

Janda, Laura A. (ed.) (2013) *Cognitive linguistics. The quantitative turn. The essential reader*. Berlin: De Gruyter Mouton (Mouton Reader). Pp. 321.

Reviewed by Emmerich Kelih

This volume comprises a number of studies which have previously been published in the journal *Cognitive Linguistics*. The articles, appearing from 2008 to 2012, were selected by editor Laura A. Janda to give a representative overview of the recent growing trend of using statistical methods within cognitive linguistics and related branches of linguistics. The omnibus selection, subtitled “The Essential Reader”, consists of ten selected papers by 19 authors and one extensive introductory chapter by the editor herself. It should be pointed out right away that the intent of my book review is not to present a comprehensive overview and review of the selected papers in detail, but rather to highlight some fundamental theoretical and methodological assumptions regarding the use of quantitative methods in cognitive linguistics and general linguistics as well. The avowed purpose of the collection, as stated in the introduction, is twofold: (a) to present papers and statistical methods that have played a leading role in defining the current state of the art in cognitive linguistics, focusing on statistical methods, and (b) to give the reader an essential overview of different but widely used statistical methods.

The introductory chapter by Laura A. Janda (pp. 1–32) gives an essential motivation and some background information on recent trends within particular branches of cognitive linguistics, which extensively make use of statistical and quantitative methods. The main reason for the suggested quantitative turn in cognitive linguistics according to the author is a combination of theoretical and historical factors. First of all cognitive linguistics utilizes a usage-based model of language structure and thus no fundamental distinction between “performance” and “competence” is posited. This position opens cognitive linguistics to systematic observations, empirical analysis and the measurement and quantitative interpretation of linguistics events and entities. Accepting introspection as an acceptable “method” of linguistics (p. 6) Laura A. Janda nonetheless highlights the absolute scientific necessity of claiming hypotheses which can be operationalized and falsified. In this respect recent developments in

some branches of linguistics seem to facilitate the use of statistical methods: namely a remarkable progress in computational linguistics (especially technological applications), the extensive technical and methodological progress in corpus linguistics and the development of multi-purpose corpora for many languages. Finally the availability of free statistical software (in particular *R*) and of relevant and highly accessible handbooks of statistical methods in linguistics accelerates the use and application of statistical methods in cognitive linguistics as well as in general linguistics. Furthermore, the introductory chapter gives a comprehensive overview of different statistical methods used by the authors of the selected papers in the main part of the omnibus volume. The range of presented methods stretches from descriptive and analytical statistics to elaborated multivariate statistical methods. It includes the very common and popular chi-square test (which is largely used in general and cognitive linguistics), the Fisher test, the exact binomial test, the t-test and ANOVA (analysis of variance), general correlation and regression models, mixed effects models and cluster analysis. Finally Laura A. Janda (p. 26) mentions some alternatives to particular regression models, such as methods of multidimensional scaling and correspondence analysis, which are quite common in linguistics but not represented in the omnibus volume. Thus it appears that the volume under review gives a representative and sophisticated overview of widely and commonly used statistical methods in linguistics.

## 1. Language acquisition

Let us now turn to the individual chapters of the volume. The emphasis of the review is not on the methods used (they are mentioned above in detail), but rather on the linguistic hypotheses and problems which are analysed by means of quantitative methods. There are two studies dealing with preemption in child language acquisition. A. Stefanowitsch (pp. 33–55) gives corpus-based data for alternating verbs which can appear in both ditransitive and prepositional constructions. Furthermore he tests whether factors which are relevant for the information structure of these verbs (for instance “givenness” is understood as referential distance, “syntactic weight” is measured by the number of orthographic words and “animacy” within one nominal phrase) play a role in the choice between the prepositional dative and the ditransitive. Subsequent to this contribution A. E. Goldberg (pp. 57–79) argues in favour of taking into consideration the probability of

these constructions and analysing whether they are actually in competition or not. The main result (again based on corpus data) shows that non-alternating verbs more likely occur in prepositional constructions, whereas the probability for alternating verbs is much lower.

Problems of the developmental processes in child language acquisition are discussed by A.L. Theakston, R. Maslen, E. Lieven and M. Tomasello (pp. 103–140), who present data of the usage (12-month sample from age 2.0 to 3.0) of transitive subject–verb–object (SVO) constructions of a child. These data are partly compared with the input of his mother. It is shown that the acquisition of these constructions seems to be a complex process, one which involves different stages of acquisition and a significant development within 12 months. Methodologically the study mainly relies on relative and absolute frequencies and the use of the chi-square test. There is no statistical modelling of the developmental data of the successive acquiring of SVO constructions, which would give further information about a law-like increase of the usage of these constructions.

E. Dąbrowska, C. Rowland and A. Theakston (pp. 197–223) address a different kind of issue, namely the acquisition of questions and constructions with long-distance dependencies in child language acquisition. It appears that in addition to a rather late acquisition of these complex constructions (after age six) it can be shown that (a) prototypical declaratives are for children the easiest construction type to imitate, (b) prototypical sentences are easier to repeat for children because they contain higher-frequency verbs and (c) adults show the prototypicality effect for questions only. The application of ANOVA (analysis of variance) seems to be appropriate, since the authors are comparing the means and variance of a group of variables.

## 2. Psycholinguistics and corpus linguistics

In the article by M. Johansson Falck and R.W. Gibbs (pp. 81–102), differences between the use of the English lexemes *path* and *road* are explored in detail. Based on experimental data, participants are required to answer specific questions about the mental image of the mentioned lexemes. In addition to these psycholinguistic surveys corpus data are represented that confirm a higher degree for metaphorical uses of *path* than of *road*. The analysis is accompanied by the use of the chi-square test.

The paper by B. Hampe (pp. 141–175) is devoted to problems of the transitive argument structure in English. Based on the concept of

“collostruction” she measures the strength of the co-occurrence of selected lexemes in transitive constructions. The analysis, which is methodologically based on a Fisher/Yates exact test, uses collostruction patterns taken from the International Corpus of English and focuses in particular on NP complements in transitive constructions. In this respect this approach can be understood as a suitable way for discovering new syntactical patterns.

The phonological similarity in multi-word units is the focus of the paper by S. Th. Gries (pp. 177–196). The idea behind it is to explore the number of phonologically similar elements in particular positions of multi-word units which have a different degree of idiomaticity. Particularly it is shown that alliterations generally play a significant role in forming idiomatic expressions. The occurrence of alliterations (explored in slightly different settings) is thus generally higher (attested by using the exact binomial test) than expected by chance.

H. Diessel (pp. 225–250) discusses general problems of iconicity as a governing principle of forming syntactical constructions. The constituent order of adverbial clauses is explored in detail. By using logistic regression analysis, the author investigates step by step the effects of various factors (length of clauses, complexity and pragmatic import) of the structuring of syntactical constructions. Generally H. Diessel’s investigations (samples are taken from the International Corpus of English) are convincing examples of the determining function of iconicity in the positioning of temporal adverbial clauses.

Problems of sociolinguistics and quantitative loanword research are tackled by E. Zenner, D. Speelman and D. Geeraerts (pp. 251–294), who rely heavily on corpus data to explore the empirical extent of Anglicisms within particular languages. The study introduces new methods and concepts of measuring the variation, the use and the success of English loanwords in Dutch. Methodologically this approach is of a general interest, since the authors use mixed effect models (recently becoming more and more common in linguistics), which can be useful for the extraction of fixed and random effects. As factors which can determine the success rate of loanwords the authors propose (a) the relative length of loanwords, (b) the affiliation to a particular lexical field, (c) the time-age of borrowing, (d) the kind of loanwords (luxury vs. necessary borrowing) and (e) the concept frequency. Generally the proposed way of thinking and the seeking of interrelations of linguistic properties reminds one of synergetic

linguistics approaches (cf. Köhler 2005), which unfortunately are not mentioned by the authors.

Finally, with L. A. Janda and V. D. Solovyev's chapter (pp. 295–321) on constructional profiles of near-synonymy lexemes in Russian we meet the application of the chi-square test and hierarchical cluster analysis. The similarity of lexemes is explored by the means of the frequency of constructions in which particular lexemes (such as Russian nouns for sadness and their synonyms *pečal'*, *toska*, *chandra*, *melancholija*, *grust'* and *unynie*) appear. In this respect the similarity of near-synonyms can be tested mathematically, which in turn allows for a comparison between the constructional profiles and the grouping in synonymy dictionaries.

Although some of the articles in the reviewed omnibus volume will appeal more to the specialist of particular problems in cognitive linguistics, others will evoke a broader interest in linguistics. Above and beyond this, the volume presents a useful overview and picture of the varied kinds of application of quantitative methods in cognitive linguistics, psycholinguistics and corpus linguistics. Thus, all in all, this volume does fulfil its intention, and will give the interested reader easy access to the state of the art of quantitative methods in cognitive linguistics.

## References

- Köhler, Reinhard (2005): Synergetic linguistics. In: Reinhard Köhler, Gabriel Altmann and Rajmund G. Piotrowski (eds.) *Quantitative Linguistik. Quantitative Linguistics. Ein internationales Handbuch. An International Handbook*, pp. 760–774. Handbücher zur Sprach- und Kommunikationswissenschaft 27. Berlin/New York: de Gruyter.

Contact information:

Emmerich Kelih  
Institute for Slavic Studies  
University of Vienna  
Spitalgasse 2, Hof 3  
1090 Wien  
Austria  
e-mail: emmerich(dot)kelih(at)univie(dot)ac(dot)at