

Unified Modeling of Diversification

in Language

by

Gabriel Altmann

2018

RAM-Verlag

Studies in Quantitative Linguistics

Editors

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

1. U. Strauss, F. Fan, G. Altmann, *Problems in quantitative linguistics 1*. 2008, VIII + 134 pp.
2. V. Altmann, G. Altmann, *Anleitung zu quantitativen Textanalysen. Methoden und Anwendungen*. 2008, IV+193 pp.
3. I.-I. Popescu, J. Mačutek, G. Altmann, *Aspects of word frequencies*. 2009, IV +198 pp.
4. R. Köhler, G. Altmann, *Problems in quantitative linguistics 2*. 2009, VII + 142 pp.
5. R. Köhler (ed.), *Issues in Quantitative Linguistics*. 2009, VI + 205 pp.
6. A. Tuzzi, I.-I. Popescu, G. Altmann, *Quantitative aspects of Italian texts*. 2010, IV+161 pp.
7. F. Fan, Y. Deng, *Quantitative linguistic computing with Perl*. 2010, VIII + 205 pp.
8. I.-I. Popescu et al., *Vectors and codes of text*. 2010, III + 162 pp.
9. F. Fan, *Data processing and management for quantitative linguistics with Foxpro*. 2010, V + 233 pp.
10. I.-I. Popescu, R. Čech, G. Altmann, *The lambda-structure of texts*. 2011, II + 181 pp.
11. E. Kelih et al. (eds.), *Issues in Quantitative Linguistics Vol. 2*. 2011, IV + 188 pp.
12. R. Čech, G. Altmann, *Problems in quantitative linguistics 3*. 2011, VI + 168 pp.
13. R. Köhler, G. Altmann (eds.), *Issues in Quantitative Linguistics Vol 3*. 2013, IV + 403 pp.
14. R. Köhler, G. Altmann, *Problems in Quantitative Linguistics Vol. 4*. 2014, VIII+148 pp.
15. Best, K.-H., Kelih, E. (eds.), *Entlehnungen und Fremdwörter: Quantitative Aspekte*. 2014. VI + 163 pp.
16. I.-I. Popescu, K.-H. Best, G. Altmann, *Unified Modeling of Length in Language*. 2014, VIII + 123 pp.
17. G. Altmann, R. Čech, J. Mačutek, L. Uhlířová (eds.), *Empirical Approaches to Text and Language Analysis dedicated to Luděk Hřebíček on the occasion of his 80th birthday*. 2014. VI + 231 pp.
18. M. Kubát, V. Matlach, R. Čech, *QUITA Quantitative Index Text Analyzer*. 2014, VII + 106 pp.
19. K.-H. Best, *Studien zur Geschichte der Quantitativen Linguistik*. 2015. III + 158 pp.

20. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, P. Mohanty, E. Kelih, R. Chen, G. Altmann, *Descriptiveness, Activity and Nominality in Formalized Text Sequences*. 2015. IV+120 pp.
21. G. Altmann, *Problems in Quantitative Linguistics Vol. 5*. 2015. III+146 pp.
22. P. Zörnig, K. Stachowski, I.-I. Popescu, T. Mosavi Miangah, R. Chen, G. Altmann, *Positional Occurrences in Texts: Weighted Consensus Strings*. 2015. II+178 pp.
23. E. Kelih, R. Knight, J. Mačutek, A. Wilson (eds.), *Issues in Quantitative Linguistics Vol. 4*. 2016. III + 231 pp.
24. J. Léon, S. Loiseau (eds.), *History of quantitative linguistics in France*. 2016. II + 232 pp.
25. K.-H. Best, O. Rottmann, *Quantitative Linguistics, an Invitation*. 2017, III+171 pp.
26. M. Lupea, M. Rukk, I.-I. Popescu, G. Altmann, *Some Properties of Rhyme*, 2017, IV + 134 pp.

ISBN: 978-3-942303-63-7

© Copyright 2018 by RAM-Verlag, D-58515 Lüdenscheid, Germany

RAM-Verlag
Stüttinghauser Ringstr. 44
D-58515 Lüdenscheid
Germany
RAM-Verlag@t-online.de
<http://ram-verlag.eu>

Preface

Quantification is a most natural development in all sciences. At the beginning, it is not very ripe but it is better to do something than to adhere to a tradition. One can distinguish several steps in quantification: (1) The use of quantitative concepts, beginning with decisions, and formalizations; then trying to ascribe to the phenomena some numbers or degrees; (2) measuring the phenomenon, especially properties. It must be noted that the properties – whatever they are – are our conceptual creations representing the reality “for us”; (3) expressing mathematically the fact found, and performing some kind of modeling; (4) testing the model on data – an enterprise that will never be finished; (5) joining the phenomenon with other phenomena, creating and deriving of hypotheses, and constructing a theory; and ultimately (6) the unification of the modeling and its simplification as far as possible.

We find all of these levels in current quantitative linguistics. In the present book we try to make a step in the last stage, namely to perform a simplification concerning classes of linguistic objects. We have collected some of the available published literature – it is enormous – and let the reader continue fitting the simple function we propose.

The approach does not mean that all the endeavours by Zipf, Mandelbrot, etc. are not correct. On the contrary, we merely bring to them a simple differential equation which is easier to interpret than the more complex ones, but all of them belong to the same unified theory. Neither does it mean that we have discovered more truth. Truth is not concealed behind formulae. They merely help us to capture the reality and to draw mechanically consequences using the rules of mathematics. But none of our models is part of the reality.

We hope that those who have many sets of data at their disposal will test the given proposal. If the data are bell shaped, the simple formula does not work. But one can add something to the differential equation and interpret it synergetically. We are persuaded that all relations in language are very simple. We do not know them but we try to capture them *viribus unitis*.

Gabriel Altmann

Contents

Preface

1.	Introduction	1
2.	Rhythm	6
	1.1. Hexamete	6
	2.2. Feet	9
3.	Parts of Speech	11
4.	Morphology	26
	4.1. Morphological complexity	31
	4.2. Derivation	32
	4.3. Cases	34
5.	Semantic Diversification	36
	5.1. Definition chains	44
	5.2. Semantic associations	46
	5.3. Pure polysemy	46
6.	Ranking	51
	6.1. Grammatical categories	53
	6.2. Adnominals	55
	6.3. Motifs of adnominals	61
	6.4. Syllable types	66
7.	Various Word Classes	67
	7.1. Special word frequencies	67
	7.2. Borrowings	68
	7.3. Parts of speech	71
	7.4. Verbs	72
	7.5. Adverbs	75
	7.6. Pronouns	77
	7.7. Individual word	78
	7.8. Personal names	78
8.	Compounds	80
9.	Syntax	85
	9.1. Sentence types	85
	9.2. Syntactic alternatives	87
	9.3. Semantic roles in dependency tree	88

10.	Environment	90
11.	Studies of Hrebs	91
12.	Belza Chains	97
13.	Word Frequencies	99
14.	Music	102
15	Conclusions	104
	References	105
	Register	115