>Lars G. Johnsen</i>	
A method for the comparison of general sequences via type-token ratio	37
<i>Vladimír Matlach, Diego Gabriel Krivochen and Jiří Milička</i>	
Quantitative analysis of syllable properties in Croatian, Serbian, Russian, and Ukrainian	55
<i>Biljana Rujević, Marija Kaplar, Sebastijan Kaplar, Ranka Stanković, Ivan Obradović and Ján Mačutek</i>	
N-grams of grammatical functions and their significant order in the Japanese clause	69
<i>Haruko Sanada</i>	
Linking the dependents: Quantitative-linguistic hypotheses on valency	93
<i>Petra Steiner</i>	
Grammar efficiency and the One-Meaning–One-Form Principle	109
<i>Relja Vulcanović</i>	
Distribution and characteristics of commonly used words across different texts in Japanese	121
<i>Makoto Yamazaki</i>	

**Part II. Empirical studies**

The perils of big data <i>Sheila Embleton, Dorin Uritescu and Eric S. Wheeler</i>	137
From distinguishability to informativity: A quantitative text model for detecting random texts <i>Maxim Konca, Alexander Mehler, Daniel Baumartz and Wahed Hemati</i>	145
A Modern Greek readability tool: Development of evaluation methods <i>George Mikros and Rania Voskaki</i>	163
Phonological properties as predictors of text success <i>Jiří Milička and Alžběta Houzar Růžicková</i>	177
Calculating the victory chances: A stylometric insight into the 2018 Czech presidential election <i>Michal Místecký</i>	195
Topological mapping for visualisation of high-dimensional historical linguistic data <i>Hermann Moisl</i>	209
Book genre and author's gender recognition based on titles: The example of the bibliographic corpus of microtexts <i>Adam Pawłowski, Elżbieta Herden and Tomasz Walkowiak</i>	225
Quantitative analysis of bibliographic corpora: Statistical features, semantic profiles, word spectra <i>Adam Pawłowski, Krzysztof Topolski and Elżbieta Herden</i>	239
Analysis of English text genre classification based on dependency types <i>Yaqin Wang</i>	257
In memory of Gabriel Altmann: Eminent linguist, a man with a brilliant mind, and friend	271
Index	277