

Problems in Quantitative Linguistics 1.

Strauss, Udo; Fan, F.; Altmann, G.

Publisher: Lüdenscheid: RAM-Verlag. [= Studies in Quantitative Linguistics, 1], 2008

Reviewer: Emmerich Kelih

Not long ago, the German publishing house RAM published the monograph *Problems in Quantitative Linguistics 1* by Udo Strauss, Fengxiang Fan and Gabriel Altmann. It is the first monograph within the newly established book series *Studies in Quantitative Linguistics*. This new book series is – besides the series *Quantitative Linguistics*¹, published by Mouton de Gruyter (Berlin, New York) – the most recent publication on the book market. In both cases the series are dealing exclusively with quantitative and synergetic linguistics.

The members of the editorial board are Fengxiang Fan (China), Peter Grzybek (Austria) und Ján Mačutek (Slovakia). The first volume gives no further information about the future contents and direction of the new series.

The book comprises nine chapters, an introduction as well as an author- and subject index. The introduction provides main information about the motivation and contents of the monograph. It is not a monograph in the common sense, but rather a compilation of recent problems and “open” questions in quantitative linguistics (QL). The target audiences are students and scientists, who are interested in a compact introduction to some recent problems in quantitative linguistics. The general recommendations given in the introduction – and this must be emphasized – are a profound mixture of epistemological, methodological, technical and practical procedures for empirical testing of quantitative-linguistic hypotheses. These recommendations are by no means rigid rules, but rather an attempt to unify the testing of QL-related problems.

For every general subject area within the nine chapters (phonemics and scripts, grammar, etc.) some current questions are discussed in more detail. For every single subtopic (e.g. accent and frequency, canonical word structure, consonant clusters, etc.) the most important hypothesis, including some references, is given. The hypothesis is followed by a paragraph on procedures, which include precise steps for future analyses (data acquisition, operationalization, appropriate methods, and adequate statistical models). By following these instructions it is very likely that “positive” solutions can be found. The mentioned procedures are mainly expressed as suggestions and their subsequent implementation offers an easier conceptual integration into the recent state of the art of QL. Needless to say, that no epistemological unification is intended by the proposed procedures, but rather a stimulation of scientific creativity. To sum up, *Problems in Quantitative Linguistics 1* could be used as a reference book.

Chapter one „Phonemics and Script“ deals with different problems of quantitative phonology and script analysis, e.g. models for phoneme frequencies, quantitative syllable structure, script complexity and simplification. The second chapter “Grammar” contains a discussion of the frequency in different areas (co-occurrence of words and cohesion, frequency of morphological irregularities). It is striking that mainly hypotheses from the “American Frequentism” (cf. Bybee/Hooper 2001) are mentioned, which have not been empirically verified or analyzed with the help of statistical/mathematical methods so far.

In the third chapter „Compounds and Lexicology“ a prominent position is given to the quantitative analysis of compounds. Definitely, there is no lack of relevant quantitative

¹ Up to now, the *Festschrift for Gabriel Altmann On the Occasion of His 75th Birthday* by Grzybek/Köhler (eds.) (2007) and the omnibus volume on quantitative script analysis by Altmann/Fengxiang (eds.) (2008) have been published. A monograph on word frequencies is in press (cf. Popescu et al. 2009).

hypotheses on compounds in general (e.g. the interrelation of composite-building and word-age, length of compounds and length of their constituents), but primarily a lack of empirical evidence. The fourth and most comprehensive chapter (“Textology”) includes more or less heterogeneous sub-areas of quantitative text analysis: problems of cotextuality and frequency, of Hrebs, of Hurst’s exponent and of so-called “word- and sentence length motives” (i.e. non-decreasing sequences of lengths, introduced by R. Köhler and S. Naumann).

The fifth chapter “Frequency and Length” is – among other well known problems like the length-frequencies of words and sentences – dedicated to some general interrelations between the frequency and the length of linguistic units. It focuses on the interdependency of letter frequency and letter utility, problems of “markedness” and the hardly analyzed relation of phoneme frequency and phoneme complexity.

The sixth chapter “Semantics, Synergetics and Psycholinguistics” includes some hypotheses about the frequency of polysemy, the degree of popularity of slang words, and some quantitative aspects of language acquisition. In this context it must be emphasized that the position of synergetic linguistics, which is mentioned in the title of this chapter, remains quite unclear: is linguistic synergetics a sub branch or a theoretical framework of QL?

The seventh chapter contains mainly questions on quantitative typology and a summing-up of more or less „well known“ hypotheses from the Prague School of Linguistics, that is mainly references on V. Skalička. Furthermore, some new and innovative indicators like the Hirsch-point are introduced. More detailed information on this will certainly be found in Popescu et al. (2009).

The eighth chapter covers different problems of QL, which obviously cannot be related to the so far mentioned areas, e.g. adequate frequency models, interrelation of entropy and size of inventories, Menzerath’s and Zipf’s law, etc. that is, the altogether more or less well known problems of QL, which are still waiting for a “final” synthesis (named by the authors “historical review”).

The last chapter can be understood as a motivation for the setting up of greater interdisciplinary research-projects (for example on the Frumkina law, Skalička’s typological system, synonymy, etc.), which are best to be analysed within a team of linguists and mathematicians. The monograph ends with a detailed author and subject index. Altogether the monograph fulfils all requirements to be an up-to-date and well arranged collection of problems on QL. Without a doubt it is a useful reference book for students and linguistics.

References

- Altmann, G.; Fengxiang, F. (eds.) (2008): *Analyses of Script. Properties of Characters and Writing Systems*. Berlin/New York. Mouton de Gruyter. [= Quantitative Linguistics 63]
- Bybee, J.; Hooper, P. (eds.) (2001). *Frequency and the emergence of linguistic structure*. Amsterdam/Philadelphia: John Benjamins. [= Typological Studies in Language 45]
- Grzybek, P.; Köhler, R. (eds.). *Exact Methods in the Study of Text and Language. Dedicated to Gabriel Altmann on the Occasion of his 75th Birthday*. Berlin/New York: Mouton de Gruyter. [= Quantitative Linguistics 62]
- Popescu, I.-I.; Altmann, G.; Grzybek, P.; Jayaram Bijapur D.; Köhler, R.; Krupa, V.; Mačutek, J.; Pustet, R.; Uhlřřov, L.; Vidya, Matummal N.; (2009). *Word frequency studies*. Berlin/New York: Mouton de Gruyter. [= Quantitative Linguistics 64]

Emmerich Kelih (University of Graz)
emmerich.kelih@uni-graz.at